

Université de Montréal

**Évaluation de la fidélité des interventions en santé publique
dans le cadre des essais randomisés en grappes dans les pays
du Sud : revue systématique et étude de cas**

Par

Myriam Cielo Pérez Osorio

Département de médecine sociale et préventive
École de Santé Publique de l'Université de Montréal

Thèse présentée en vue de l'obtention du grade de Philosophiae Doctor - Doctorat (Ph.D.)
en Santé publique, option Santé mondiale

Septembre 2020

© Myriam Cielo Pérez, 2020

Université de Montréal
Département de médecine sociale et préventive
École de santé publique de l'Université de Montréal

Cette thèse intitulée

**Évaluation de la fidélité des interventions en santé publique
dans le cadre des essais randomisés en grappes dans les pays
du Sud : revue systématique et étude de cas**

présentée par

Myriam Cielo Pérez Osorio

a été évaluée par un jury composé des personnes suivantes

Béatrice Godard
Présidente-rapporteuse

Dre Mira Johri
Directrice de recherche

Dr Valéry Ridde
Codirecteur de recherche

Aude Motulsky
Membre du jury

Yaya Sanni
Examineur externe

Résumé

La santé publique fondée sur des données probantes doit être basée sur les meilleures preuves disponibles pour prendre des décisions éclairées, afin de mettre en place des interventions dirigées vers le maintien et l'amélioration de la santé, ainsi que vers le bien-être de toute la population. Les essais contrôlés randomisés (ECR) sont souvent utilisés en recherche clinique pour tester les effets d'un médicament, d'une thérapie ou d'une intervention sur un groupe expérimental qui bénéficiera de l'intervention, en le comparant à un groupe contrôle qui recevra un placebo ou aucun traitement. Bien que le débat persiste, les essais randomisés constituent une source importante et, apparemment, de haute qualité pour évaluer l'efficacité des interventions en santé. Dû à de multiples facteurs, les essais randomisés en grappes (ERG) sont largement utilisés pour évaluer la prestation des services de santé et des interventions en santé publique. Dans ce type d'essai, ce ne sont plus des individus qui sont randomisés, mais des groupes d'individus tels que les familles, les médecins, les villages qui vont recevoir l'intervention. Ces interventions peuvent varier pendant la mise en œuvre en raison de divers facteurs liés à la conception de l'intervention, aux participants, aux intervenants ainsi qu'aux facteurs du contexte qui influencent les résultats. Ces facteurs doivent être pris en compte au moment de l'évaluation, et avant la réplique dans d'autres contextes. L'évaluation de la fidélité de la mise en œuvre, outil clé de l'évaluation du processus et élément essentiel du processus de mise à l'échelle, vise à mesurer le degré selon lequel une intervention a été implantée telle que conçue par les concepteurs.

Cette thèse a comme objectif principal examiner la fidélité de la mise en œuvre des interventions en santé publique dans le cadre des essais randomisés en grappes, pour savoir si les interventions mises en place sous un modèle contrôlé doivent prendre en compte ce type d'évaluation pour renforcer ces résultats et faciliter leur réplique à grande échelle. Cette thèse comporte deux volets : une revue systématique et une étude de cas unique à trois unités d'analyse selon une approche mixte concomitante. Le premier article évalue la pratique de la fidélité de la mise en œuvre des interventions en santé publique dans le cadre des essais randomisés en grappes des études publiées qui ont été identifiées et incluses dans la révision systématique. La révision systématique met en lumière que les interventions mises en place sous ce modèle ne tiennent pas

compte de cette évaluation de façon systématique, que la façon de la faire est très hétérogène, et que l'évaluation n'est pas bien documentée. Les deuxième et troisième articles sont les résultats de recherche de l'évaluation d'une intervention, à travers une étude de cas comme méthode de recherche, qui a été menée, dans un premier temps, pour examiner la plausibilité de la théorie de l'intervention, et, dans un deuxième temps, pour évaluer leur fidélité de la mise en œuvre et leur acceptabilité auprès des participants dans le but de l'améliorer, si nécessaire, avant sa mise en place à grande échelle. L'évaluation de l'intervention met en lumière plusieurs aspects. D'abord, la théorie sous-jacente et le modèle de l'intervention évaluée sont bien conçus pour parvenir aux résultats visés. L'évaluation fournit des points clés et des actions à prendre en considération, pendant le développement des interventions, pour servir les communautés difficiles à atteindre, et pour améliorer les résultats en matière de santé. Ensuite, les résultats ont démontré une fidélité de mise en œuvre élevée. La clarté de la théorie de l'intervention, la motivation et l'engagement des intervenants, ainsi que les réunions périodiques des superviseurs avec les intervenants-terrain expliquent largement le haut niveau de fidélité obtenu. Des facteurs contextuels tels que la distance géographique, l'accès à un téléphone portable, le niveau d'éducation et les normes de genre ont contribué à l'hétérogénéité de la participation du groupe cible de l'intervention. Finalement, cette évaluation souligne que la plateforme mobile combinée à la mobilisation communautaire, composantes clés de l'intervention, ont été bien accueillies par les participants, et pourraient être mis en place à grande échelle.

Cette thèse contribue au développement des connaissances sur le plan méthodologique concernant l'évaluation de la fidélité de la mise en œuvre des interventions en santé publique en mettant en relief des lacunes dans ce domaine, et en suggérant un outil pour faire avancer cette pratique évaluative. Cette thèse participe également au renforcement de la recherche dans les sciences de l'implémentation, et apporte sur le plan empirique des éléments clés essentiels pour évaluer la fidélité de la mise en œuvre de ce type d'intervention à l'aide des essais randomisés en grappes, évaluation de cette fidélité qui est l'objet de cette recherche doctorale.

Mots-clés : Intervention en santé publique; essais randomisés en grappes; revue systématique; évaluation de la fidélité; évaluation du processus; santé mondiale; pays du Sud.

Abstract

Evidence-based public health should be based on the best available evidence to make informed decisions and to implement interventions aimed at maintaining and improving the health and well-being of all people. Randomized controlled trials (RCTs) are often used in clinical research to test the effects of a drug, therapy, or intervention on an experimental group that may benefit from the intervention, comparing it to a control group that received either a placebo or no intervention treatment. Although the debate persists, randomized controlled trials are an important and objectively high quality method for evaluating the effectiveness of health interventions. Due to multiple factors, cluster randomized trials (CRTs) are widely used to assess the delivery of health services and public health interventions. In this type of trial, it is no longer individuals who are randomized, but groups of individuals such as families, doctors, and village communities who receive the intervention. These interventions may differ during implementation as a result of various factors related to the complexity of the intervention design, context, participants, and stakeholders involved. These factors should be considered at the time of assessment and before replication in other contexts. Implementation fidelity assessment, a key tool in process evaluation, examines study processes to assess the extent to which the intervention was carried out as originally intended. The fidelity of implementation is an essential part of the scale-up process.

This thesis aimed to examine the fidelity of implementation of public health interventions in the context of cluster randomized trials, to determine whether the interventions implemented under a controlled model should consider this type of evaluation to strengthen their results and facilitate their replication on a large scale. This thesis has two parts: a systematic review and a single case study with three units of analysis using a mixed triangulated approach. The first article assessed the implementation fidelity of public health interventions in the context of cluster randomized trials. The systematic review highlighted the finding that public health interventions implemented under this model did not systematically consider this type of evaluation, that the way of doing it was very heterogeneous, and that the evaluation was not adequately documented. The second and third articles were the research findings of the evaluation of an intervention, using a case study as the research method, that was conducted to first examine the plausibility of the

intervention theory and to better understand the design and context of the intervention being evaluated, and second, to evaluate implementation fidelity and its acceptability among the participants with the aim of making improvements (if necessary) before large-scale replication. The evaluation of the case study highlighted several key findings. First, the results of the evaluation reflected that the underlying theory and model of the public health intervention were well designed to achieve the desired results. The evaluation provided key points and actions to consider during intervention development to serve hard-to-reach communities and improve health outcomes. Further, it was shown that the results demonstrated a high degree of implementation fidelity. The clarity of the theory of the intervention, the motivation and commitment of the stakeholders as well as the periodic meetings of supervisors with the field team largely explained the high level of fidelity obtained. Contextual factors such as geographical distance to the intervention, access to a mobile phone, level of education, and gender norms contributed to the heterogeneity of the participation of the intervention target group. Finally, this evaluation underlined the finding that the mobile platform coupled with community mobilization, both key components of the intervention, were well received by the participants and may be an effective means of improving health knowledge and changing health-related behaviors.

This thesis contributes to the development of methodological knowledge concerning the evaluation of the fidelity of implementation of public health interventions by identifying gaps in this field, and by suggesting a tool that facilitates advancing this evaluation practice. This thesis also contributes to the strengthening of research in implementation sciences, and empirically provides key elements essential to assess the fidelity of the implementation of this type of intervention using CRT studies and evaluation of this fidelity, which is the subject of this doctoral research.

Keywords: Public health intervention; cluster randomized trials; systematic review; fidelity assessment; process evaluation; global health; developing countries.

Table des matières

Résumé	3
Abstract.....	5
Table des matières	7
Liste des tableaux	11
Liste des figures.....	12
Liste des sigles.....	13
Glossaire	14
Remerciements	16
Introduction	19
Chapitre 1 – État des connaissances	25
1.1 Les interventions en santé publique dans les pays du Sud.....	26
1.2 Des essais randomisés en grappes pour évaluer l’efficacité des interventions en santé publique.....	29
1.2.1 L’évaluation des interventions dites « complexes » en santé publique.....	34
1.3 Le but de l’évaluation et ses typologies	36
1.4 L’évaluation de la mise en œuvre d’une intervention	39
1.5 L’évaluation de la fidélité de la mise en œuvre des interventions en santé publique.....	43
1.5.1 Avantages pour les chercheurs d’évaluer la fidélité de la mise en œuvre de l’intervention	47
1.5.2 Avantages d’évaluer la fidélité du point de vue méthodologique	50
1.5.3 Des critères pour évaluer la fidélité de la mise en œuvre des interventions.....	52
1.6 Fidélité versus Adaptation	57

1.7 Évaluer la fidélité de la mise en œuvre de l'intervention Tika Vaani dans un contexte d'une étude pilote d'essais randomisés en grappes : étude de cas	61
1.7.1 Le but de la fidélité dans l'étude pilote	61
1.7.2 Contexte de l'intervention Tika Vaani	63
1.7.3 L'Intervention Tika Vaani	65
1.8 Synthèse du cadre théorique de la thèse au sein de la fidélité de la mise en œuvre	67
Chapitre 2 – Démarche scientifique, objectifs, méthode de recherche, rôle de l'étudiante et considérations éthiques	70
2.1 Premier volet : Revue systématique	71
2.1.1 Démarche scientifique	71
2.1.2 Objectif de recherche	72
2.1.3 Méthode de recherche	72
2.1.3.1 Article 1a Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review PROTOCOL	76
2.2 Deuxième volet : Étude de cas	111
2.2.1 Démarche scientifique	111
2.2.2 Objectif de recherche	113
2.2.3 Méthode de recherche	113
2.3 Rôle de l'étudiante	124
2.4 Considérations éthiques	126
Chapitre 3 – Résultats	128
3.1 Article 1B Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review	129

3.2 Article 2 Development of an mHealth Behavior Change Communication Strategy: A case-study from rural Uttar Pradesh in India.....	224
3.3 Article 3 Implementation fidelity and acceptability of an intervention to improve vaccination uptake and child health in rural India: A mixed methods evaluation of a pilot cluster randomized controlled trial.....	253
Chapitre 4 – Discussion	317
4.1 Synthèse, contribution des résultats et pistes d’action	318
4.2 Contribution méthodologique concernant l’évaluation de la fidélité de la mise en œuvre de l’intervention.....	321
4.2.1 Outil pour évaluer la fidélité de la mise en œuvre de l’intervention dans les essais randomisés en grappes	324
4.3 Contribution empirique à la pratique évaluative de la fidélité de la mise en œuvre de l’intervention.....	337
4.4 Recommandations pour de futures recherches.....	342
4.5 Forces et limites de la thèse	344
Conclusion.....	351
Annexes	374
Annexe 1 Grille pour l’extraction de données-Revue systématique	374
Annexe 2 Grille structurée pour les petites rencontres communautaires	376
Annexe 3 Grille d’entretien pour guider les rencontres communautaires	388
Annexe 4 Guide d'entretien pour guider les entrevues avec les travailleurs de la santé.....	392
Annexe 5 Grille d'entretien pour guider les entrevues et les focus groups avec les intervenants-terrain.....	396
Annexe 6 Grille pour évaluer la fidélité de la mise en œuvre de l’intervention.....	402
Annexe 7 Liste de control GRAMMS pour le rapport des études utilisant des méthodes mixtes	403

Annexe 8 Certificats d'éthique de la recherche	404
Annexe 9 Note de synthèse des résultats de la révision systématique	407
Annexe 10 Synthèse du panorama des chercheurs concernant l'évaluation de la fidélité des études inclues dans la révision systématique	410

Liste des tableaux

Tableau 1. –Interprétation de l'étude selon l'évaluation de la fidélité de la mise en œuvre de l'intervention.....	49
Tableau 2. –Synthèse sur les données collectées dans le cadre de la thèse doctorale.....	122
Tableau 3. –Outil pour évaluer la fidélité de la mise en œuvre des interventions en santé publique	327
Tableau 4. –Points saillants de la pratique évaluative de la fidélité de la mise en œuvre de l'intervention.....	337

List of tables

Article 1A

Table 1. Conceptual framework for implementation fidelity used in this review.....	80
--	----

Article 1B

Table 1. Conceptual framework for implementation fidelity used in this review.....	135
Table 2. General characteristics of included studies (n=65)	146
Table 3. Intervention fidelity assessment among CRTs of public health interventions in LMICs	147
Table 4. Fidelity components and moderating factors identified* (n=65)	147

Article 2

Table 1. Key findings of the formative research.....	238
Table 2. Basic cost regarding the mobile strategy.....	244

Article 3

Table 1. Outcome variables and data sources for the Tika Vaani pilot study.....	260
Table 2. Intervention components and beneficiaries for the Tika Vaani Intervention.....	263
Table 3. Role of the implementers and the research group in the Tika Vaani Intervention.....	264
Table 4. Overview the conceptual framework to assess acceptability, fidelity implementation and type of data collection for each dimension.....	266
Table 5. Content dimension.....	279
Table 6. Coverage dimension.....	281

Liste des figures

Figure 1. – Composantes et activités de l'intervention Tika Vaani	66
Figure 2. – Synthèse du cadre théorique de la thèse au sein de la fidélité de la mise en œuvre ...	69

List of figures

Article 1B

Figure 1. Flow diagram summarising study selection process.....	145
Figure 2. Summary of risk of bias in included RCTs.....	149

Article 2

Figure 1. IVR (Interactive Voice Response) structure for Tika Vaani.....	232
--	-----

Article 3

Figure 1. Mixed methods study flow diagram of collection, analysis, and integration of study data	275
Figure 2. Logic model for the "Tika Vaani" to improve vaccination uptake and other health outcomes	277
Figure 3. Participation of mothers to the small group meetings.....	282
Figure 4. Reasons for non-participation of mothers in small group meetings.....	283
Figure 5. Reasons given by meeting participants for not listening to the educational capsules via mobile phone.....	284
Figure 6. Factors influencing participation of women in community mobilization activities (n=184)	287
Figure 7. Factors influencing participation of women in educational capsules via mobile phone (n=184)	287

Liste des sigles

CME: Contexte-Mécanisme-Effet

ECR: Essai contrôlé randomisé

ERG: Essai randomisé en grappes

MRC: Medical Research Council du Royaume-Uni

NFHS: Enquête nationale sur la santé et la famille

OMS: Organisation Mondiale de la Santé

Glossaire

Adaptation: Modification du modèle d'origine d'une intervention par les personnes responsables de sa mise en œuvre.

Adhésion : Mise en œuvre du programme comme prévu en évaluant les composantes essentielles de l'intervention quant au *contenu* et à la *dose* de l'intervention reçue par les participants aussi souvent (*la fréquence*) et aussi longtemps (*la durée*) qu'elle aurait dû l'être.

Étude pilote : étude de plus petite taille qui vise à étudier la faisabilité des éléments cruciaux de l'étude principale, et à guider la planification d'une intervention à grande échelle.

Essais randomisés en grappes: Essais dans lesquels on ne randomise pas individuellement des sujets, mais de groupes de sujets qu'on appelle des « grappes ». Ces unités de randomisation peuvent être des unités sociales, ou des grappes d'individus tels que des familles, des écoles, des villages, etc.

Évaluation de la fidélité de la mise en œuvre : Évaluation qui vise à mesurer le degré selon lequel une intervention a été implantée telle que conçue par ses concepteurs. Cette évaluation vise à mesurer *l'adhésion* et à *explorer tous les facteurs qui la modèrent* pour mieux comprendre les résultats obtenus.

Évaluation du processus : Évaluation qui vise la mise en œuvre d'une intervention, afin de mieux comprendre la relation entre les éléments spécifiques de l'intervention et leurs résultats. Cette évaluation vise à analyser les différents facteurs qui peuvent agir sur l'intervention et influencer leurs résultats.

Fidélité de la mise en œuvre: Degré avec lequel une intervention est mise en œuvre comme prévue par ses concepteurs. Les termes : intégrité au traitement, conformité au traitement, fidélité à l'intervention, intégrité procédurale, qualité de la mise en œuvre sont souvent utilisés de manière interchangeable dans différents domaines, pour signifier le degré auquel les interventions sont réalisées telles que conçues et planifiées par les chercheurs.

À mes chers parents, Luz Elena et Alfredo,
qui m'ont transmis la valeur de l'éducation.
Merci d'avoir cultivé en moi l'exemple de
l'effort et du courage, pour ne pas craindre
l'adversité.

Remerciements

La réalisation de cette thèse a été possible grâce à l'appui de plusieurs personnes, à qui je voudrais témoigner toute ma reconnaissance.

Je voudrais tout d'abord adresser toute ma gratitude à ma directrice de recherche, Dre Mira Johri qui m'a encadrée et orientée tout au long de ce processus de formation. Je vous remercie pour votre gentillesse, pour votre patience et pour votre disponibilité en tout temps. Merci pour les nombreux encouragements que vous m'avez prodigués, et pour l'appui scientifique pour me soutenir au cours de l'élaboration de cette thèse. Merci pour vos conseils qui ont contribué à alimenter ma réflexion et à grandir sur le plan professionnel et personnel. Merci de croire en moi et de me donner l'opportunité de travailler au sein d'une merveilleuse équipe riche, autant sur le plan humain que professionnel. Merci pour votre générosité et pour votre soutien financier qui m'a permis de m'engager et de partager mes travaux de recherche, ainsi que de découvrir et d'agrandir mon réseau académique. Merci du fond de mon cœur.

J'adresse tous mes remerciements à Dr Valéry Ridde de l'honneur qu'il m'a fait en acceptant d'être mon codirecteur de cette thèse. Merci pour vos nombreux et précieux conseils et encouragements pour bien continuer dans ce processus, dont cette thèse est l'accomplissement. Vous êtes une source d'inspiration pour ne pas abandonner et pour donner mon maximum pour continuer. Votre expertise dans le domaine m'a motivée et m'a donné le goût d'explorer et d'approfondir ce champ de l'évaluation qui ne fait que commencer. Toute ma gratitude et tout mon respect.

Je tiens aussi à remercier toute l'équipe Tika Vaani qui m'a accueillie au sein du travail sur le terrain, et qui m'a permis de concilier la recherche théorique et pratique pendant cette thèse. Merci de me faire sentir en famille, de me faire grandir sur le plan humain et académique, et d'accroître ma passion pour travailler au sein des communautés. Je remercie spécialement Dinesh Chandra et Aaditeshwar Seth qui ont répondu avec patience aux questions quotidiennes pour réduire la distance entre le terrain et la recherche, et faciliter ainsi ma réflexion.

J'exprime ma gratitude à toutes les personnes qui ont facilité mon travail d'analyse et d'écriture pour accomplir ce projet de thèse. Merci aux coauteurs des articles qui ont contribué significativement à élaborer des articles avec des bases solides. Merci à Daniela Ziegler, bibliothécaire à l'Hôpital de l'Université de Montréal (CHUM), pour son aide dans la stratégie de recherche de base de données. Merci à Pawan Kumar pour son aide pour la conduite des entretiens en langue locale (Hindi) avec les intervenants-terrain. Merci à Kalpa Sharma pour la traduction des données en anglais, et merci à Frédérique Sauvé pour son soutien à la gestion des données. Un Merci spécial au Dr Beto pour s'être porté volontaire pour la lecture et la correction du français de cette thèse, qui permettent aux lecteurs de mieux comprendre le contenu de cette thèse. Dr Beto, votre engagement altruiste et votre disponibilité en tout temps n'ont pas de limites. Merci mille fois d'avoir pris le temps de bien comprendre ma pensée pour m'aider à mieux m'exprimer, cela signifie vraiment beaucoup pour moi.

Je suis également reconnaissante envers toutes les personnes de la communauté qui ont eu la gentillesse de participer à des entretiens, et envers celles qui ont accepté de répondre aux questions posées lors de la collecte des données pour faire avancer mon analyse. Merci pour votre temps et votre patience.

Merci aux différentes institutions pour le soutien financier qui m'a permis de me consacrer à mon projet doctoral, et de rendre possible ce projet de formation. Merci au IC-IMPACTS (Le Centre Inde-Canada pour des partenariats multidisciplinaires innovants pour accélérer la transformation et la durabilité des communautés) qui a financé mes études sous la forme de bourses de doctorat à travers le concours financier de ma directrice de recherche. Merci au Centre de recherche du CHUM pour l'allocation de formation, et à l'Institut de recherche en santé Canada IRSC pour l'allocation de voyage qui a facilité encore plus ce processus de recherche et de diffusion de connaissances.

Je tiens à remercier toutes les personnes avec qui j'ai partagé mes études, et notamment ces années de thèse. Merci aux « mousquetaires » pour les échanges d'apprentissage, de réalisations, de doutes et de frustrations qui ont rendu ce processus plus humain.

Mes remerciements vont aussi à toute ma famille et à mes amis pour leurs encouragements permanents, leur soutien moral tout au long de mon parcours universitaire et pour leur sincère amitié et confiance. Monica, Laura, Isa, Ana, Robert, Andres, Diego, Sandra, Marcela, et Roman vos questions répétées pour savoir : Comment va la thèse et quand vas-tu la soutenir? ainsi que votre motivation et votre écoute attentive, lorsque j'en avais besoin, m'ont permis de surmonter des moments difficiles et des moments de doute pendant mon parcours doctoral. À vous tous, je vous aime fortement et je vous présente mes remerciements et ma gratitude.

Enfin, je remercie de tout mon cœur mon cher fils Mathieu et mon cher époux Felipe pour leur soutien quotidien, autant pour mener mes travaux à terme que pour la vie en général. Tous les deux m'avez donné le temps, qui vous appartenait sans reproche, pour achever cette thèse. Mathieu, ta tendresse et ton sourire ont été une source indispensable tout au long de ce processus qui m'ont encouragée à continuer. Felipe, tu as vécu en silence mes moments difficiles et incompréhensibles pour toi. Tu étais toujours à mes côtés pour me réconforter. Merci pour ton soutien et ta patience dans ce projet d'étude. Aucun mot ne serait assez fort pour exprimer ma gratitude et mon amour envers toi. Vous êtes, tous les deux, une source d'inspiration qui me motive à poursuivre la meilleure version de moi-même.

Introduction

En vue de répondre à l'appel mondial de travailler ensemble pour éradiquer la pauvreté, assurer le bien-être des populations et réduire des inégalités persistantes, la mise en place des interventions pour faire des progrès significatifs et pour atteindre les Objectifs de Développement Durable (ODD) devient un élément clé. Dans ce contexte, l'identification et la priorisation des interventions constituent un volet indispensable afin de pousser les pays à avancer vers leurs cibles locales et nationales. Cette approche est encore plus importante pour les pays du Sud qui font face à un ensemble de difficultés en matière de santé publique au niveau, entre autres, de l'accès, du coût, des ressources et de la qualité des soins qui affectent la santé des populations (Jacob, 2009; Severino, 2008; Wynn, Dutta, & Nelson, 2006). Dans ce contexte, des interventions en santé publique sont souvent mises en œuvre dans le cadre de l'aide au développement en se confrontant à des enjeux de pérennité qui rendent plus difficile la continuité des activités liées à l'intervention subventionnée (Samb, Ridde, & Queuille, 2013; Severino, 2008; Tizio, 2004).

Devant cet ensemble de circonstances, le besoin d'établir des priorités, de choisir les interventions les plus efficaces, de mieux comprendre les conditions dans lesquelles les interventions sont menées, et de participer à l'amélioration de ces interventions avant sa mise en échelle, est imminent. L'évaluation peut y contribuer en facilitant un processus décisionnel transparent et fondé sur des connaissances scientifiques solides (Brownson, Fielding, & Maylahn, 2009; H. T. Chen & Chen, 2005). Elle peut également mieux répondre aux questions qui expliquent comment et pourquoi les interventions de santé publique fonctionnent (Steckler, Linnan, & Israel, 2002). À cet égard, l'évaluation devra concentrer ses efforts pour examiner et comprendre les forces et les faiblesses de ces interventions et son interaction dans un contexte donné, en allant au-delà de la mesure de l'efficacité des interventions, afin d'optimiser la conception de l'évaluation et leur utilisation (M. Patton, 2008). Ainsi, avant de dépenser les ressources souvent limitées dans les pays du Sud, une évaluation qui vise à comprendre comment une intervention fonctionne dans son contexte, pour aider les chercheurs et décideurs à l'améliorer, est d'autant plus exigée.

L'essai contrôlé randomisé (ECR) est souvent utilisé en recherche clinique pour tester les effets d'un médicament, d'une thérapie ou d'une intervention sur un groupe expérimental, qui

bénéficiera de l'intervention en la comparant à un groupe contrôle qui recevra un placebo ou aucun traitement. La sélection des participants de chaque groupe se fait de façon aléatoire (Kendall, 2003). Une personne qui participe à un essai contrôlé randomisé est appelée un participant ou un sujet. Les chercheurs vont mesurer et comparer les résultats après que les participants ont reçu l'intervention, ou non, dans un ordre aléatoire. Dans ce type d'étude, les chercheurs peuvent employer différentes techniques pour réduire des biais, permettre ainsi d'estimer une causalité et de conclure que le résultat est dû à l'intervention étudiée. Pour certains auteurs, les principes de la randomisation et de la comparaison contrôlée réduit les biais, et fournit un outil rigoureux pour examiner les relations de cause à effet, en le plaçant ainsi comme un des outils les plus simples et les plus puissants de la recherche clinique (Akobeng, 2005; Hariton & Locascio, 2018). Pour d'autres, la randomisation ne constitue pas la « norme d'or » pour la recherche en éducation médicale (Sullivan, 2011) où ils considèrent que des études observationnelles bien conçues peuvent fournir des informations cliniquement pertinentes, par exemple sur l'efficacité et l'innocuité des médicaments à long terme qui ne sont pas nécessairement fournis par les ECR (Benson & Hartz, 2000; Silverman, 2009). Bien que le débat persiste et bien que certains questions ne peuvent pas être résolues par l'essai contrôlé randomisé, ces études bien menées et bien conçues constituent une preuve importante pour documenter les décisions en matière de soins de santé (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996).

Sous cette perspective, et face au besoin de développer des interventions en santé publique qui visent à améliorer la santé des populations, un grand nombre d'essais randomisés en grappes est davantage utilisé dans les pays du Sud par rapport aux essais individuels. D'un côté pour limiter, entre autres, le risque de contamination et éviter une sous-estimation de l'effet réel de l'intervention mais aussi pour atteindre à un plus grand nombre des communautés entières et favoriser l'efficacité administrative ou la réduction de coûts. (Bland, 2004; M. J. Campbell, Donner, & Klar, 2007; Handlos, Chakraborty, & Sen, 2009; Isaakidis & Ioannidis, 2003; Moberg & Kramer, 2015; Valéry Ridde & Dagenais, 2019). Dans ce type d'essai, ce ne sont plus des individus qui sont randomisés, mais des groupes d'individus ou des unités de randomisation collective tels que les familles, les villages, les écoles, les hôpitaux qui vont recevoir l'intervention. Les essais randomisés en grappes sont largement utilisés dans l'évaluation des interventions du système de santé, compte tenu de leur importance en tant que source de données probantes (M.K.

Campbell, Mollison, Steen, Grimshaw, & Eccles, 2000; Osrin et al., 2009) pour mesurer l'effet global d'une intervention au niveau de la population (Hayes, Alexander, Bennett, & Cousens, 2000; Isaakidis & Ioannidis, 2003), ainsi que pour sa rigueur en évaluant l'efficacité des interventions, quelle que soit leur complexité (Oakley, Strange, Bonell, Allen, & Stephenson, 2006; Richards & Hallberg, 2015). Cependant, des défis méthodologiques dans ce type d'étude qui favorisent des biais devront être pris en compte à partir des méthodes statistiques appropriées, afin d'attribuer les effets de l'intervention avec un niveau de confiance plus élevé (Duflo, Glennerster, & Kremer, 2007) autant dans le contexte de soins primaires au niveau local (M. K. Campbell et al., 2000) que dans le contexte des interventions de santé mondiale (Dumont, 2019).

Selon Rychetnik et collègues. (2002), les interventions en santé publique ont tendance à être «complexes, programmatiques et dépendantes du contexte», ce qui posent des défis pour leur évaluation (Rychetnik, Frommer, Hawe, & Shiell, 2002). Les interventions « complexes » ont été utilisées dans la littérature en santé comme raccourcis, pour décrire des interventions construites à partir d'un certain nombre de composantes qui agissent de manière indépendante et interdépendante et qui peuvent varier pendant la mise en œuvre (Medical Research & Council., 2008). Cette variation peut se présenter par un ensemble de facteurs internes et externes à l'intervention : soit par la nature de la conception de l'étude, soit par le contexte, soit par les participants et soit par les intervenants, qui vont contribuer au succès ou à l'échec de l'intervention (Anderson et al., 2013; Craig et al., 2008; Noyes et al., 2013; Wells, Williams, Treweek, Coyle, & Taylor, 2012). Pour les essais randomisés en grappes, la prise en compte des interactions entre les composantes de l'intervention et tous ces facteurs est encore plus importante, car les interventions peuvent être dispensées à différents niveaux et dans des environnements communautaires, où il y a divers paramètres qui empêchent ou facilitent la mise en œuvre (Duffy et al., 2015; Grant, Treweek, Dreischulte, Foy, & Guthrie, 2013).

Des revues systématiques ont soulevé que les interventions complexes, comme dans le cas des essais randomisés en grappes, peuvent être à risque d'éprouver des écarts du protocole et de la non-adhésion, et compromettre la validité des conclusions (Bonell, Oakley, Hargreaves, Strange, & Rees, 2006; S. Eldridge, Ashby, Bennett, Wakelin, & Feder, 2008). Pour atténuer ces problèmes méthodologiques, qui peuvent affecter la validité interne et externe des résultats des essais

randomisés, l'évaluation des processus y compris l'évaluation de la fidélité de la mise en œuvre de l'intervention devient fondamentale (Brierley, Brabyn, Torgerson, & Watson, 2012; Isaakidis & Ioannidis, 2003; Moore et al., 2015; Oakley et al., 2006; S. Puffer, D. Torgerson, & J. Watson, 2003). En effet, les initiatives de développement exigent des évaluations de haut niveau pour savoir si les programmes fonctionnent et comment les améliorer (Banerjee, Banerjee, & Duflo, 2011; Duflo et al., 2007). L'évaluation du processus répond à ce besoin en surveillant la mise en œuvre d'une intervention (M. Patton, 2008). Cette évaluation aide les parties prenantes à expliquer les succès, les échecs et les changements d'une intervention, car elle vise à analyser les différents facteurs qui peuvent agir sur l'intervention et influencer leurs résultats (Jean-Louis & Champagne, 1990; Moore et al., 2015; Steckler et al., 2002). Un élément essentiel dans l'évaluation de processus est la dimension de la fidélité, afin de comprendre quelles raisons ont provoqué la réussite ou l'échec de l'intervention (Carroll et al., 2007; Medical Research & Council., 2008; Richards & Hallberg, 2015).

Au sens large, la fidélité de la mise en œuvre de l'intervention évoque le degré selon lequel une intervention a été implantée telle que conçue par ses concepteurs (Carroll et al., 2007; Dane & Schneider, 1998; Dusenbury, Brannigan, Falco, & Hansen, 2003). Cette évaluation permet de savoir si les effets mesurés peuvent être attribués aux interventions ou à des facteurs exogènes (Carroll et al., 2007; H. T. Chen, 2014; Gearing et al., 2011; Jean-Louis & Champagne, 1990; Proctor, 2004; Steckler et al., 2002). Cette compréhension est importante pour aider les évaluateurs à éviter de commettre l'erreur de type III (Dobson & Cook, 1980), soit de conclure qu'une intervention n'est pas efficace, alors qu'en réalité, c'est plutôt la mise en œuvre qui est défectueuse (Hulscher, Laurant, & Grol, 2003). Ainsi, l'évaluation de la fidélité de la mise en œuvre de l'intervention est une étape essentielle vers la création "d'interventions durables" (Breitenstein et al., 2010). Cependant, pour entreprendre une telle évaluation, il est nécessaire de savoir d'abord si l'intervention à évaluer repose sur des fondements théoriques solides pour établir la plausibilité de la théorie du programme (Brousselle, Champagne, Contandriopoulos, & Hartz, 2011). Ces fondements théoriques supportent l'enchaînement causal qui explique les effets escomptés (Weiss, 1997), et donnent ainsi confiance aux décideurs que les effets observés sont dus au programme.

Au vu de l'état actuel des connaissances, l'objectif général de cette recherche doctorale est d'examiner la fidélité de la mise en œuvre des interventions en santé publique dans le cadre des essais randomisés en grappes, pour savoir si les interventions mises en place sous un contexte contrôlé doivent tenir compte de ce type d'évaluation. Donc, nous souhaitons explorer spécifiquement :

a) Comment l'évaluation de la fidélité de la mise en œuvre des interventions en santé publique, utilisant des essais randomisés en grappes (ERG) dans les pays du Sud, a été rapportée? Cette question sera abordée à travers une revue systématique.

b) Comment tenir compte de l'évaluation de la fidélité de la mise en œuvre d'une intervention en santé publique dans le cadre d'un essai randomisé en grappes dans les pays du sud en vue de l'améliorer, si nécessaire, avant sa mise en place à grande échelle?

Cette question sera abordée par l'évaluation de l'intervention Tika Vaani mise en place dans le contexte d'un essai randomisé en grappes dans la zone rurale d'Hardoi (Uttar Pradesh) en Inde. Cette évaluation sera faite à travers une étude de cas unique à trois unités d'analyse, selon une approche mixte concomitante. Cette évaluation vise trois objectifs spécifiques suivants:

- a) Comprendre la conception de l'intervention Tika Vaani afin d'analyser leur plausibilité
- b) Mesurer le degré de fidélité de la mise en œuvre de l'intervention Tika Vaani
- c) Comprendre le rôle des facteurs modérateurs dans le degré d'hétérogénéité et de la fidélité de la mise en œuvre de l'intervention Tika Vaani.

La thèse est organisée en quatre chapitres :

Le premier chapitre dresse un état de connaissances sur des sujets qui touchent de façon transversale cette recherche doctorale. D'abord, nous faisons un survol sur la mise en place des interventions de santé publique dans les pays du Sud, sur leur complexité et sur l'importance de les

évaluer. Ensuite, une section propose un survol des différentes perspectives théoriques concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention, ainsi que de l'importance de cette évaluation dans des interventions évaluées par des essais randomisés en grappes. Ce survol fait ressortir le cadre conceptuel retenu pour cette thèse. Enfin, un portrait général du contexte et de l'intervention à évaluer sera présenté plus en détail dans le Chapitre Résultats.

Le deuxième Chapitre présente la démarche scientifique de la recherche doctorale, et énonce les objectifs, la méthodologie de recherche, le rôle de l'étudiante-chercheuse dans ce processus de recherche évaluative, ainsi que les considérations éthiques de la recherche. Dans ce Chapitre, le protocole publié du premier article est présenté pour répondre à l'objectif général de cette thèse doctorale.

Un troisième Chapitre présente les résultats de recherche correspondant aux résultats de la revue systématique rapportés sous la forme d'un article scientifique, et présente deux autres articles scientifiques correspondant aux résultats de l'étude du cas du développement et de l'évaluation de la fidélité de la mise en œuvre d'une intervention mise en place dans le cadre d'un essai randomisé en grappes dans un contexte des ressources limitées.

Enfin, le quatrième Chapitre aborde la discussion des résultats de recherche et leur contribution, ainsi que les principales forces et limites de la thèse, pour finalement aborder quelques avenues futures de recherche dans le domaine de la recherche évaluative.

Chapitre 1 – État des connaissances

L'objet de ce chapitre est de capitaliser sur les savoirs dans le monde de la recherche concernant les différents sujets qui touchent de façon transversale cette thèse. Ainsi, ce chapitre de la thèse est présenté en trois grandes sections.

La première section met en lumière le rôle de la santé publique et de sa pratique à travers ses interventions, pour ensuite soulever les défis auxquels ces interventions sont confrontées dans les pays du Sud.

La deuxième section met en valeur le champ de l'évaluation en soulevant la pertinence de l'approche de l'évaluation de la fidélité de la mise en œuvre, et tout le fondement de cette pratique évaluative.

Enfin, la troisième section présente le portrait de l'intervention de santé publique mise en place dans le cadre d'un essai randomisé en grappes, qui fait l'objet de l'évaluation de la fidélité de la mise en œuvre de cette intervention et qui correspond à notre étude de cas.

1.1 Les interventions en santé publique dans les pays du Sud

La santé publique, c'est «la science et l'art de prévenir les maladies, de prolonger la vie et d'améliorer la santé physique et mentale à un niveau individuel et collectif» (Nutbeam, 1998). Elle se voit façonnée par la structure sociale dans laquelle elle se développe comportant des enjeux sanitaires, socioéconomiques, démographiques, épidémiologiques, environnementaux et politiques, qui font de ce domaine un champ riche de connaissances propice à l'action (Fassin, 1996, 2005; Pommier & Grimaud, 2007).

Sous cette perspective de la santé publique, vue comme un champ d'action vers lequel convergent de nombreuses disciplines, la santé publique doit viser des actions supportées par l'éducation, la recherche et l'élaboration des politiques pour créer des interventions dirigées vers le maintien et l'amélioration de la santé et le bien-être de toute la population. Donc, au contraire de la médecine clinique qui opère au niveau individuel, l'essence de la santé publique consiste à adopter une perspective basée dans la population, qui exige des efforts coordonnés visant à améliorer la santé des populations (Valery Ridde, 2007). C'est pourquoi ce domaine demande des stratégies axées sur la population, qui dépendent fortement des structures gouvernementales et des partenariats intersectoriels qui les soutiennent et qui les financent pour apporter des changements qui améliorent la santé.

Bien que la santé publique ait connu des évolutions aussi bien sur le plan conceptuel, institutionnel, normatif et en matière de santé, des enjeux restent encore à surmonter, pouvant faire l'objet de diverses interventions selon les réalités de chaque pays. Pour Rychetnik et collègues (2002), une intervention en santé publique, en tant qu'intervention de prévention de la maladie ou de promotion de la santé, c'est une intervention appliquée à une grande partie, à la plupart ou à tous les membres d'une communauté, qui vise à procurer un avantage net à la communauté, ainsi qu'aux individus (Rychetnik et al., 2002). Ces interventions, inscrites dans les fonctions essentielles de la santé publique, facilitent un terrain de rencontre auquel plusieurs acteurs peuvent identifier leur rôle pour y contribuer (Pommier & Grimaud, 2007). Cependant, les interventions en santé publique varient d'un pays à un autre en fonction de sa réalité locale. Dans les pays du Sud,

ces interventions sont souvent conditionnées par des choix politiques, par la disponibilité de ressources à tous les niveaux et par de nombreux obstacles à la mise en œuvre, qui mettent en péril le plan d'action pour maintenir et améliorer la santé de la population.

Les Objectifs du Millénaire pour le Développement (OMD) ont fourni un contexte pour l'établissement d'engagements et d'initiatives importants pour améliorer la santé de la population à l'échelle mondiale. Le rapport final a révélé des progrès importants en matière de santé en vue de réaliser les Objectifs du Millénaire (Nations-Unies, 2015). Deux constats, confirmant cette réussite, révèlent que la pauvreté et la faim ont été réduites presque de moitié dans le monde depuis 1990. Cependant, malgré des progrès importants remarquables dans tous les indicateurs cibles permettant d'améliorer les conditions de vie des personnes, ces progrès ont été inégaux au sein de la population favorisant des écarts importants entre les pays et entre les régions d'un même pays. Les principaux groupes vulnérables, tels que les populations les plus pauvres et les plus isolées, ont été exclus et marginalisés de ces progrès (Lozano et al., 2018).

En effet, malgré qu'à l'échelle mondiale, l'espérance de vie s'améliore à un rythme de plus de 3 ans par décennie depuis 1950 (WHO, 2016), des défis restent à surmonter à plusieurs niveaux. Par exemple, les maladies non transmissibles sont responsables d'environ 38 millions de décès par an, ce qui représente 68% de tous les décès dans le monde (WHO, 2016). En ce qui concerne les maladies transmises par des vecteurs, quant à elles, sont responsables de plus de 17% des maladies infectieuses, et provoquent plus d'un million de décès chaque année (National Academies of Sciences & Medicine, 2016). En plus des facteurs climatiques, d'autres facteurs tels que le contexte environnemental, le contexte social et comportemental, ainsi que des caractéristiques propres de l'individu telles que l'âge, le genre et la fonction immunitaire, peuvent également affecter la vulnérabilité en influençant la sensibilité à l'infection (Beard et al., 2016). L'ensemble de ces conditions peut contribuer à accroître le taux de mortalité dû aux maladies vectorielles. Nous pouvons le constater actuellement avec la pandémie en cours de COVID-19, qui exige toujours de prendre des mesures de précaution en matière de santé publique (Heymann & Shindo, 2020), mais lesquelles devront être pensées et adaptées, selon le contexte de chaque pays et avec une vision globale.

Un autre constat, malgré les progrès, c'est qu'au moins la moitié de la population mondiale ne reçoit pas des services de santé essentiels dont elle a besoin et, chaque année, environ 100 millions de personnes sont plongées dans l'extrême pauvreté en raison des dépenses personnelles en santé (WHO, 2017). Pour que les individus puissent accéder à des services de santé de qualité, sans difficultés financières individuelles ou familiales, il faut des améliorations substantielles des systèmes de santé et des stratégies de financement pour garantir la disponibilité efficace des services de santé (Mukherjee et al., 2019). Cependant, cela implique aussi faire des efforts importants du côté de la demande, si on veut atteindre toute la population et ainsi réduire des inégalités en matière de santé.

D'autres progrès, en matière de santé, laissent voir que le taux de mortalité infantile a diminué de façon importante au niveau mondial (H. Wang et al., 2016). Néanmoins, les progrès n'ont pas été égaux pour tous les âges de l'enfant. Il y a encore 5,4 millions d'enfants qui meurent avant leur cinquième anniversaire (Roser, Ritchie, & Dadonaite, 2013), et ce sont les enfants vivant dans les zones rurales qui ont contribué le plus à la mortalité infantile, comparativement aux enfants vivant dans les zones urbaines qui ont de meilleurs résultats de santé (Fink, Günther, & Hill, 2014). Donc, bien que des progrès réels aient été faits en matière de santé au cours des dernières décennies, des inégalités persistent et elles se traduisent par des écarts de santé qui exigent des actions permettant de mieux contribuer à les prévenir ou les réduire.

« Ne laisser personne de côté » est un appel universel à l'action pour travailler à ce que tout le monde profite de la volonté d'atteindre les Objectifs de Développement Durable du Programme 2030, adopté en 2015 par tous les États membres des Nations Unies. Donc, cet appel lancé évoque l'importance de déployer des efforts globaux et coordonnés pour intensifier des interventions efficaces, qui demandent des approches innovantes, intégrées, transdisciplinaires et avec des multipartenaires. Cependant, concevoir une intervention réussie nécessite de savoir ce qui fonctionne bien dans un environnement particulier, et de savoir quels sont les problèmes les plus susceptibles de se poser. La littérature démontre que les pays du Sud font toujours face aux défis en matière de santé publique qui exigent, entre autres, le développement et la mise en place des interventions pour atteindre les objectifs ciblés. Cependant, ces interventions devraient faire l'objet d'une attention particulière au moment de son évaluation car des facteurs socioéconomiques,

d'infrastructures, de ressources humaines et culturels, par exemple, peuvent influencer sa mise en œuvre. À ces enjeux, s'ajoute le fait que la plupart de ces interventions sont souvent dépendantes de l'aide des ressources humaines et économiques qui proviennent de l'extérieur du pays, en les confrontant à des enjeux de continuité (Valéry Ridde & Capelle, 2011; Samb et al., 2013; Severino, 2008). Donc, l'évaluation de sa mise en œuvre est nécessaire, voire fondamentale, avant de dépenser les ressources souvent très limitées, et avant de définir ses effets escomptés.

1.2 Des essais randomisés en grappes pour évaluer l'efficacité des interventions en santé publique

L'évaluation des interventions en santé publique est essentielle pour éclairer la prise de décision fondée sur des preuves, et pour mieux dégager quelles sont les interventions les plus efficaces et les plus bénéfiques pour tous. Cette évaluation permettra aussi de mieux informer l'allocation des ressources, et de cerner les priorités ou les besoins de concevoir des nouvelles stratégies visant à améliorer la santé. Bien que l'amélioration de la santé de la population continue d'être une priorité pour tous les pays, dans les pays du Sud, ces évaluations devront s'imposer afin de garantir la mise en œuvre des interventions qui peuvent contribuer à réduire les inégalités persistantes en santé (Severino, 2008; Wynn et al., 2006). Donc, éclairer les futures décisions qui facilitent les connaissances sur l'efficacité des interventions est nécessaire pour orienter les efforts de santé publique, et devient une priorité pour tous les décideurs.

En médecine, la pratique fondée sur des preuves comprend trois éléments clés : les données probantes fondées sur la recherche, l'expertise et les compétences par l'intervenant, et les valeurs et les préférences du patient (Masic, Miokovic, & Muhamedagic, 2008). L'avantage de cette approche est que les prestataires « écoutent » les préoccupations des patients et les « prennent en considération » pour déterminer le plan de traitement approprié. En santé publique, cette approche est un processus d'intégration des interventions scientifiques qui répondent aux besoins de la communauté pour améliorer la santé des populations, et c'est un élément clé pour la formulation de programmes et de politiques concernant la santé, ainsi que pour évaluer les interventions et les progrès en matière de santé (Glasgow & Emmons, 2007; Jenicek, 1997). En outre, en santé

publique, toute stratégie d'intervention est déterminée non seulement par son efficacité, mais aussi par la mesure dans laquelle elle est adoptée, de sorte qu'elle ait un impact sur la santé des groupes d'individus dans la communauté et de la population dans son ensemble. Également, ces interventions devront avoir la capacité d'agir sur un large éventail de facteurs sociaux et environnementaux qui influencent la santé (Parsaik & Singh, 2013; Rychetnik, Hawe, Waters, Barratt, & Frommer, 2004; Upshur, 2002).

Étant donné l'éventail des mesures et des méthodes pour établir l'efficacité d'une intervention (C. H. Brown et al., 2017), il y a divergence d'opinion sur quelles sont les meilleures données probantes qu'on devrait privilégier pour améliorer la santé de la population cible. Des chercheurs soulignent que des modèles d'études, qui n'incluent pas des essais randomisés individuels, peuvent mieux répondre aux questions de recherche, et vont également fournir des informations clés pour prendre des décisions concernant les soins des personnes (Sackett et al., 1996). Bien que la randomisation vise à éliminer tout biais de confusion pour qu'il soit possible de déterminer la causalité, elle ne supprime pas au complet les causes de contamination importantes, pour établir que ce type d'étude soit la seule source de « bonnes » preuves scientifiques (Scriven, 2008). Par exemple, Scriven fait ressortir deux facteurs pouvant invalider l'inférence du traitement expérimental: (i) la difficulté de séparer les effets « intrinsèques » d'un traitement des effets psychologiques de l'administration du traitement; ainsi que (ii) la difficulté à dire quel sera l'effet du traitement expérimental s'il est administré en dehors du contexte expérimental, car on ne peut évaluer combien ce contexte ajoute aux effets (Scriven, 2008). Selon Bédécarrats et collègues, l'utilisation de ce type d'étude se voit massivement répandue par les exigences de la communauté des donateurs et du monde universitaire, qui répondent à une combinaison politique et économique pour le développement des nouvelles interventions qui ne tient pas compte parfois d'autres méthodes, pour identifier des phénomènes nouveaux et inattendus aussi importants pour contextualiser les politiques de développement (Bédécarrats, Guérin, & Roubaud, 2018). Donc, la richesse du champ de la recherche pour évaluer des interventions peut s'appuyer sur des approches méthodologiques distinctes, mais à la fois complémentaires pour étudier l'effet causale d'une intervention.

Évaluer des interventions dans des études qui vont au-delà des études expérimentales ou quasi expérimentales telles que des études observationnelles peut être considéré, car elles sont reconnues, entre autres, pour : être moins coûteuses, refléter la diversité des populations, offrir l'avantage de mieux répondre aux besoins de la pratique quotidienne, s'appuyer sur immenses bases de données des patientes et faire un longue période de suivi, apporter de plus amples informations sur l'efficacité et la sécurité du traitement, et probablement plus facile à implémenter que les essais contrôlés randomisés (Ligthelm et al., 2007; Scriven, 2008; Silverman, 2009). De même, pour Concato et collègues, les études observationnelles peuvent présenter moins d'hétérogénéité des résultats que les essais contrôlés randomisés sur un même sujet d'étude (Concato, Shah, & Horwitz, 2000). Sanson-Fisher et collègues soulèvent que le choix des méthodologies doit tenir compte des différences entre la collecte, l'analyse et l'application des preuves au niveau clinique et au niveau de la population. En effet, les chercheurs considèrent que les interventions axées sur la santé de la population sont plus susceptibles de rencontrer des limites concernant la disponibilité de la population, le délai de suivi, la validité externe, le coût, l'éthique et le consentement éclairé en s'appuyant sur des essais contrôlés randomisés. (Sanson-Fisher, Bonevski, Green, & D'Este, 2007). Pour ces auteurs, des méthodes autres que des essais contrôlés randomisés peuvent être utilisées pour répondre aux différentes questions de recherche. D'autres chercheurs ont suggéré des essais contrôlés randomisés de type « réalistes » pour combler entre autres les lacunes des problèmes de généralisation des résultats des essais contrôlés randomisés qui évaluent des interventions de santé publique, en particulier lorsque les interventions sont de nature complexe (Jamal et al., 2015). Ce type d'intervention vise à préserver les forces des essais contrôlés randomisés pour fournir des preuves avec une forte validité interne dans l'estimation des effets, et, également, profite de la force de l'approche réaliste pour mieux expliquer les processus sous-jacents de la mise en œuvre et des mécanismes d'action, et comment ceux-ci varient en fonction des caractéristiques contextuelles des participants et du site d'intervention (Jamal et al., 2015). Cependant, cette approche ne permet pas d'identifier l'interaction dynamique entre l'intervention, les acteurs, le contexte, les mécanismes et les résultats, qui est au cœur de la recherche réaliste (Van Belle et al., 2016). En effet, pour Van Belle et collègues, les différences paradigmatiques dans le traitement de la causalité entre la logique post-positiviste et réaliste ne permettent pas cette fusion. Selon ces auteurs, la causalité fonctionne à travers des effets interactifs qui sont essentiellement basés sur des interactions entre les personnes. Donc, les explications causales

nécessitent d'analyser les interventions dans une perspective systémique avec une orientation basée sous la forme de configuration Contexte-Mécanisme-Effet (CME) par opposition à une orientation basée sur des variables (Van Belle et al., 2016).

Bien que le débat persiste pour déterminer si les essais randomisés sont des études considérées « gold standard » (Grossman & Mackenzie, 2005), il est reconnu que ce type d'étude, par sa rigueur méthodologique, permet bien de tester l'efficacité d'une intervention, car la randomisation réduit les biais et fournit un outil rigoureux pour examiner les relations de cause à effet entre une intervention et les résultats. (Akobeng, 2005; Hariton & Locascio, 2018). Depuis le début du siècle, l'essai randomisé en grappes a été largement implémenté dans la recherche médicale (Bland, 2004), et il est couramment utilisé pour évaluer des interventions en santé publique (M. J. Campbell et al., 2007; Isaakidis & Ioannidis, 2003). L'essai randomisé en grappes est un essai dans lequel des unités sociales, ou des grappes d'individus, plutôt que des individus eux-mêmes, sont randomisées dans un groupe d'intervention (Donner & Klar, 2000; Handlos et al., 2009). Ce type d'étude teste l'intervention dans un groupe d'individus soit par exemple, des groupes de services de soins, de professionnels, des écoles ou des groupes d'habitants d'un même village. Des interventions en santé publique sous ce modèle sont souvent implémentées dans les pays du Sud pour plusieurs raisons (Handlos et al., 2009; Osrin et al., 2009). Cela s'avère être une alternative pour améliorer l'impact sur l'équité en santé et pour fournir des informations clés aux décideurs politiques pour une possible mise en œuvre de l'intervention à grande échelle (Hayes et al., 2000; Ranson, Sinha, Morris, & Mills, 2006). Il semble que ce type d'étude, pour des raisons éthiques, logistiques et administratives, soit une alternative plus efficace, pratique et peu coûteuse pour mesurer l'effet global de la mise en œuvre d'une intervention au niveau populationnel. Cependant, ces mêmes enjeux peuvent devenir un problème au moment de choisir ce type d'étude, soit par le coût associé pour le recrutement des participants et du personnel pour mettre en place l'intervention dans une longue période du temps, ou soit par exemple pour des raisons logistiques et administratives, voire éthiques, qui se génèrent au moment de la randomisation des unités ou des grappes ciblées par l'étude.

En effet, bien que les essais randomisés en grappes offrent plusieurs avantages par rapport aux essais randomisés individuels, des lacunes méthodologiques ont été soulevées dans la

littérature en utilisant ce type d'étude. Ces défis peuvent conduire aux problèmes de biais de recrutement et produire une estimation biaisée de l'effet du traitement. Ces biais devront être considérés au moment de la conception de l'étude (Hahn, Puffer, Torgerson, & Watson, 2005). Par exemple, faire la distinction entre l'unité de randomisation et l'unité qui sera prise en compte pour l'analyse, limite parfois l'extrapolation des résultats. Le fait de considérer les participants comme une seule unité d'analyse plutôt que comme différentes unités indépendantes fait perdre la puissance statistique. Cela amène les chercheurs à augmenter le nombre des participants dans les grappes pour avoir des conclusions plus solides, et à éviter la confusion de savoir si les résultats observés sont dus à l'intervention ou aux caractéristiques des participants du groupe qui reçoivent l'intervention (Leyrat, Morgan, Leurent, & Kahan, 2018). De même, les chercheurs, en référence aux essais randomisés en grappes, devront s'assurer que les groupes d'intervention et de contrôle ne soient pas systématiquement différents les uns par rapport aux autres en termes de caractéristiques pour ne pas générer des estimations biaisées. (S. M. Eldridge, Ashby, Feder, Rudnicka, & Ukoumunne, 2004). Ces défis méthodologiques peuvent souvent être compensés ou surmontés par des outils statistiques pour renforcer la conception de ce type d'étude et leurs analyses (Leyrat et al., 2018; van Breukelen & Candel, 2018).

D'autres auteurs ont relevé des défis éthiques liés à la mise en place des interventions sous ce modèle concernant, par exemple, le consentement du responsable de grappe ou la nécessité du consentement des individus au sein des groupes pour la collecte de données, l'obtention de l'approbation éthique des organisations ou du gouvernement local pour mettre en place l'intervention, ainsi que le temps d'attente des personnes du groupe témoin à recevoir l'intervention testée dans une phase postérieure (Osrin et al., 2009). Donc, il est fortement suggéré que des enjeux éthiques à l'intérieur de la recherche soient considérés sous une approche « globale, transdisciplinaire et inclusive » permettant de trouver un équilibre entre la population et les parties prenantes, afin de ne pas reproduire des inégalités sociales au sein de ces études (Hamelin et al., 2018).

Des revues systématiques ont révélé des améliorations dans la conception de l'étude des essais randomisés en grappes, mais ont aussi souligné la persistance des lacunes, dans la mise en œuvre de ce type d'étude, qui peuvent affecter la validité des résultats (Díaz-Ordaz, Kenward,

Cohen, Coleman, & Eldridge, 2014; S. Eldridge et al., 2008; Suezann Puffer, David Torgerson, & Judith Watson, 2003). Le contexte joue un rôle fondamental dans ce type d'intervention, car des différences dans la mise en place de l'intervention peuvent varier d'une unité à une autre et devront être prises en considération. Des défis soulevés dans la phase d'implémentation devront être documentés par les chercheurs, afin de produire des résultats valables qui facilitent la prise de décision des parties prenantes, concernant le choix des interventions pour résoudre les problèmes auxquels les pays en développement sont confrontés. En effet, le niveau de preuve d'efficacité d'une intervention ne devrait pas être le seul critère pour les choix de décideurs. Pour prendre une décision éclairée, il faut aussi considérer les conditions de mise en œuvre d'une intervention et le contexte dans laquelle elle existe, pour mieux comprendre comment et pourquoi cette intervention produit des effets sur la santé de la population cible, avant de mettre en place ces interventions dans d'autres contextes ou de concevoir d'autres interventions. En somme, la compréhension de la mise en œuvre des interventions sous ce modèle en combinaison aux mesures statistiques pour atténuer les enjeux méthodologiques va renforcer la validité des résultats, et ainsi renforcer l'utilité et l'importance de ces études comme un outil clé pour fournir des évidences scientifiques.

1.2.1 L'évaluation des interventions dites « complexes » en santé publique

Les interventions « complexes » ont été largement utilisées dans la littérature comme raccourcis pour décrire des interventions construites à partir d'un certain nombre de composantes qui peuvent varier pendant la mise en œuvre, soit les participants, soit les personnes qui mènent les interventions, et soit tous les facteurs contextuels qui peuvent l'influencer (Craig et al., 2013). Les interventions dites complexes se caractérisent entre autres par la multiplicité des finalités et des niveaux d'analyse, ainsi que par le nombreux interrelations entre les acteurs ou participants et le contexte (N. C. Campbell et al., 2007).

Le Medical Research Council du Royaume-Uni (MRC) fournit des lignes directrices pour aider les bailleurs de fonds, les chercheurs et les décideurs à prendre des décisions appropriées concernant l'élaboration, l'évaluation et la mise en œuvre des interventions complexes. Le MRC a défini les interventions complexes comme des interventions qui se composent de plusieurs

éléments pouvant agir de façon indépendante ou interdépendante (Moore et al., 2015). Ces types d'interventions comprennent un ensemble d'actions pour résoudre une situation problématique, qui peut avoir plusieurs finalités et niveaux d'analyses, visant à agir sur des problèmes dans un contexte et dans un temps donné (M. Campbell et al., 2000). En effet, des interventions en santé publique agissent sur les déterminants de la santé pour influencer l'état de santé de la population. Cette interaction évoque une approche complexe, dynamique et en interaction constante avec les éléments contextuels où l'intervention est mise en place. Au fur et à mesure que l'intervention devient encore plus complexe, le risque des écarts par rapport à la conception de l'étude augmente (Horner, Rew, & Torres, 2006). En effet, les interventions complexes interagissent avec le contexte, ce qui signifie que leur efficacité dépendra de différents facteurs tels que les conditions socio-économiques et environnementales, le contexte de la politique et de la population cible.

Tel que mentionné dans la section précédente, les interventions de santé publique sont généralement complexes, programmatiques et contextuelles, et leur évaluation doit tenir compte de cette complexité (Craig et al., 2013; Rychetnik et al., 2002). Donc, l'évaluation d'une intervention en santé publique dans le cadre des essais randomisés en grappes doit être suffisamment exhaustive pour englober cette complexité, afin de déterminer si la réussite ou l'échec de l'intervention est dû à la conception de la théorie de l'intervention du programme ou à d'autres facteurs qui ont influencé la mise en place de l'intervention. Comprendre le contexte est donc essentiel pour interpréter les résultats d'une évaluation spécifique, ce qui permet de la généraliser à d'autres contextes. Même lorsqu'une intervention elle-même est relativement simple, son interaction avec son contexte peut être très complexe. En effet, des facteurs contextuels jouent un rôle potentiellement important, car l'efficacité d'une intervention ne peut être isolée du contexte environnemental dans laquelle elle intervient. Sous cette perspective, ces évaluations pourraient fournir des résultats pour de futurs développements d'interventions, mais aussi pour les chercheurs et les décideurs qui cherchent des informations scientifiquement valides et crédibles, afin d'appliquer la même intervention dans un autre contexte (Contandriopoulos, Rey, Brousselle, & Champagne, 2011; Craig et al., 2008).

1.3 Le but de l'évaluation et ses typologies

L'évaluation des interventions est largement reconnue comme un élément essentiel de toute intervention pour porter un jugement, pas seulement de l'intervention, mais de tout ce qui la concerne, afin d'aider à la prise de décision. La définition d'évaluation n'est pas nécessairement homogène et varie dépendamment de la génération où elle a été formulée (Brousselle et al., 2011). Les travaux de Madaus (1989), Guba et Lincoln (1984) ont regroupé ces générations par période qui sont caractérisées par des préoccupations ou des cibles différentes: la *génération I* est caractérisée par la mesure du succès à partir de différents tests; la *génération II* est caractérisée par la description pour expliquer les écarts entre les objectifs et les résultats finaux; la *génération III* est caractérisée par la description pour porter un jugement neutre sur l'objet évalué; et la *génération IV* est caractérisée par la négociation qui tient compte du point de vue de différents acteurs (Brousselle et al., 2011). Une génération plus récente est apparue, la *génération V*, laquelle est caractérisée par la négociation et le partenariat (Fontan, 2001). Cette génération implique une mobilisation d'une variété d'acteurs y compris des chercheurs universitaires qui utilisent l'évaluation comme un outil de gestion critique sur le contexte de la réalisation (Fontan, 2001).

Le concept de l'évaluation a évolué dans le but de mieux répondre au besoin d'information des décideurs, et son essence, comme un outil de vérification pour l'atteinte des objectifs et du développement des interventions, est toujours présente. Par exemple, Patton considère l'évaluation comme un processus qui examine de façon critique un programme, afin de porter un jugement sur le programme évalué, d'améliorer leur efficacité et/ou de déterminer le chemin à suivre (M. Q. Patton, 2014). Selon Brousselle et collègues, l'évaluation permet, à partir de la compilation des informations valides et légitimes, de prendre position sur l'intervention pour construire un jugement qui puisse se traduire en action (Brousselle et al., 2011). Dans la même ligne, Suchman considère l'évaluation comme un jugement sur la valeur ou l'utilité de l'action (Suchman, 1967). Pour d'autres, l'évaluation est considérée soit comme un moteur de changement (V. Ridde, Pommier, & Jabot, 2010), ou soit comme un outil de recherche, et non pas comme un synonyme de mesure de rendement (Bowen, 2012). D'autres auteurs conçoivent la définition de l'évaluation à l'intérieur du déroulement de l'intervention, en la définissant comme une méthode pour juger

l'état ou le degré d'adoption d'un programme (F. Champagne, & Denis, J. L., 1990), ou comme une stratégie pour ne pas juste porter un jugement, mais aussi pour pouvoir améliorer l'intervention (M. Patton, 2008). Bref, l'évaluation permet de décrire et d'évaluer entre autres, ce qui était prévu dans l'intervention, ce qui s'est passé et quels résultats ont été atteints. C'est ainsi que Pluye et collègues (2004) considèrent l'évaluation comme une composante essentielle de toute intervention, autant que la planification, la mise en œuvre et la pérennisation (P. Pluye, L. Potvin, et al., 2004).

Également, l'évolution de la notion du concept a apporté un ensemble de méthodes de recherche et de typologies associées ayant un but distinctif. De façon globale, pour Brousselle et collègues (2011), une intervention peut faire l'objet d'une évaluation normative ou de la recherche évaluative. L'évaluation normative vise à confronter les résultats à une norme, à une référence tandis que la **recherche évaluative** vise à explorer et à comprendre les relations de cause à effet qui existent entre les composantes d'une intervention (Brousselle et al., 2011). Sous un autre angle, la littérature énumère un certain nombre de différents types d'évaluation selon *le niveau, la finalité et la perspective* qui vise l'évaluation. Par exemple, *selon le niveau* où l'évaluation touche l'intervention, elle peut être classifiée comme une **évaluation du processus**, une évaluation d'impact, une évaluation des effets ou une évaluation économique (Richards & Hallberg, 2015; Valéry; Ridde & Dagenais, 2009; Steckler et al., 2002). Ensuite, *selon la perspective*, l'évaluation peut avoir un angle avec une vision de développement, de gestion ou expérimentale (Nebot, López, Ariza, Villalbí, & García-Altés, 2011). Enfin, *selon leur finalité*, l'évaluation peut être classifiée en deux grandes catégories: l'évaluation sommative et l'évaluation formative (Bowen, 2012; Valéry; Ridde & Dagenais, 2009). L'évaluation sommative porte un jugement sur l'efficacité des résultats, sur les effets ou sur l'utilité d'une intervention, tandis que **l'évaluation formative** vise à améliorer l'intervention dans la phase de développement, pour explorer la faisabilité et ainsi l'améliorer (Bowen, 2012; F. Champagne & Denis, 1992; Nebot et al., 2011; M. Patton, 2008; Valéry; Ridde & Dagenais, 2009). C'est sous cette dernière approche, *l'évaluation formative*, que l'évaluation de l'intervention s'inscrit dans le cadre de cette thèse.

En effet, selon la littérature, l'objectif principal de **l'évaluation formative**, aussi connue sous le nom de recherche formative, aide à créer des interventions plus efficaces dès les premières étapes de l'élaboration du programme. Cette évaluation a pour but de clarifier les objectifs, et

d'affiner et de guider le processus de la conception du projet, y compris la stratégie d'évaluation et les exigences en matière de données, pour l'améliorer et s'assurer que l'intervention soit acheminée à atteindre les résultats visés (Dehar, Casswell, & Duignan, 1993; Nebot et al., 2011; M. Patton, 2008; Stetler et al., 2006; Wynn et al., 2006). Patton et collègues (2008 : page 116-117) appellent ce type d'évaluation une « évaluation orientée vers l'amélioration », laquelle ne se focalise pas juste dans le processus, mais analyse aussi si tout progresse bien pour s'assurer des résultats désirés (M. Patton, 2008). Pour d'autres, cette évaluation permet de déterminer l'étendue, la fidélité et la qualité de la mise en œuvre d'une d'intervention, ainsi que de comprendre l'expérience des personnes directement touchées par les efforts de mise en œuvre (Havas, Anliker, Damron, Feldman, & Langenberg, 2000; Stetler et al., 2006).

Plusieurs études utilisent l'évaluation formative pour différentes raisons, mais toujours dans le but ultime d'améliorer l'intervention. On peut constater à travers différentes études le rôle essentiel qu'a joué cette évaluation dans l'amélioration de l'intervention. Gittelsohn et collaborateurs (1999) ont utilisé l'évaluation de type formative pour cibler les composantes de l'intervention de prévention de l'obésité en milieu scolaire, et pour identifier et classer les facteurs de risque du comportement afin d'élaborer des stratégies acceptables pour tous les sites ciblés (Gittelsohn et al., 1999). Cette évaluation, par exemple dans une étude pilote, peut favoriser la compréhension des forces structurelles, contextuelles et personnelles en jeu, afin de développer des interventions culturellement appropriées (Kumanyika et al., 2003). Pour d'autres, cette évaluation, en combinaison à une évaluation sommative, permet d'optimiser l'intervention pour envisager des effets à long terme et pour l'appliquer à d'autres contextes (Jurg, De Meij, Van der Wal, & Koelen, 2008). Également, cette évaluation permet de se concentrer sur l'identification des contraintes et des ressources déjà disponibles pour développer une intervention adaptée au système de santé. Par exemple, Napúa, M., et collègues (2016) ont utilisé ce type d'évaluation pour produire des données spécifiques de services de soin prénatal sur le volume et l'écoulement des patients dans les services, sur les ressources humaines, sur les ruptures de stock et sur le temps d'attente des patients (Napúa et al., 2016).

Bref, on peut constater que l'évaluation de type *formative* peut conduire à une meilleure compréhension des facteurs qui influent sur les résultats du programme (Dehar et al., 1993), car

elle est axée sur le processus interne de l'intervention afin de mieux comprendre ses forces, ses faiblesses et les changements qui se produisent dans le temps (M. Patton, 2008). Donc, étant donné l'intérêt des concepteurs de l'intervention à mettre en place l'intervention à grande échelle, *une évaluation* sous une perspective, *de type formative*, est totalement justifiée par ses multiples bénéfices déjà soulevés.

1.4 L'évaluation de la mise en œuvre d'une intervention

La recherche évaluative a conduit à une prolifération des approches d'évaluation qui peuvent être réalisées en fonction du but de l'évaluation (Richards & Hallberg, 2015; Steckler et al., 2002). Selon Brousselle et collègues (2011; p 63), une intervention peut être l'objet d'une évaluation en analysant « sa pertinence, sa cohérence, sa production, ses effets, son efficacité et sa mise en œuvre » (Brousselle et al., 2011). Des premières recherches sur la mise en œuvre se retrouvaient dans des écoles supérieures américaines avant les années 1970 pour aborder des questions dans des domaines de l'éducation, de la santé, environnementaux, sociaux et économiques sous une perspective initialement nationale, par une approche plus internationale et globale (Saetren, 2005). Ainsi, la mise en œuvre d'une intervention est une phase cruciale, qui devrait toujours faire l'objet d'une évaluation, car elle sert à plusieurs fins (Contandriopoulos et al., 2011; M. Patton, 2008). Par exemple, selon des études menées dans des pays africains, la mise en œuvre a été une approche adoptée dans plusieurs recherches, pour mieux analyser les écarts entre les politiques publiques et la pratique, qui ne se limite pas au domaine de la santé mais bien à d'autres secteurs (de Sardan & Ridde, 2014). Pour Ridde et collègues, le développement de la mise en œuvre est « complexe et dynamique », exigeant une évaluation qui permettra de mieux comprendre les différentes interactions qui se produisaient pendant sa phase de réalisation (Valéry Ridde & Dagenais, 2019).

Donc, plusieurs alternatives peuvent être considérées pour aborder l'évaluation de la mise en œuvre qui devra être guidée selon le but de l'évaluation. Cette évaluation est encore plus importante dans les pays du Sud, car les interventions de santé mises en œuvre dans ce contexte sont souvent financées par des acteurs de pays à revenu élevé, ce qui entraîne des luttes de pouvoir

et des multiples relations, qui influencent la mise en œuvre des interventions (Valéry Ridde, Pérez, & Robert, 2020). Un type d'évaluation, entre autres, est **l'évaluation du processus**, laquelle est utilisée pour surveiller la mise en œuvre d'une intervention, afin de mieux comprendre la relation entre les éléments spécifiques de l'intervention et leurs résultats (M. Patton, 2008). Pour entreprendre une telle évaluation, les chercheurs peuvent s'appuyer sur des théories de l'intervention et des cadres descriptifs, pour fournir des explications causales sur la manière dont différents facteurs peuvent influencer la mise en œuvre et les résultats (Valéry Ridde et al., 2020).

L'évaluation de processus vise à comprendre les facteurs qui ont eu un incidence sur la mise en place de l'intervention et aussi sur les résultats, dans le but d'améliorer ce processus (M. Patton, 2008; Weiss, 1997). Selon Weiss, on peut s'appuyer sur la théorie de l'intervention, laquelle se focalise à expliquer le chemin causal qui vise à obtenir les effet prévus à partir du plan du programme, et de la manière dont les participants réagissent (Weiss, 1997). Pour Chen, cette évaluation se voit favorisée avec l'inclusion des parties prenantes tout au long du processus d'évaluation, pour bien concevoir la collecte de données et l'utilisation des informations dans l'évaluation du programme (H. T. Chen & Chen, 2005). Durlak et DuPre (2008) ont suggéré d'accorder une attention particulière à trois autres aspects de l'étude de la mise en œuvre, à savoir: (i) le suivi du contrôle, c'est-à-dire la nature et la quantité des services reçus par les participants; (ii) la portée du programme, à savoir le taux d'implication et la représentativité des participants; et (iii) l'adaptation, c'est-à-dire les changements et les modifications apportés au cours de la mise en œuvre (Durlak & DuPre, 2008). Pour certains auteurs, l'évaluation du processus est fondamentale, car elle améliore l'interprétation des données recueillies, et aide les évaluateurs à éviter de commettre l'erreur de type III (Bradley, Wiles, Kinmonth, Mant, & Gantley, 1999; Dobson & Cook, 1980; Steckler et al., 2002). Une erreur de type III se produit quand on conclut qu'une intervention n'est pas efficace, alors qu'en réalité c'est plutôt la mise en œuvre défectueuse que l'intervention elle-même qui a conduit à l'échec, tel que nous l'avons mentionné auparavant.

Bien qu'il n'y ait pas un consensus dans la littérature sur la conceptualisation de l'évaluation du processus, tous les chercheurs convergent sur l'importance de son évaluation, en recourant à différentes approches méthodologiques dans le but de mieux comprendre les effets de l'intervention. À ce propos, l'évaluation du processus aide les parties prenantes à expliquer les

succès, les échecs et les changements d'une intervention, car elle vise à analyser les différents facteurs qui peuvent agir sur l'intervention et influencer leurs résultats (F. Champagne & Denis, 1992; M. Patton, 2008). L'évaluation du processus peut être développée de différentes façons selon la question de recherche à laquelle on veut répondre (M. Patton, 2008; Richards & Hallberg, 2015). Le Medical Research Council du Royaume-Uni soulève que l'évaluation du processus peut être utilisée pour évaluer la théorie de l'action à partir de laquelle l'intervention a été développée, pour évaluer la fidélité et la qualité de la mise en œuvre, pour clarifier le mécanisme causal ainsi que pour explorer le contexte dans lequel l'intervention a été développée (Craig et al., 2008; Moore et al., 2015).

L'évaluation du processus va au-delà de l'évaluation d'impact d'une intervention qui est basée sur le modèle de la « boîte noire » (H.-T. Chen & Rossi, 1989) . Cette évaluation d'impact conçoit une intervention insensible aux facteurs contextuels au moment de la mise en œuvre, alors que l'évaluation du processus s'intéresse à fournir l'information des conditions qui façonnent et influencent les résultats. Cette évaluation du processus décrit la qualité d'une intervention, l'expérience des personnes exposées, le contexte, la portée, la fidélité et l'acceptabilité de l'intervention (Bradley et al., 1999; F. Champagne & Denis, 1992; Hulscher et al., 2003; Moore et al., 2015; M. Patton, 2008; Steckler et al., 2002). Une révision systématique a souligné l'importance de tenir compte des facteurs contextuels dans l'évaluation des essais randomisés en grappes, car ces facteurs ont des effets importants sur les résultats de l'intervention (Blacklock et al., 2016). Donc, il est nécessaire d'évaluer si l'intervention a été développée comme elle a été programmée (**évaluation de la fidélité**), si des adaptations ou des modifications ont été faites, et si tous les moyens ont été utilisés pour leur mise en œuvre (Steckler et al., 2002). Cette évaluation pour Valéry et collègues fait référence à l'évaluation de la mise en œuvre qui est plus « normative », comparativement à l'évaluation du processus qui est plus « dynamique » (Valéry Ridde & Dagenais, 2019). Cependant, évaluer la fidélité de la mise en œuvre implique non seulement recueillir des données sur le programme lui-même, mais aussi recueillir des données décrivant l'environnement du programme. Cela est particulièrement important pour des interventions multi sites, afin de ne pas tirer une conclusion erronée sur la raison pour laquelle le même modèle d'intervention produit des résultats différents dans différents sites (Orwin, 2000).

Bien qu'il n'y ait pas de consensus sur la meilleure façon de diviser l'étude de la mise en œuvre en sous-composantes clés, le Medical Research Council du Royaume-Uni recommande une phase de faisabilité et de pilotage après le développement d'une intervention (Moore et al., 2015). À ce stade, l'évaluation du processus peut jouer un rôle essentiel pour comprendre la faisabilité de l'intervention et optimiser sa conception, tandis que l'évaluation de la fidélité peut fournir l'information sur comment l'intervention a été livrée, pour aider les chercheurs et les décideurs à déterminer si l'intervention peut être répliquée à grande échelle (Moore et al., 2015; Steckler et al., 2002). L'évaluation de la fidélité constitue « une étape préalable à l'évaluation du processus », afin de mieux comprendre tous les facteurs qui ont contribué à la mise en œuvre observée (V Ridde & Haddad, 2013). Donc, cette évaluation constitue une forme ciblée d'évaluation du processus, car elle possède une composante descriptive pour décrire la mise en œuvre du programme, ainsi qu'une composante d'évaluation normative, puisqu'elle comporte des attentes de performance ou d'adhésion à un plan d'action prescrit (Orwin, 2000). Au stade de la faisabilité ou du pilotage d'une intervention, des mesures quantitatives de base de la mise en œuvre peuvent être combinées à des données qualitatives approfondies pour fournir une compréhension détaillée du fonctionnement de l'intervention à petite échelle, et offrir ainsi aux chercheurs la possibilité d'être plus confiants dans leurs résultats (M. Campbell et al., 2000; Grant et al., 2013; Richards & Hallberg, 2015).

En somme, une évaluation de la mise en œuvre, qui s'intéresse particulièrement à la dimension de la fidélité de la mise en œuvre, est tout à fait justifiée et pertinente pour déterminer si l'intervention a été mise en œuvre de la façon dont elle a été conçue, avant de pouvoir y attribuer des effets et d'affirmer son efficacité. Cette évaluation constitue une étape essentielle pour amener les utilisateurs à une meilleure interprétation des résultats (H. T. Chen, 2014). En effet, une meilleure connaissance sur la fidélité de l'intervention peut contribuer à la conception d'une intervention mieux orientée vers leurs besoins. En retour, cette meilleure compréhension aide à promouvoir la mise en place de l'intervention à grande échelle afin d'atteindre le but poursuivi.

1.5 L'évaluation de la fidélité de la mise en œuvre des interventions en santé publique

La notion d'évaluation de la fidélité de la mise en œuvre de l'intervention est définie et appliquée différemment, ce qui entraîne souvent une confusion pour l'aborder. La définition la plus communément acceptée concernant cette pratique évaluative est celle qui vise à mesurer le degré auquel un traitement / intervention est mis en œuvre comme prévu par les concepteurs (B. Borrelli, 2011; Carroll et al., 2007; Dane & Schneider, 1998). Donc, cette évaluation vise à mesurer si les composantes de base d'une intervention ont été mises en œuvre telles que prévues par les chercheurs. Des termes comme intégrité au traitement, conformité au traitement, fidélité à l'intervention, intégrité procédurale, qualité de mise en œuvre et intégrité de l'intervention sont souvent utilisés de manière interchangeable, pour signifier le degré auquel les interventions sont livrées telles que conçues et planifiées par les chercheurs.

La notion de fidélité de la mise en œuvre n'est pas un nouveau concept. Cette notion a émergé dans les années 1900 comme une approche pour évaluer et documenter l'intégrité du traitement, et comme un concept clé au moment d'évaluer des interventions pour mesurer leur efficacité dans le domaine psychiatrique (Moncher & Prinz, 1991). Puis, la notion de fidélité des interventions a été largement examinée dans différents domaines pour mieux expliquer les résultats obtenus dus à son potentiel d'ouvrir complètement la boîte noire, en mesurant les processus conduisant à la mise en œuvre et aux résultats. Il est reconnu qu'une mise en œuvre réussie peut et doit, de manière optimale, inclure à la fois la dispensation et la réception du traitement, car les deux sont généralement nécessaires pour changer le comportement ciblé (Orwin, 2000). Donc, si la programmation destinée à faciliter l'atteinte des résultats de l'intervention n'est pas mise en œuvre comme prévue, les mesures des résultats ne peuvent pas bien révéler l'efficacité du programme prévu. De là, l'importance de cette évaluation pour renforcer les résultats d'une intervention.

Au cours des dernières décennies, les disciplines des sciences sociales et comportementales ont pris en compte la fidélité de la mise en œuvre dans un large éventail de contextes, notamment les interventions en réadaptation psychiatrique (Bond, Evans, Salyers, Williams, & Kim, 2000), les

interventions comportementales en milieu scolaire (Dusenbury et al., 2003), les interventions sur les comportements de la santé (Belinda Borrelli et al., 2005) et les interventions psychosociales (McArthur, Riosa, & Preyde, 2012). Bien que l'évaluation de la fidélité de la mise en œuvre soit un domaine d'étude en pleine croissance, les résultats de ces études suggèrent que la fidélité de la mise en œuvre des interventions est toujours un domaine de la recherche scientifique qui nécessite une attention particulière. Par exemple, Durlak et Wells (1997) ont constaté dans une révision de plus de 1 200 études publiées, qui visent des interventions de prévention des maladies, que seulement 5% des études ont fourni des données sur la fidélité de la mise en œuvre (Durlak & Wells, 1998). Par sa part, Dane et Schneider (1998) ont constaté que seulement 24% (39/162) des interventions examinées contenaient des informations sur la fidélité du programme, et que seulement 33% (13/39) de ces études ont examiné l'impact de la fidélité sur les résultats. (Dane & Schneider, 1998).

L'expansion de cette pratique à différents domaines fait de ce concept une notion multiforme sans une approche unifiée entre les disciplines de recherche, conduisant les chercheurs à employer une variété de termes ainsi que des méthodes plus variées pour évaluer la fidélité de la mise en œuvre. En effet, plusieurs revues de littérature soulèvent que le terme fidélité varie selon le domaine d'étude (en général l'éducation, la psychologie et la santé publique), ainsi que selon les outils pour la mesurer (Corley & Kim, 2016; Gearing et al., 2011; Ibrahim & Sidani, 2015; O'Donnell, 2008; Slaughter, Hill, & Snelgrove-Clarke, 2015). Cependant, parmi ces auteurs, il y a un consensus à propos de ce qui est prévu lorsque la recherche se réfère à l'évaluation de la fidélité de la mise en œuvre. Cette évaluation cherche à évaluer (i) si la mise en œuvre est conforme à la théorie (ii), si la dose de l'exécution du programme pendant leur mise en œuvre est mesurée, et (iii) si des facteurs externes affectent le niveau de la fidélité obtenu.

L'adhésion à l'intervention est définie, si le programme est fourni tel qu'il a été conçu ou s'il est conforme au protocole de l'intervention (S. Mihalic, 2004), en évaluant les composantes essentielles de l'intervention, quant au contenu et à la dose de l'intervention reçus par les participants aussi souvent et aussi longtemps que prescrits (Carroll et al., 2007). Concernant l'adhésion qui est souvent mesurée (Paquette, Joly, & Tourigny, 2009), une révision de la littérature a rapporté que cette dimension est souvent mal mesurée, mal rapportée ou souvent absente dans les

interventions dites complexes (Dusenbury et al., 2003). L'évaluation de la fidélité de la mise en œuvre est cruciale pour les études d'efficacité, en particulier lorsque les chercheurs veulent savoir si les mauvais résultats du programme sont dus à un manque de la mise en œuvre ou à une intervention mal conçue ou mal alignée avec la théorie du programme. En effet, une intervention peut avoir des effets limités, soit en raison de faiblesses dans sa conception théorique ou soit parce qu'elle n'est pas correctement mise en œuvre (Moore et al., 2015). Donc, étudier la pratique évaluative concernant la fidélité de la mise en œuvre est fondamental dans les interventions de santé publique qui sont « complexes », pouvant varier pendant leur mise en œuvre, afin de mieux comprendre et de renforcer les résultats de l'intervention.

Au fur et à mesure que les programmes sont développés et validés dans des études d'efficacité, les développeurs doivent s'assurer que les ingrédients actifs (les composantes spécifiques à l'intervention servant de leviers clés du changement) sont explicitement identifiés pour faciliter la mise en œuvre de l'intervention (Sidani & Sechrest, 1999). Donc, si un élément de base n'est pas lié aux résultats escomptés, il peut être considéré comme moins important ou non essentiel. Un élément clé pour garantir la fidélité du programme est de s'assurer que le modèle du programme et ses ingrédients actifs sont clairement définis dès le départ, afin que ce modèle du programme soit bien compris par les intervenants (Abry, Hulleman, & Rimm-Kaufman, 2015; Bond et al., 2000). Les pratiques de fidélité au traitement liées à la conception de l'étude visent à garantir qu'une étude peut tester correctement ses hypothèses, par rapport à la théorie sous-jacente et aux processus cliniques (Resnick et al., 2005). Cependant, la mesure de l'adhésion ne devrait pas être suffisante pour bien expliquer les résultats obtenus. Pour mieux comprendre comment et pourquoi une intervention fonctionne, et dans quelle mesure les résultats peuvent être améliorés, il est nécessaire d'explorer tous les facteurs qui peuvent l'affecter (Carroll et al., 2007). Donc, il ne s'agit pas de simplement documenter les échecs de la mise en œuvre. Une approche beaucoup plus utile pour les chercheurs consiste à: (i) vérifier que la théorie de l'intervention soit bien comprise par ceux qui vont la mettre en place; (ii) utiliser des mesures de fidélité; et (iii) explorer les facteurs qui peuvent modifier la fidélité de la mise en œuvre, permettant ainsi d'établir l'efficacité d'une intervention. Cela va aider à guider le développement et à maintenir les programmes sur la bonne voie.

En outre, la littérature fait ressortir l'importance d'aller au-delà de l'intégrité du traitement, c'est-à-dire ne pas seulement déterminer si un traitement est mis en œuvre comme prévu, mais aussi de savoir si les conditions de traitement diffèrent les unes des autres de la manière prévue. La « différenciation » permet de déterminer le degré auquel un traitement peut être distingué d'un autre (Schoenwald, Henggeler, Brondino, & Rowland, 2000). Pour Moncher et collègues (1991), le degré de différenciation est important à considérer, car il est possible que les traitements diffèrent, mais qu'ils ne soient toujours pas mis en œuvre comme prévus, ou que les traitements individuels aient une intégrité élevée, mais qu'ils ne soient pas suffisamment distincts (Moncher & Prinz, 1991). Donc, déterminer s'il existe des différences importantes entre le groupe qui reçoit l'intervention et celle du groupe contrôle est aussi important pour établir une différence dans les résultats (Morello et al., 2017; O'Donnell, 2008).

Dans les cas des essais randomisés en grappes, les résultats peuvent être dilués par des effets de « contamination ». Par exemple, si les participants du groupe témoin participent à des activités du groupe d'intervention ou à d'autres services similaires à l'intervention étudiée, ils vont diminuer les effets observés. Ainsi, des mesures de fidélité doivent être utilisées pour évaluer les conditions expérimentales et de contrôle. Du côté expérimental, la question est de savoir si le programme a été mis en œuvre comme prévu, ou s'il ressemble en réalité aux pratiques habituelles de prestation de services. Du côté contrôle, la question est généralement de savoir si le groupe témoin a été contaminé en imitant les stratégies du programme expérimental (Bond et al., 2000). Donc, l'évaluation de la fidélité peut être utile pour explorer pourquoi une intervention fondée sur une étude pilote réussie n'atteint pas les mêmes résultats durables dans le projet à grande échelle (R. Allison et al., 2017).

Ainsi, évaluer la fidélité d'intervention est aussi fondamentale dans un contexte d'un essai randomisé, car des différences contextuelles au niveau de chaque grappe pourraient favoriser l'adaptation locale de l'intervention. Ces adaptations peuvent conduire à l'hétérogénéité de la fidélité d'intervention (Mann et al., 2016). Cette hétérogénéité est aussi importante à explorer pour comprendre le niveau de fidélité atteint et le contexte dans lequel se développe l'intervention (Davidoff, 2009). L'évaluation de la fidélité d'intervention peut aider à évaluer l'ampleur de ces problèmes, et à en atténuer les effets négatifs qui peuvent affecter la validité interne et externe des

résultats des essais randomisés en grappes. Enfin, pour les interventions en santé publique mises en place dans le cadre des essais randomisés en grappes, il est tout à fait pertinent et important d'entreprendre ce type d'évaluation, et cela pour ses multiples avantages, autant du plan opérationnel que méthodologique, qui seront abordés dans la prochaine section.

1.5.1 Avantages pour les chercheurs d'évaluer la fidélité de la mise en œuvre de l'intervention

En plus de tout ce qui a déjà été mentionné dans la section précédente, l'importance d'étudier la fidélité de la mise en œuvre est largement reconnue, dans la littérature, pour une variété de raisons : (i) elle évite de commettre l'erreur de Type III (évaluer une intervention qui n'a pas été mise en œuvre telle que prévue), afin d'empêcher de tirer des conclusions potentiellement fausses sur l'efficacité d'une intervention; (ii) elle permet souvent d'expliquer pourquoi les innovations ont été un succès ou un échec ; (iii) elle permet aux chercheurs d'identifier ce qui a été changé dans un programme et comment les changements ont des répercussions sur les résultats; (iv) elle révèle des informations importantes sur la faisabilité d'une intervention et recommande les futures modifications à l'intervention si nécessaire; (v) elle explique pourquoi certaines études génèrent des résultats différents, même si elles semblent viser les mêmes effets; (vi) elle permet en même temps de réduire les coûts financiers et sociaux investis dans la mise en œuvre, en reconnaissant les moyens par lesquels se sont produits ou non les effets; (vii) elle permet d'identifier les ingrédients essentiels du programme, et d'identifier les éléments critiques qui déterminent la réussite d'un programme; enfin (ix) cette évaluation va assurer aux chercheurs que les services sont mis en œuvre comme le souhaitaient leurs bailleurs de fonds.(Bond et al., 2000; Carroll et al., 2007; Corley & Kim, 2016; Dusenbury et al., 2003; Gearing et al., 2011; O'Donnell, 2008; Orwin, 2000; Slaughter et al., 2015).

L'évaluation de la fidélité aide les chercheurs à savoir si l'intervention a fonctionné et à comprendre les détails nuancés de la mise en œuvre de l'intervention concernant le qui, le quoi, le quand, le où, et le pourquoi (Corley & Kim, 2016). De même, l'évaluation de la fidélité donne des avantages à la recherche primaire, car elle renforce les résultats obtenus de l'intervention testée.

De même, elle donne des avantages pour la synthèse des résultats lors des révisions systématiques, car en évaluant des études qui ont tenu compte de cette évaluation, les auteurs seront plus confiants au moment de la synthèse des études, en minimisant les biais inhérents des études originales, et en renforçant ainsi les arguments pour la prise de décision et la production des recommandations issus de la revue systématique.

Bien que les chercheurs comprennent clairement les dangers d'une erreur de type I (un résultat faux positif ou rejeter incorrectement une hypothèse nulle qui est vraie), et l'erreur de type II (résultat faux négatif ou ne pas rejeter l'hypothèse nulle lorsqu'elle est fausse), un autre type d'erreur important, l'erreur de type III, a reçu peu d'attention. Une erreur de type III a été décrite comme obtenant la bonne réponse pour la mauvaise raison, en ce qui concerne la relation avec la mesure de la fidélité, lorsqu'une incompatibilité se produit entre l'intervention envisagée et l'intervention livrée. Les chercheurs finissent par évaluer quelque chose qui ne s'est pas produit, ce qui confond la question de recherche réelle (erreur de type III). Lorsqu'une erreur de type III se produit, les résultats de la recherche peuvent fournir des informations trompeuses. Par exemple, les parties prenantes peuvent décider d'arrêter des programmes efficaces sur la base de résultats négatifs qui découlent d'une mauvaise mise en œuvre (en raison de la fidélité inadéquate de l'intervention prévue) plutôt que d'une intervention inefficace. Sans une évaluation appropriée de la fidélité de la mise en œuvre de l'intervention, les chercheurs et les cliniciens ne peuvent pas déterminer si le manque d'impact est dû à l'intervention elle-même (par exemple, mal conceptualisée, mauvaise théorie ou intervention invalide) ou à un manque de compréhension de l'intervention à la phase de la mise en œuvre. Cela fait de cette pratique évaluative un domaine riche à privilégier par les chercheurs dans la planification des évaluations des interventions en santé publique pour aider à interpréter les résultats de l'étude (Tableau1), sans oublier tous les autres domaines qui peuvent en bénéficier.

Tableau 1. – Interprétation de l'étude selon l'évaluation de la fidélité de la mise en œuvre de l'intervention

Impact attendu de l'intervention	Niveau de fidélité de la mise en œuvre de l'intervention	Ce qu'on peut déduire
Positif	Haut niveau de fidélité	L'intervention peut contribuer à atteindre les résultats escomptés, car elle a été mise en œuvre comme prévue et les résultats ont été atteints, de sorte que l'intervention peut être efficace.
Positif	Bas niveau de fidélité	L'intervention telle que prévue ne peut pas être créditée d'avoir contribué aux résultats, car elle n'a pas été mise en œuvre comme prévue, et donc il est difficile de déclarer que le programme prévu était efficace.
Négatif	Haut niveau de fidélité	Le programme a été mis en œuvre comme prévu, mais les résultats escomptés n'ont pas été observés. Les résultats de l'évaluation des résultats devraient contribuer à apporter des changements au programme planifié par les parties prenantes, car une faible fidélité de mise en œuvre peut être exclue comme la raison des mauvais résultats.
Négatif	Bas niveau de fidélité	Si le programme formulé n'a pas été mis en œuvre comme prévu et que les résultats escomptés n'ont pas été observés, il faudra (a) penser à une nouvelle étude qui est menée avec une fidélité accrue à la mise en œuvre, pour évaluer l'efficacité du programme et de son fondement théorique avant de déterminer son inefficacité; ou (b) revoir la pertinence théorique

		du programme pour déterminer si les activités et les stratégies mises en place correspondent au fondement théorique du but ciblé par l'intervention.
--	--	--

1.5.2 Avantages d'évaluer la fidélité du point de vue méthodologique

L'évaluation de la fidélité est un élément essentiel pour augmenter la validité interne et la validité externe de l'intervention, deux aspects critiques pour répliquer et généraliser toute intervention (B. Borrelli, 2011; Ibrahim & Sidani, 2015; O'Donnell, 2008). Ainsi, cette évaluation favorise la validité théorique et la puissance statistique des résultats de l'intervention (Moncher & Prinz, 1991). Par exemple, quand on teste une intervention, si des résultats significatifs ont été trouvés, mais que la fidélité n'a pas été vérifiée, le résultat pourrait être dû à une intervention efficace ou à des contaminants inconnus, ajoutés à l'intervention. Inversement, avec des résultats non significatifs et aucune information sur la fidélité, on ne sait pas si l'intervention a été inefficace ou mal mise en œuvre, tel que nous l'avons fait ressortir dans la section précédente. Donc, cette évaluation permet de savoir si les effets sont dus à l'intervention elle-même ou s'ils sont dus à d'autres facteurs externes à l'intervention (la validité interne).

D'un autre côté, la réplification de l'étude et la comparaison entre les études similaires sont entravées par des descriptions inadéquates de la théorie de l'intervention, par des changements dans la mise en œuvre qui ne sont pas documentés, et par des données insuffisantes sur l'intervention telle qu'elle a été réellement mise en œuvre. Donc, une description détaillée du protocole de l'intervention et la compréhension même de l'intervention par le personnel qui met en place l'intervention vont prévenir des problèmes d'interprétation et de confusion associés à d'autres variables (comme la durée d'une activité, les caractéristiques du personnel qui met en place l'intervention, les intérêts des participants, par exemple), qui vont affecter le degré de fidélité de l'intervention. Donc, l'établissement de la validité conceptuelle (validité théorique) d'une intervention est une étape essentielle et critique pour assurer de mettre en place l'intervention telle que planifiée. Ces éléments ajoutés à une description détaillée de la mise en œuvre vont fournir une feuille de route pour la reproductibilité de l'intervention dans un autre contexte (la validité

externe). En effet, un manque de données sur la mise en place de l'intervention limite la capacité de reproduire et de diffuser l'intervention, ce qui rend difficile, pour les chercheurs, de dupliquer même les interventions menées avec succès.

La fidélité de la mise en œuvre a également des implications importantes sur la puissance statistique d'une intervention. Si la fidélité est élevée, la probabilité de changement du comportement visé est élevée. Lorsque le contenu de l'intervention n'est pas uniforme entre les intervenants et que l'investigateur a traité la variable de l'intervention de manière dichotomique (c'est-à-dire intervention reçue versus absence d'intervention), le résultat est une diminution de la puissance et une plus grande probabilité d'erreur de type II (refuser l'efficacité de l'intervention alors que l'intervention aurait pu être utile). (Bond et al., 2000; Moncher & Prinz, 1991; O'Donnell, 2008; Shadish, Cook, & Campbell, 2002). Le suivi et l'optimisation de la fidélité de l'intervention au cours d'une série d'études peuvent augmenter la taille de l'effet, et réduire le nombre de sujets requis lors d'études ultérieures, réduisant ainsi les coûts et améliorant l'efficacité d'un programme de recherche interventionnelle (Bellg et al., 2004).

En plus de son importance pour la recherche scientifique, la documentation de la fidélité est importante pour des raisons pratiques et éthiques. La surveillance de la mise en place de l'intervention favorise la détection précoce et la correction des erreurs, ce qui à son tour réduit les coûts et améliore le traitement. Le coût d'une fidélité inadéquate peut être le rejet de programmes de traitement puissants ou l'acceptation de programmes impuissants. De plus, une mise en œuvre inadéquate de l'intervention peut même nuire aux participants ciblés par l'intervention. Tous ces éléments confirment que l'évaluation de la fidélité de la mise en œuvre d'une intervention est une condition essentielle de l'évaluation du processus pour les programmes qui visent la reproductibilité dans d'autres contextes, et son importance est accentuée dans un contexte où les ressources sont limitées, comme dans le cas des pays du Sud.

Bref, bien que les essais randomisés en grappes soient des études clés pour déterminer la relation causale entre une intervention et ses résultats présumés, la conception de l'étude, à elle seule, ne peut suffire comme critère principal pour leur attribuer leur crédibilité maximale comme outil clé en tant qu'évidence des données probantes. Dans le cas des interventions de santé

publique, étant donné l'existence des facteurs qui peuvent influencer la mise en œuvre de l'intervention, l'évaluation de la fidélité peut contribuer à améliorer et à accroître sa performance autant sur le plan pratique que méthodologique.

1.5.3 Des critères pour évaluer la fidélité de la mise en œuvre des interventions

Un instrument de fidélité universelle ne peut pas être possible parce que les mesures de fidélité doivent être conçues avec la théorie de programme spécifique (Ibrahim & Sidani, 2015). Cependant, il est conseillé aux chercheurs de suivre un cadre ou un ensemble de lignes directrices pour développer ces mesures, et pour procéder à une évaluation complète de la fidélité avec laquelle les interventions sont mises en œuvre (Ibrahim & Sidani, 2015; O'Donnell, 2008). Plusieurs auteurs ont établi des guides pour évaluer la fidélité de la mise en œuvre de l'intervention qui ont été suggérés/modifiés et sont disponibles dans la littérature (B. Borrelli, 2011; Belinda Borrelli et al., 2005; Carroll et al., 2007; Dabbs et al., 2011; Dane & Schneider, 1998; Gearing et al., 2011; Hasson, 2010; D. Perez, Van der Stuyft, Zabala, Castro, & Lefevre, 2016).

Dans le cadre de fidélité de la mise en œuvre de l'intervention, Dane et Schneider (1998) ont identifié cinq aspects essentiels à la mise en œuvre de toute innovation, à savoir: (i) la conformité: si les composantes du programme sont exécutées comme prévues; (ii) le dosage: le nombre, la fréquence et le temps consacrés à la mise en œuvre; (iii) la qualité: la manière dont les composantes du programme sont exécutées; (iv) la réactivité: la mesure dans laquelle le programme stimule l'intérêt des participants; et (v) la différenciation du programme: la mesure dans laquelle le programme est unique et différent des autres interventions (Dane & Schneider, 1998).

Pour leur part, Borrelli et collègues (2005, 2011) ont développé un guide pour évaluer la fidélité au traitement des essais randomisés en grappes qui visent un changement de comportement en matière de santé. À partir d'une série de quinze études du groupe Behavior Change Consortium, les chercheurs ont suggéré des lignes directrices comprenant l'évaluation de la fidélité du traitement en ce qui concerne: (i) la conception de l'étude: visant à garantir qu'une étude puisse tester correctement ses hypothèses par rapport à la théorie sous-jacente; (ii) la formation des

interventionnistes: pour s'assurer qu'ils aient été formés de manière satisfaisante pour fournir l'intervention aux participants à l'étude; (iii) la livraison: pour contrôler que l'intervention soit réalisée comme prévu; (iv) la réception: pour garantir que le traitement ait été reçu et compris par les participants; et (v) l'adoption de l'intervention: pour surveiller que la personne exécute les compétences comportementales et les stratégies cognitives liées à l'intervention dans les contextes de la vie réelle, comme prévu. (B. Borrelli, 2011; Belinda Borrelli et al., 2005).

Dans le cadre de la recherche en intervention en travail social, Corley et Kim (2016) ont voulu évaluer si les études ont relevé des problèmes liés à la fidélité de la mise en œuvre des interventions à partir des critères développés par les chercheurs, pour évaluer si les études rapportent la fidélité au traitement de l'étude. Les résultats de cette étude suggèrent que le rapport de la fidélité de la mise en œuvre des interventions dans la discipline de travail social a suivi une légère amélioration dans le temps, mais que cette évaluation est toujours un domaine qui nécessite une attention particulière. Par exemple, 65% (21/32) des études ont effectivement utilisé le terme de fidélité de la mise en œuvre (ou un terme similaire: intégrité ou adhésion au traitement) dans l'évaluation des résultats (Corley & Kim, 2016). Pour améliorer cette pratique, les auteurs ont suggéré une liste de contrôle avec huit questions pour guider les chercheurs à répondre au moment d'évaluer une intervention dans une étude ciblée: (i) l'intervention a-t-elle été pré-testée sur un sous-ensemble de l'échantillon de la population?; (ii) l'éthique a-t-elle été mentionnée?; (iii) des manuels ou protocoles de traitement ont-ils été utilisés?; (iv) une formation a-t-elle été offerte aux praticiens?; (v) la supervision des praticiens a-t-elle été offerte?; (vi) une description détaillée de l'intervention a-t-elle été fournie?; (vii) le terme fidélité de la mise en œuvre de l'intervention a-t-il été utilisé?; (viii) des détails spécifiques ont-ils été fournis sur la façon dont l'intervention a été surveillée?.

Par sa part, Carroll et collègues (2007), à partir d'une révision exhaustive de la littérature, ont compilé les dimensions soulevées par la littérature et en ont ajouté d'autres, en proposant un cadre conceptuel pour expliquer comment la fidélité de la mise en œuvre d'une intervention doit être évaluée (Carroll et al., 2007). Selon ce cadre de référence, la fidélité de la mise en œuvre de l'intervention se mesure par l'adhésion au programme, c'est-à-dire si le programme a été mis en œuvre de la façon dont il a été conçu. L'adhésion se divise en quatre dimensions: contenu,

fréquence, durée et couverture. Ainsi, l'adhésion dépend de la question de savoir si les participants ont reçu les composantes actives de l'intervention aussi souvent et aussi longtemps que prévu initialement. Il s'agit également de savoir si toutes les personnes qui participaient ou bénéficiaient des avantages d'une intervention sont atteintes. En effet, la couverture fait référence aux sujets abordés dans le cadre du programme. La fréquence fait référence au nombre de périodes consacrées au programme sur une période de temps donnée. La durée correspond au temps accordé au programme. Enfin, le contenu correspond au degré d'application du programme.

Par ailleurs, selon ce cadre de référence, certains éléments modérateurs peuvent influencer le niveau de fidélité de la mise en œuvre d'un programme. Donc, les auteurs suggèrent également d'évaluer la description détaillée de l'intervention, les stratégies pour faciliter la mise en œuvre, la qualité de la prestation et la réponse des participants. Ces facteurs peuvent interagir entre eux en affectant l'implémentation de l'intervention, et donc sa fidélité. Par exemple : (i) la complexité du programme : plus un programme est simple et détaillé, plus il risque d'être mis en œuvre fidèlement par les intervenants; (ii) les stratégies de facilitation de la mise en œuvre : un programme qui s'accompagne d'outils ou de guides favorisant sa mise en œuvre est plus susceptible d'être mis en œuvre de la façon dont il a été conçu; (iii) la qualité de la mise en œuvre : soit la façon dont les intervenants mènent les activités, des intervenants motivés et engagés envers un programme augmentent les chances de ce dernier d'être mis en œuvre fidèlement. Finalement, le dernier facteur modérateur concerne la réponse des participants : plus les participants sont engagés et participent au programme, plus ses chances d'être mis en œuvre fidèlement augmentent.

Le cadre de référence de Carroll et collègues (Carroll et al., 2007) a été amélioré par Hasson (2010) en ajoutant deux éléments au niveau des facteurs modérateurs: le recrutement et le contexte (Hasson, 2010). En effet, ces dimensions ont été ajoutées, car l'importance de prendre en compte le contexte et les liens entre les acteurs, les organisations et les participants, par exemple, qui influencent la mise en œuvre du programme, est bien connue. De même, les procédures, qui ont été utilisées pour attirer des participants potentiels au programme, peuvent influencer la non-participation parmi les participants potentiels de l'étude, en influençant la mise en œuvre de l'intervention et ainsi la généralisation des résultats de l'étude. Par exemple, si l'intervention se déroule dans un contexte où les femmes se sentent plus à l'aise de participer en voyant des femmes,

une équipe de recrutement des femmes pourraient faciliter leur participation dans les différentes activités. Ainsi, toutes les dimensions qui peuvent influencer la fidélité d'une intervention pour avoir une meilleure compréhension des effets à la suite de la mise en place d'une intervention devront être évaluées.

D'autres modifications ont été apportées au cadre de Carroll et collègues pour le rendre utile pour les interventions adaptatives. Ce type d'intervention est définie comme des interventions pour lesquelles les parties prenantes sont autorisées, voire encouragées, à apporter des modifications à la conception originale d'une intervention existante (D. Perez et al., 2016). Perez et collègues (2016) ont constaté que les dimensions de la fidélité classique n'abordent pas la question de savoir comment adapter une intervention tout en conservant son efficacité pour les interventions dites « adaptatives », tout en conservant en même temps ses composantes actives. (D. Perez et al., 2016). Pour cette raison, et en considérant que, pour ce type d'intervention, les deux concepts, fidélité vs adaptabilité, sont intrinsèquement liés, les auteurs ont suggéré d'évaluer les informations dans ce type d'intervention en tenant compte : (a) de l'idée claire des résultats escomptés; (b) des principes de fonctionnement ou de la théorie du changement; (c) des descripteurs spécifiques de fidélité; et (d) des questions pour identifier les adaptations basées sur la description de l'intervention. Finalement, sous cette perspective d'adaptation, les auteurs suggèrent de mettre en place des stratégies, non pas dans le but d'augmenter « l'adhésion stricte », mais pour contribuer à une adaptation de fidélité adéquate et équilibrée.

D'autres auteurs se sont inspirés par les travaux de Carroll et collègues pour suggérer des conseils théoriques et pratiques pour mieux définir et suivre la fidélité des interventions basées sur la technologie (Dabbs et al., 2011). Étant donné la croissance de l'utilisation de la technologie dans le domaine de la santé pour promouvoir des changements de comportement, Dabbs et collègues (2011) ont réalisé que ce cadre permet l'inclusion de composantes supplémentaires importantes pour évaluer la fidélité dans des interventions qui visent l'utilisation de la technologie. Les auteurs ont suggéré que des facteurs humains, qui expliquent l'acceptation de la technologie, et des facteurs prédictifs de l'intention d'utiliser et d'adopter une technologie sont des aspects nécessaires à considérer au moment de son évaluation. Donc, ces auteurs ont souligné l'importance de mesurer autant *l'acceptation de la technologie*, car les participants doivent accepter la technologie avant

d'envisager de l'utiliser (l'adopter), que *la perception d'utilité de la technologie* et sa facilité d'utilisation, car cette technologie va influencer leur propension à l'utiliser.

En résumé, tous ces travaux, qui visent à évaluer la fidélité de la mise en œuvre, ont en commun certains éléments qui doivent être considérés pour son évaluation tels que: (i) évaluer l'adhésion au programme; (ii) les facteurs qui affectent la livraison du programme; et (iii) la réactivité des participants. En plus, dans le cas d'un essai randomisé en grappes, il faut évaluer si des événements se sont déroulés dans le groupe contrôle, car ces événements peuvent influencer l'efficacité de l'intervention. Cependant, tous les auteurs suggèrent fortement plus de recherches empiriques pour tester et pour valider les différents guides suggérés. Le cadre de Carroll et collègues, modifié par Hasson, a été testé par d'autres auteurs en soulignant que c'est un outil d'évaluation utile pour évaluer la fidélité de la mise en œuvre des interventions complexes (R Allison et al., 2017; Augustsson, von Thiele Schwarz, Stenfors-Hayes, & Hasson, 2015; Gautier, Pirkle, Furgal, & Lucas, 2016; Hasson, Blomberg, & Dunér, 2012). Le cadre proposé par Carroll et collègues (Carroll et al., 2007), modifié par Hasson (Hasson, 2010), conceptualise la fidélité de la mise en œuvre comme étant multidimensionnelle, englobant à la fois l'intervention et son exécution. Ce cadre conceptuel offre un outil potentiel pour évaluer la mise en œuvre d'une intervention en développement ou pour une intervention déjà bien établie mais ce cadre ne fournisse aucune indication sur la manière d'explorer les relations entre ses éléments clés.

Bien que la littérature établisse des dimensions à considérer pour évaluer la fidélité de l'intervention, il n'y pas un plan d'opérationnalisation spécifique pour évaluer chaque concept. Quelques chercheurs ont suggéré que la fidélité de l'intervention peut être évaluée soit par des mesures génériques ou soit par des mesures spécifiques à l'intervention. Les mesures génériques de la fidélité évaluent les attributs généraux de l'intervention qui sont essentiels pour toute intervention (par exemple, le rapport travailleur de santé-famille), et les stratégies qui sont couramment utilisées pour des interventions similaires. (Breitenstein et al., 2010; Rezza, Law, Eva, Pollock, & Gorter, 2013). Bien que les mesures génériques de fidélité puissent être largement appliquées à des interventions théoriquement similaires, elles sont insuffisantes pour saisir les composantes uniques d'une intervention qui la distinguent des autres interventions. Les mesures de fidélité spécifiques à l'intervention évaluent les composantes clés d'une intervention particulière

pour permettent de reproduire les effets de l'intervention. Les mesures de fidélité spécifiques peuvent être surveillées et incluses dans l'analyse statistique (par exemple, le nombre de fois qu'un sujet a été contacté par téléphone pour évaluer la mise en œuvre d'une intervention téléphonique), et ainsi augmenter les connaissances sur la façon dont les composantes d'une intervention influencent le résultat.

Au vu de l'état actuel des connaissances, bien que l'évaluation de la fidélité d'une intervention ait été reconnue comme un élément essentiel de toute évaluation du processus, un élément important rarement rencontré dans la littérature est le fait d'évaluer toutes les composantes qui affectent la fidélité d'une intervention pendant sa phase de mise en œuvre. Actuellement, il y a peu de recherche empirique sur les facteurs qui influencent la fidélité, car la plupart des études antérieures sur la fidélité se sont axées uniquement sur la mesure de l'adhésion. Également, il existe actuellement un consensus limité sur la meilleure façon de définir et de mesurer la fidélité de la mise en œuvre. Le besoin d'études, qui donnent un sens au concept de fidélité et qui clarifient la fonction des facteurs modérateurs affectant la fidélité (la description détaillée de l'intervention, les stratégies pour faciliter la mise en œuvre, la qualité de la prestation, la réponse des participants, le recrutement et le contexte) et leur relation les uns aux autres, est fondamental.

Bref, le développement d'un outils de mesure de fidélité devrait être adapté à l'intervention à évaluer, et ces outils doivent être bien conçus afin qu'ils reflètent les éléments clés de l'intervention. Cependant, il faut tenir compte autant des facteurs que mesurent l'adhésion au programme ainsi que des facteurs modérateurs déjà soulevées auparavant pour donner un aperçu complet de la fidélité de la mise en œuvre. À la lumière de la littérature, aucun travail n'a fait l'objet d'une évaluation complète de la fidélité de la mise en œuvre d'une intervention.

1.6 Fidélité versus Adaptation

Dans le domaine de la santé, plusieurs études ont constaté que la fidélité avec laquelle une intervention est mise en œuvre affecte la façon dont elle performe (Dane & Schneider, 1998; Hansen, Graham, Wolkenstein, & Rohrbach, 1991; Hasson et al., 2012; Keith, Hopp, Subramanian,

Wiitala, & Lowery, 2010). En effet, des études ont démontré l'association entre la fidélité d'intervention et sa contribution aux résultats obtenus en affirmant que, plus le programme est fidèle, meilleurs sont les résultats obtenus (Dane & Schneider, 1998; Durlak & DuPre, 2008; Dusenbury et al., 2003). Par exemple, une revue systématique évaluant les données de plus de 500 études des programmes de promotion et de prévention a révélé que, lorsque les interventions sont bien mises en œuvre (avec une grande fidélité), l'ampleur de la taille moyenne des effets est au moins 2 à 3 fois plus élevée que lorsqu'il existe une faible fidélité d'intervention (Durlak & DuPre, 2008). Cependant, les interventions sont rarement mises en œuvre en alignement complet avec le modèle de programme en raison des adaptations. Les interventions en santé publique ne font pas exception, car elles sont complexes et dépendantes du contexte.

Donc, un débat permanent dans la littérature sur la fidélité est celui de la « fidélité stricte » par rapport à « l'adaptation », et la question non résolue de « combien » la fidélité est nécessaire ou bénéfique. Les adaptations sont les processus par lesquels le contenu ou la livraison d'une intervention est modifié pour mieux s'adapter aux conditions locales (von Thiele Schwarz, Hasson, & Lindfors, 2015). Selon Paquette et collègues (2009), les adaptations peuvent être identifiées afin qu'elles puissent être incorporées à une nouvelle intervention, et faire de l'intervention adaptée, un nouveau standard pour guider la fidélité (Paquette et al., 2009). Pour Orwin, les écarts par rapport au programme tel que prévu ne sont pas nécessairement négatifs, car si des événements externes incontrôlables s'immiscent dans le programme, des changements adaptatifs au plan initial peuvent être à la fois appropriés et nécessaires pour le déroulement de l'intervention (Orwin, 2000). Donc, un contrôle trop étroit de la forme de l'intervention risque de compromettre plutôt son efficacité.

Reconnaître l'équilibre entre le maintien de la fidélité et l'adaptation, en fonction du contexte et des toutes ces interactions entre les différentes parties prenantes, mérite une attention particulière dans les évaluations des programmes. Par exemple, si des activités sont programmées auprès des travailleurs de santé pour améliorer la qualité d'un service, et si une loi s'impose (lors de la phase de la mise en œuvre) pour fusionner des services à l'intérieur d'une institution, probablement que des adaptations pour s'ajuster au nouveau règlement devront être considérées en cours de route pour adapter l'intervention et la rendre plus efficace. Aussi, des adaptations pour mieux répondre aux besoins des participants ou pour mieux s'adapter au contexte, pour accroître

le sentiment d'appropriation parmi les intervenants ou pour améliorer les résultats du programme, sont des raisons invoquées pour adapter des interventions (Blakely et al., 1987; Hasson et al., 2012). Mesurer la fidélité en continu au cours d'une étude afin de documenter les adaptations et faire de corrections en cours de route, si nécessaire, sont des façons de capturer des modifications qui peuvent se présenter pendant la mise en place de l'intervention (Bond et al., 2000).

Donc, l'adaptation est un processus inévitable et parfois nécessaire qui peut augmenter la probabilité que l'innovation réponde aux besoins locaux des participants. Bien que plus la mise en œuvre d'une intervention fidèle à sa conception a l'avantage d'avoir de meilleurs résultats, il est difficile voire impossible que des interventions soit implémentées avec une fidélité à 100%. (Durlak & DuPre, 2008). Pour les essais pragmatiques, par exemple, une fidélité uniforme ou stricte à 100% peut ne pas être réalisable ni nécessairement souhaitable. Un consensus général dans la littérature est d'assigner une fidélité élevée aux interventions, quand la valeur de son évaluation mesure entre 80 à 100%, mais il n'y a pas suffisamment de preuves à l'appui (Perepletchikova & Kazdin, 2005). Donc, la conception d'évaluation de la fidélité doit être suffisamment flexible pour s'adapter, mais suffisamment explicite pour expliquer et pour documenter les variations qui se sont présentées pendant la mise en place de l'intervention.

L'utilisation de la théorie du changement, aussi appelée théorie du programme, model logique d'intervention (H.-t. Chen & Rossi, 1992), théorie de l'implantation (Weiss, 1997) ou théorie de l'action (Schorr, 1998), est un terme utilisé pour désigner le méthode ou le modèle conceptuel de la manière dont une intervention est censée produire son impact, est utile voire indispensable, pour identifier les variations qui se produisent lors de la mise en place de l'intervention (De Silva et al., 2014). La théorie du changement est un outil pour permettre aux évaluateurs de mesurer ce qui est livré dans la pratique, et de relever les adaptations et les changements qui minent la fidélité de l'intervention (Moore et al., 2015). Bien que la théorie du changement ait été utilisée pour concevoir, évaluer et fournir des explications causales dans les interventions de santé publique, une revue systématique, qui vise à évaluer son utilisation, n'a trouvé que quelques études qui l'utilisent pour identifier des adaptations qui puissent expliquer des variations dans la mise en œuvre de l'intervention (Breuer, Lee, De Silva, & Lund, 2015).

Une description détaillée des éléments essentiels du programme va favoriser une meilleure compréhension entre l'adhésion au modèle du programme original, et va aussi permettre d'identifier l'adaptabilité aux conditions locales. Donc, pour ne pas tomber dans la piège de la flexibilité, Carroll et collègues suggèrent que la mise en œuvre du programme puisse être flexible tant qu'il y a fidélité aux éléments dits « essentiels » d'une intervention (Carroll et al., 2007). L'absence de ces éléments aurait des effets négatifs importants sur la capacité d'une intervention à atteindre ses objectifs, et il serait donc difficile d'atteindre une fidélité de mise en œuvre élevée. Concilier la fidélité et les modifications pour adapter un programme sera possible dans la mesure qu'on comprenne d'abord la façon dont le programme réalise ses effets, et ensuite faire des possibles modifications au niveau de la mise en œuvre uniquement lorsque le mécanisme causal est préservé (Bauman, Stein, & Ireys, 1991).

Bien que les éléments essentiels d'une intervention devraient être standardisés pour ne pas les modifier, le mécanisme par lequel ceux-ci sont opérationnalisés pourrait être changé en permettant une flexibilité dans la mise en œuvre selon le contexte où l'intervention est testée (Kilbourne, Neumann, Pincus, Bauer, & Stall, 2007). Donc, un programme peut être modifié soit par des modifications du contenu du programme en omettant, modifiant ou ajoutant des composantes, soit par la manière ou l'intensité avec laquelle les composantes d'intervention sont livrées. (Castro, Barrera, & Martinez, 2004). Chaque composante essentielle de l'intervention peut avoir un point de coupure différent qui devra exiger des niveaux de fidélité plus élevés. Cela pourrait servir de guide pour de futures interventions, par exemple, en indiquant aux intervenants quels aspects d'un programme ils peuvent adapter et dans quelle mesure. Cela signifie que la fidélité et l'adaptation peuvent coexister sans que cette dernière soit perçue comme une menace pour la fidélité. La clé est de bien identifier les éléments du programme, d'être clair sur la quantité de changement ou d'adaptation permise, et d'enregistrer les variations dans la mise en œuvre, afin que la fidélité puisse être évaluée par rapport au degré de standardisation requis par le protocole de l'étude.

Finalement, afin de chercher un point d'équilibre entre la fidélité et l'adaptation des interventions, le cadre conceptuel de Carroll et collègues, soulevé dans la section précédente, a été modifié pour tenir compte des caractéristiques spécifiques à la fois de la fidélité et de l'adaptation,

au lieu d'évaluer seulement des sous-catégories exclusives à l'adhésion (D. Perez et al., 2016). Cependant, le cadre de Carroll et collègues (2007), modifié par Hasson (2012), a été utilisé dans d'autres études en démontrant son utilité pour évaluer de façon exhaustive la fidélité de la mise en œuvre, mais aussi pour comprendre les adaptations apportées pendant la mise en place d'une intervention (von Thiele Schwarz et al., 2015). Ainsi, ce cadre de référence est un outil pertinent pour guider l'évaluation de la fidélité de la mise en œuvre d'une intervention en santé publique, à la fois pour mesurer la fidélité et pour mieux comprendre le niveau atteint, mais aussi pour relever les adaptations qui peuvent se présenter en cours de route.

1.7 Évaluer la fidélité de la mise en œuvre de l'intervention Tika Vaani dans un contexte d'une étude pilote d'essais randomisés en grappes : étude de cas

Cette section vise à présenter l'importance d'évaluer la fidélité de la mise en œuvre des interventions qui sont dans le contexte d'une étude pilote, en commençant par clarifier le but de ce type d'étude et l'importance de concevoir ce type d'évaluation, pour continuer par la suite à mettre en contexte le développement et l'intervention même, qui se voit bonifiée par l'évaluation.

1.7.1 Le but de la fidélité dans l'étude pilote

Dans le domaine de la recherche, le terme étude pilote semble, pour certains auteurs, un synonyme d'étude de faisabilité, mais, pour d'autres, ces études ont différentes approches. De façon générale, une étude pilote est définie comme la réplique d'une petite étude afin de tester la conception ou la méthodologie de l'étude proposée (Craig et al., 2008; Richards & Hallberg, 2015; Thabane et al., 2010). En ce qui concerne une étude de faisabilité, elle rassemble des études régulièrement centrées pour répondre aux questions sur les aspects au niveau du processus et de la conception de l'étude, afin de vérifier que le projet soit techniquement faisable et économiquement viable (Richards & Hallberg, 2015). Une récente enquête Delphi a été utilisée pour mieux définir la faisabilité et les études pilotes utilisées pour la préparation des essais contrôlés randomisés. À

ce propos, les auteurs ont développé un cadre conceptuel et ont défini les études pilotes comme un sous-ensemble d'études de faisabilité (S. M. Eldridge et al., 2016). Une étude de faisabilité demande si quelque chose peut être fait, si nous procédons avec elle, et si oui, comment. Une étude pilote pose les mêmes questions, mais possède une caractéristique de conception spécifique pour mener à petite échelle la totalité ou une partie d'une étude future (S. M. Eldridge et al., 2016). Donc, dans le cadre de cette thèse, nous considérons une étude pilote comme une étude de plus petite taille qui vise à étudier la faisabilité des éléments cruciaux de l'étude principale, et à guider la planification d'une intervention à grande échelle (Thabane et al., 2010).

Plusieurs raisons sont attribuées à l'utilisation des études pilotes. En effet, elles peuvent servir à: (i) identifier ou affiner une question de recherche; (ii) vérifier si la conception est adéquate; (iii) déterminer les critères d'admissibilité et les méthodes de recrutement; (iv) estimer combien de temps et de ressources seront nécessaires pour compléter la version plus grande; et (v) déterminer la meilleure façon de mener un projet de recherche à grande échelle (Arain, Campbell, Cooper, & Lancaster, 2010; Shanyinde, Pickering, & Weatherall, 2011; Thabane et al., 2010). Pour d'autres auteurs, étant donné le rôle que joue ce type d'étude dans la recherche en santé publique, les études pilotes vont fournir un cadre de référence pour encourager les chercheurs à assurer la rigueur méthodologique de ce type d'étude, afin qu'elle soit scientifiquement valide et publiable (Lancaster, Dodd, & Williamson, 2004). Une étude pilote peut être un outil clé également pour examiner le respect du protocole par les intervenants et les participants, ainsi que pour examiner les procédures de collecte et la qualité des données lors de l'étude (Brooks & Stratford, 2009). Les études pilotes, dans son rôle de planification dans la recherche à grande échelle, fournissent des informations très importantes pour optimiser les ressources humaines, techniques, et économiques avant de les implémenter à grande échelle.

Les études pilotes sont vues comme un outil de la science de la mise en œuvre. La science de la mise en œuvre peut être vue comme un domaine de méthodes visant à promouvoir l'adoption systématique des résultats de la recherche dans la pratique (Eccles & Mittman, 2006) Donc, elle comble le fossé entre la théorie et la pratique en étudiant la façon dont les programmes fondés sur des preuves peuvent être intégrés pour maximiser les résultats positifs (Fixsen & Blase, 2009). La recherche sur la mise en œuvre prend en compte autant le processus, que les facteurs affectant la

mise en œuvre, ainsi que les résultats de la mise en œuvre pour les améliorer et promouvoir leur utilisation à grande échelle et leur durabilité (Peters, Adam, Alonge, Agyepong, & Tran, 2013). En étudiant le succès et l'échec de l'adoption des interventions, dans diverses disciplines, cette approche scientifique permet de mieux comprendre comment les stratégies accréditées peuvent être transférées avec succès dans de nouveaux contextes (Fixsen et al., 2005).

En raison de la nécessité perçue de passer à la mise en œuvre des interventions efficaces à grande échelle, les études pilotes sont un outil de la science de la mise en œuvre pour soutenir ce processus, et pour augmenter la probabilité de sa durabilité dans le temps. Cependant, pour savoir si ces études sont mises en œuvre telles que conçues, une évaluation qui vise à fournir ces informations devient essentielle avant d'attribuer sa faisabilité à grande échelle. Donc, sous la perspective « évaluative », l'évaluation d'une intervention devrait être menée tout au long des différentes étapes d'une intervention, en commençant par la conception du projet et en terminant par une évaluation des résultats finaux. L'évaluation formative pendant la phase de développement du projet, tel que dans une étude pilote, est cruciale pour clarifier les objectifs et pour affiner la conception du projet, y compris la stratégie d'évaluation et les exigences en matière de données, tout en tenant compte des facteurs locaux qui influencent la façon dont le projet est mis en œuvre.

1.7.2 Contexte de l'intervention Tika Vaani

En 2018, l'Organisation mondiale de la santé (OMS) a constaté qu'environ 19,4 millions d'enfants dans le monde ne sont pas encore complètement vaccinés. Environ 60% (11,7 millions) de ces enfants vivent dans des pays à revenu faible ou intermédiaire, et 29% (2,6 millions) en Inde (WHO, 2018). L'Inde a fait des progrès substantiels pour sauver des vies et pour améliorer la qualité de vie de millions de personnes dans le pays et dans le monde. Des enquêtes démographiques sur la santé, connue sous le nom d'Enquête nationale sur la santé et la famille (NFHS), lancée en Inde en 1992-1993 et jusqu'en 2015-2016, fournissent des informations sur plusieurs aspects tels que la fertilité, la mortalité, la planification familiale et les aspects importants de la nutrition, de la santé et des soins de santé dans la population. Des analyses à partir de ces enquêtes ont été faites pour déterminer le taux de progression annuel et le taux de progression

requis pour comprendre le progrès entre 2005-06 et 2015-16. (Panda & Mohanty, 2019). L'analyse de ces enquêtes laisse voir des progrès importants vers les Objectifs de Développement Durable. Par exemple, dans le domaine infantile, environ 43,5% des enfants âgés de 12 à 23 mois ont été entièrement vaccinés en 2005-2006, ce qui est passé à 62% en 2014-2015, avec un taux de progression annuel de 3,27%. Cependant, le taux de progression annuel requis pour réaliser la vaccination universelle est de 4,1% (Panda & Mohanty, 2019).

Bien que l'Inde ait considérablement amélioré la couverture vaccinale des enfants, il reste d'importants défis à surmonter pour combler le déficit de vaccination, en particulier du côté de la demande. Notamment en Uttar Pradesh, où malgré les efforts, la couverture vaccinale complète reste faible. Selon la plus récente série d'Enquête nationale sur la santé et la famille, dans le district d'Uttar Pradesh, le pourcentage des enfants âgés de 12 à 23 mois entièrement immunisés est de seulement 51%, et ce pourcentage tombe à 39% dans la ville d'Hardoi (IIPS, 2016), la zone ciblée de l'étude. Uttar Pradesh est l'État le plus peuplé de l'Inde et a enregistré le plus grand nombre de décès d'enfants de moins de 5 ans, contribuant à 27,1% (325 500) du total national (1 200 998), principalement en raison de causes qui pourraient être potentiellement traitées par les vaccins (Liu et al., 2019).

Malgré plusieurs programmes et les efforts du gouvernement pour réduire la mortalité infantile et juvénile, Uttar Pradesh est l'un des districts présentant le plus grand écart de couverture vaccinale parmi les sous-groupes de population défavorisés (les parents les plus pauvres et les moins instruits, la caste sociale inférieure et les croyances religieuses). Une récente étude a révélé qu'en Inde, parmi les castes et tribus répertoriées (représentant les groupes ethniques les plus opprimés et les plus vulnérables de la société indienne), seulement 55,8% et 63,2% de leurs enfants, respectivement, étaient entièrement immunisés, tandis qu'en Uttar Pradesh, ces chiffres étaient respectivement de 34,1% et 52,1% (Vishwakarma, Shekhar, Dutta, & Yadav, 2019). D'autres études ont constaté que des facteurs sociodémographiques, des niveaux de connaissances et des croyances publiques peuvent affecter la couverture vaccinale des enfants (Gurnani et al., 2018; Shrivastwa, Wagner, & Boulton, 2019). Donc, pour atteindre une couverture élevée et équitable, des actions ciblées sont nécessaires pour garantir l'accès à la vaccination, en particulier pour les

populations vulnérables. Des approches éducatives ciblant la population générale et adaptées aux croyances locales sont nécessaires pour surmonter ces obstacles (Gurnani et al., 2018).

Le gouvernement de l'Inde veut s'assurer que la vaccination touche tous les enfants du pays et ainsi contribuer à la réalisation des Objectifs de Développement Durable qui vise, entre autres, à réduire le taux de mortalité infantile. La vaccination figure parmi les interventions sanitaires les plus efficaces et les plus rentables pour combattre les maladies infectieuses, et réduire le taux de mortalité chez les enfants (OMS, 2012). La vaccination des enfants a grandement contribué aux progrès accomplis dans la réalisation du quatrième Objectif du Millénaire, à la réduction de la mortalité infantile, et cette vaccination est au cœur de l'objectif de santé dans les Objectifs de Développement Durable post-2015. Pour continuer le progrès contre la mortalité infantile, des stratégies concentrées davantage sur les interventions qui peuvent protéger les plus jeunes sont requises (Roser et al., 2013). Plusieurs revues systématiques ont mis en lumière le fait que les stratégies du côté de la demande, telles que l'éducation et le transfert de connaissances, ainsi que l'utilisation des technologies sont des outils clés pour promouvoir la prise du vaccin, et accroître le taux de vaccination infantile dans le pays du Sud (Batt, Fox-Rushby, & Castillo-Riquelme, 2004; Johri et al., 2015; Oyo-Ita et al., 2016; Shea, Andersson, & Henry, 2009; Sondaal et al., 2016).

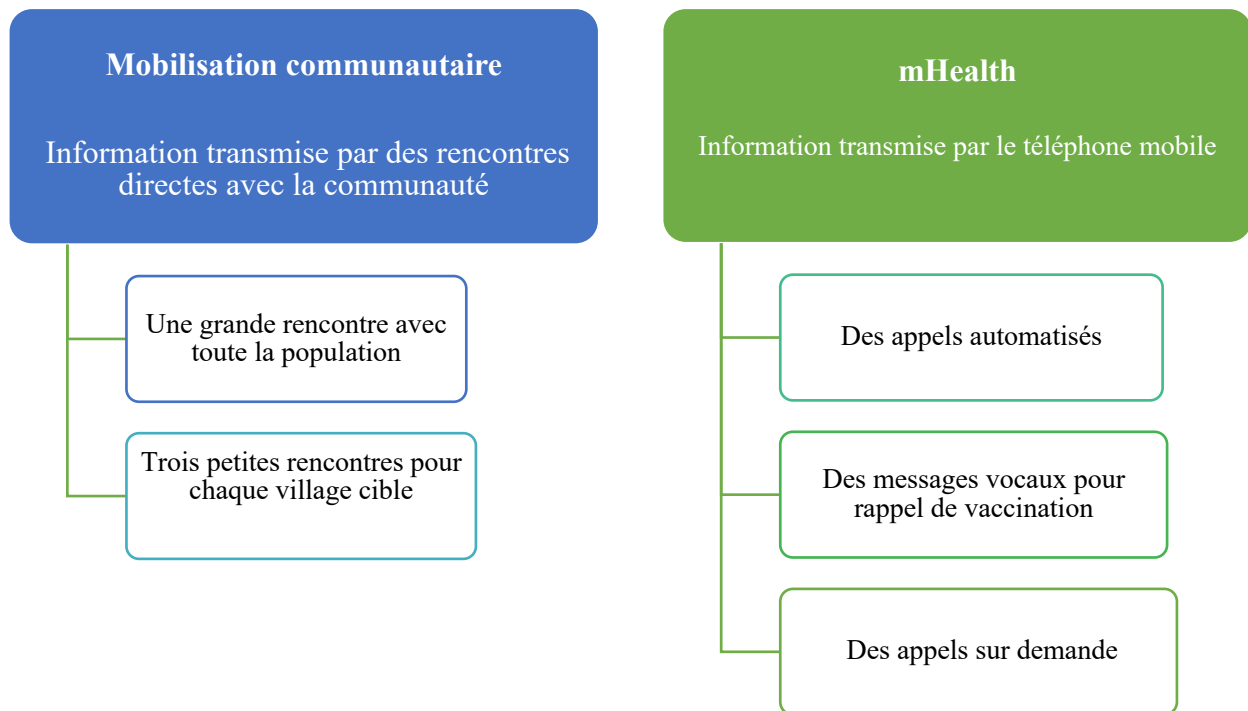
1.7.3 L'Intervention Tika Vaani

L'intervention Tika Vaani (« voix de la vaccination » en Hindi), en utilisant une plateforme mobile interactive combinée à une mobilisation communautaire, a été conçue pour éduquer les bénéficiaires sur des thèmes de base de la santé infantile et sur la vaccination des enfants. L'objectif principal de cette intervention est d'augmenter la couverture vaccinale des enfants de 0 à 2 ans vivant dans les zones rurales d'Uttar Pradesh, en Inde. En outre, dans le but de renforcer les soins de santé primaires, l'intervention vise à améliorer les connaissances en santé de tous les membres de la communauté, des agents de santé communautaires et particulièrement des familles avec de jeunes enfants.

Le développement de cette intervention est présenté en détail dans un des articles dans la section des résultats. De façon globale, l'intervention a été développée à partir des résultats d'une recherche formative pour éclairer la conception de l'intervention et pour la rendre plus adaptée au contexte local et aux populations défavorisées. Le développement du contenu a été conçu par Gram Vaani, une entreprise Indienne spécialisée dans les plateformes de médias pour les populations rurales peu alphabétisées. Cependant, le développement du contenu a suivi un processus interactif avec des experts scientifiques au niveau national et international, et par des membres du gouvernement local en santé. Le contenu consiste en capsules d'information liée à l'importance de la vaccination infantile, la prévention, la reconnaissance et la gestion des maladies infectieuses telles que pneumonie, diarrhée, dengue, chikungunya, ainsi que sur des sujets concernant la nutrition et l'eau, l'assainissement et l'hygiène.

Le contenu développé par l'intervention Tika Vaani a été transmis par deux stratégies: (1) stratégie de mobilisation communautaire et (2) stratégie mHealth (Figure 1).

Figure 1. – Composantes et activités de l'intervention Tika Vaani



La mobilisation communautaire consiste dans des rencontres en personne avec la communauté. Dans un premier temps, les intervenants-terrain ont mené une grande réunion avec tout le village pour présenter le projet aux communautés. Par la suite, trois petites réunions offertes à chaque participant couvrant des thèmes spécifiques en matière de santé ont été offertes par les intervenants-terrain, selon un calendrier établi au départ par les chercheurs. Ces rencontres permettent à la fois de vérifier la compréhension du contenu abordé selon la programmation, ainsi que d'évaluer l'expérience de l'utilisation de la plateforme mobile. En ce qui concerne le volet *mHealth*, des capsules audio éducatives en format ludo-éducatif adaptés au contexte ont été transmises par téléphone mobile, comme des appels sortants automatisés deux fois par semaine, ainsi que des rappels vocaux pour les séances de vaccination basés sur la date de naissance aux familles du groupe cible. De même, tous les participants des communautés ciblées pouvaient faire un appel téléphonique pour accéder à la plate-forme, et, par conséquent, recevoir un rappel leur permettant d'accéder à tous les contenus, d'enregistrer leurs requêtes / commentaires ou de se connecter à un expert en direct.

1.8 Synthèse du cadre théorique de la thèse au sein de la fidélité de la mise en œuvre

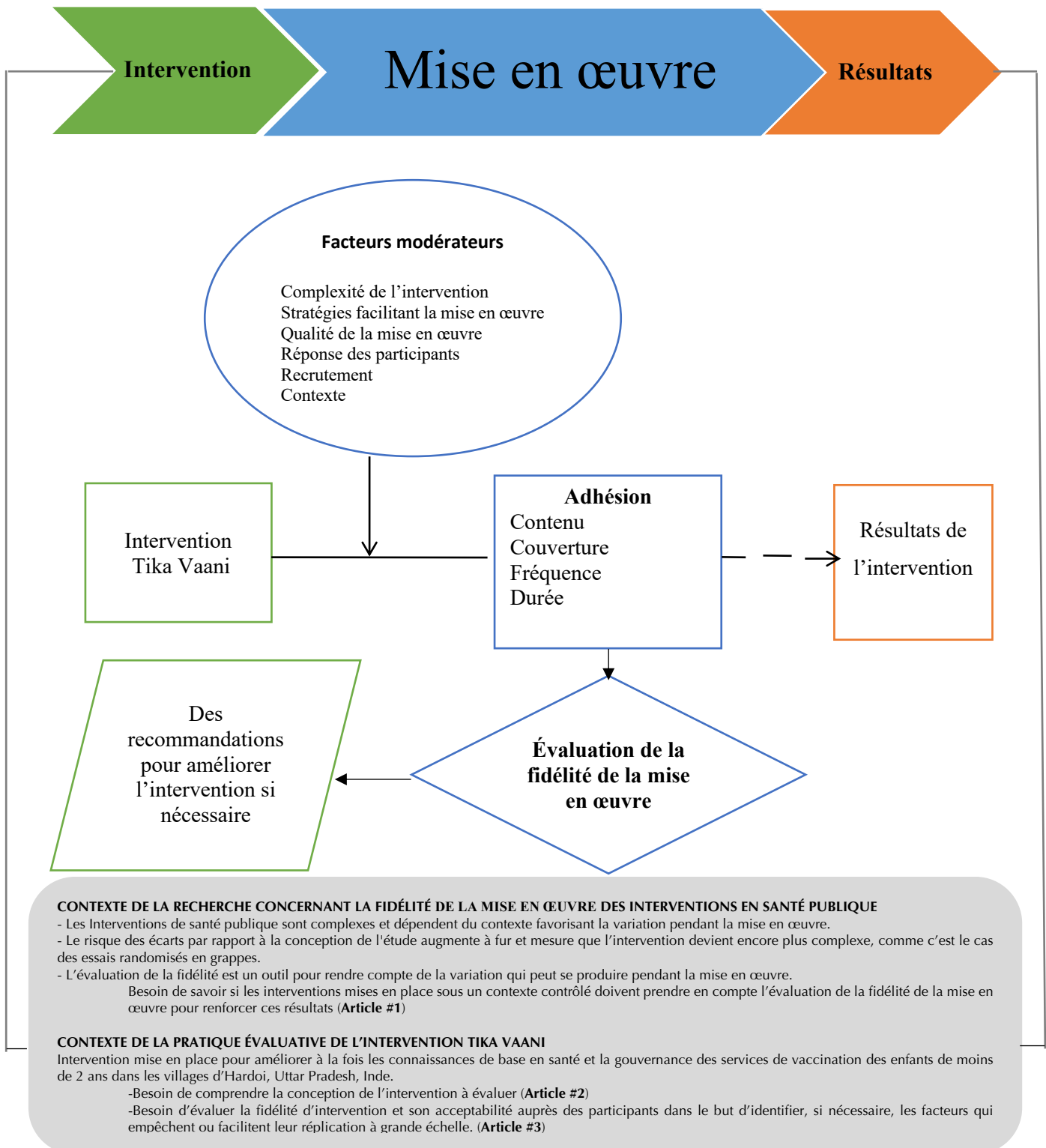
L'objectif principal de la réalisation d'un essai randomisé en grappes est de vérifier l'efficacité d'une intervention en la comparant à un groupe contrôle. Bien que l'essai contrôlé randomisé représente une méthode fiable pour établir la causalité, les interventions complexes, comme dans le cas des essais randomisés en grappes, peuvent être à risque d'éprouver des écarts du protocole et la non-adhésion au traitement, compromettant ainsi la validité des conclusions. Étant donné que les interventions de santé publique sont « complexes » et « dépendantes du contexte », des écarts du protocole, par rapport à ce qui a été planifié, peuvent se présenter lors de la mise en œuvre. Dans un contexte d'un essai randomisé en grappes, cette évaluation est aussi importante, car des différences contextuelles au niveau de chaque grappe pourraient favoriser l'adaptation locale de l'intervention. Ces variations peuvent être encore plus importantes dans les pays du Sud où les ressources sont souvent limitées. Une évaluation qui permet de tenir compte du

niveau de fidélité de la mise en œuvre et des adaptations mises en place, et qui fait ressortir les facteurs qui produisent ce niveau de fidélité, est fondamental voire indispensable dans les interventions de santé publique.

Les programmes innovants ont besoin d'un temps de planification important avant le démarrage du programme, car ils mettent du temps à se développer et à mûrir (Orwin, 2000). Les études pilotes sont un outil précieux pour affiner l'intervention. Donc, l'évaluation de la fidélité d'intervention de l'étude pilote, sous une perspective formative, favorise une rétroaction à la direction du programme sur le niveau atteint de fidélité de la mise en œuvre, ainsi que des recommandations pour l'améliorer. L'évaluation de la fidélité devient un atout dans ce type d'études pour identifier les facteurs qui limitent ou facilitent leur reproduction. Cette évaluation est particulièrement nécessaire dans les pays à faible revenu, qui doivent optimiser les ressources pour bénéficier au plus grand nombre de personnes possible, et pour améliorer la santé de la population.

Dans le but de réaliser un futur essai randomisé en grappes afin d'évaluer l'impact de l'intervention Tika Vaani sur la couverture vaccinale, les concepteurs de l'intervention ont décidé de mener un essai pilote et veulent évaluer la fidélité de l'intervention visée. Bien que l'évaluation de la fidélité et de l'adaptation sont deux aspects importants à évaluer dans la mise en œuvre des interventions de santé mondiale, car ils ont une incidence sur l'efficacité d'une intervention, l'intervention Tika Vaani est une nouvelle intervention qui est testée dans un premier temps pour évaluer sa faisabilité, mais aussi pour évaluer si ses composantes essentielles devront être mise en œuvre afin que l'intervention produise ses effets. Donc, évaluer l'adaptation que sont les modifications du modèle d'origine, qui se produisent pendant la phase de la mise en œuvre, ne sont pas une évaluation pertinente pour une intervention qui doit valider d'abord son modèle initial, avant d'évaluer son efficacité et son passage à grande échelle. Enfin, cette évaluation répond à la volonté des concepteurs pour améliorer, si nécessaire, l'intervention dans sa phase pilote avant un passage à grande échelle.

Figure 2. – Synthèse du cadre théorique de la thèse au sein de la fidélité de la mise en œuvre*



*Cadre théorique de la thèse adapté du modèle conceptuel pour l'évaluation de la fidélité de la mise en œuvre proposé par Carroll et collègues 2007, modifié par Hasson (2010)

Chapitre 2 – Démarche scientifique, objectifs, méthode de recherche, rôle de l'étudiante et considérations éthiques

L'état des connaissances présentées dans cette thèse a mis en relief la variation que les interventions en santé publique peuvent présenter dans leur mise en œuvre, même pour celles qui sont testées, sous le modèle proclamé « standard de haute qualité », pour évaluer l'efficacité des programmes comme dans le cas des essais randomisés en grappes. Des lacunes dans la mise en œuvre dans ce type d'étude peuvent compromettre la validité de résultats de l'étude, ce qui soulève l'importance d'évaluer leur processus pour atténuer ces problèmes méthodologiques. Cette variation peut être plus importante dans le contexte de ressources limitées, étant donné que ce type d'intervention dépend du contexte où elle intervient. Un outil pour évaluer ce type d'intervention et pour renforcer la validité des résultats est l'analyse de la fidélité de la mise en œuvre, qui vise à évaluer si l'intervention a été mise en place selon les concepteurs du programme. Le survol des théories concernant les guides pour évaluer la fidélité de la mise en œuvre de l'intervention nous confirme que cette évaluation est un outil indispensable, entre autres, pour évaluer les interventions avant de les mettre en échelle dans d'autres contextes. De plus, un consensus sur la théorie de l'évaluation de la fidélité est de mesurer le niveau d'adhésion mais, aussi, d'analyser tous les différents facteurs (la description détaillée de l'intervention, les stratégies pour faciliter la mise en œuvre, la qualité de la prestation, la réponse des participants, le recrutement et le contexte) qui peuvent l'influencer pour mieux comprendre les résultats obtenus.

Sous cette perspective, cette thèse s'inscrit dans le champ de la recherche évaluative sous l'angle d'analyse de la mise en œuvre lequel vise à examiner les relations entre l'intervention, ses composantes et le contexte dans lequel elle intervient (Contandriopoulos, Champagne, Denis, & Pineault, 1993). Cette thèse vise donc à offrir deux types de contributions : premièrement, la contribution méthodologique dans le domaine des sciences d'implémentation concernant l'évaluation de la fidélité des interventions en santé publique, sous le modèle des essais randomisés en grappes, et deuxièmement, une contribution empirique à l'intervention évaluée avant sa mise en échelle. Ainsi, cette thèse comporte deux volets qui correspondent aux deux objectifs principaux de recherche: une revue systématique et une étude de cas unique à trois unités d'analyse selon une

approche mixte concomitante. Cette section sera présentée en fonction de chaque volet de recherche.

2.1 Premier volet : Revue systématique

2.1.1 Démarche scientifique

Une revue systématique est une approche méthodologique clé de haute qualité, rigoureuse et reproductible, dans l'approche fondée sur des données probantes (Stevens, 2001). Cette approche permet d'identifier, d'analyser et de synthétiser les revues scientifiques primaires d'un sujet en particulier pour répondre à une question précise de recherche (Clarke, 2011). Les résultats de cette approche ont différentes utilités, et sont souvent requis par des chercheurs et des bailleurs de fonds pour établir l'état des connaissances existantes, pour estimer l'effet d'une intervention, pour mettre en évidence des lacunes existantes dans les données scientifiques. Ces résultats sont fréquemment utilisés dans l'élaboration des lignes directrices, pour l'élaboration des politiques dans différents domaines pour orienter les recherches futures. La réalisation d'une revue systématique doit suivre une démarche scientifique semblable à toute étude de recherche. Elle doit être rédigée dans un protocole enregistré ou publié pour limiter les biais de report, et assurer la transparence de l'étude. De plus, elle fait une évaluation critique des études de qualité et la synthèse des résultats, selon une méthode prédéterminée (Clarke, 2011; Moher D, Liberati, Tetzlaff, Altman, & Group, 2009; Stevens, 2001).

Afin de guider le développement des protocoles de revues systématiques et de méta-analyses des études à évaluer, les auteurs sont encouragés à utiliser la guide PRISMA-P pour décrire la justification et l'objectif prévu de la revue systématique, ainsi que les approches méthodologique et analytique qui seront abordées (Moher D et al., 2009). En effet, une revue systématique doit suivre une série d'étapes qui seront définies lors de l'élaboration du protocole, afin que cette revue systématique soit reproductible (Higgins & Wells, 2011). Finalement, une revue systématique peut, ou non, inclure une méta-analyse, c'est-à-dire une analyse statistique des résultats des études indépendantes, qui vise à produire une synthèse de connaissances sur un sujet donné.

La base de données PROSPERO, une base internationale de données des revues systématiques enregistrées en matière de santé et de services sociaux a été créée pour faciliter le registre des protocoles, afin d'assurer que les revues systématiques soient faites de façon transparente, reproductible et conforme aux normes de qualité reconnues à l'échelle internationale (Booth et al., 2012). Étant donné que notre revue systématique met l'accent sur des questions méthodologiques plutôt que sur des résultats liés à la santé, elle ne peut pas s'enregistrer dans la base internationale de données PROSPERO. Cependant, nous avons élaboré et publié le protocole pour cette revue systématique (M. C. Perez, Minoyan, Ridde, Sylvestre, & Johri, 2016), et ainsi éviter que ces travaux de synthèse soient influencés par les données trouvées dans les études incluses dans la revue (Higgins & Wells, 2011; Reviews & Dissemination, 2009). Cet article (Article 1 A) est présenté dans la section de méthode de recherche pour ce premier volet.

2.1.2 Objectif de recherche

La revue systématique permettra d'étudier comment l'évaluation de la fidélité de la mise en œuvre des interventions en santé publique, utilisant des essais randomisés en grappes (ERG) dans les pays du Sud, a été rapportée. Par ailleurs, cette revue permettra d'identifier s'il y a des lacunes, afin d'offrir des recommandations, si nécessaire, pour améliorer le rapport de ces études. Des recommandations pour améliorer la pratique de l'évaluation de la fidélité de la mise en œuvre des interventions pour faire avancer le domaine des sciences de la mise en œuvre sont présentées dans le Chapitre Discussion.

2.1.3 Méthode de recherche

Nous avons suivi les lignes directrices PRISMA-P 2015 pour l'élaboration et le rapport des revues systématiques, ainsi que le guide Cochrane pour évaluer le risque de biais des essais randomisés en grappes des interventions en santé publique dans les pays du Sud, qui ont planifié et/ou évalué la fidélité de la mise en œuvre de l'intervention dans le protocole et/ou dans l'article principal. Nous avons développé la démarche méthodologique pour cet objectif, et elle est

présentée en détail dans cette section sous forme d'article scientifique (Article 1A) afin de présenter le processus mis en place pour répondre à l'objectif de recherche de ce premier volet de la thèse.

De façon synthétique, nous présentons ici la démarche méthodologique pour la revue systématique en sept étapes :

- i. **Caractéristiques des études incluses :** Nous avons inclus les essais randomisés en grappes des interventions de santé publique dans les pays du Sud qui ont planifié l'évaluation et/ou évalué la fidélité de la mise en œuvre de l'intervention dans le protocole et/ou dans l'article principal (ou dans un document associé). Les études admissibles sont des études faites dans les pays à faible revenu et à revenu intermédiaire (LMIC), selon la classification de la Banque Mondiale (The World & Bank.) et publiées en anglais, en espagnol ou en français à partir du 1er janvier 2012 (dernière mise à jour des guides CONSORT pour améliorer le rapport des essais randomisés en grappes). Nous avons exclu les études qui n'ont pas les caractéristiques mentionnées ci-haut ainsi que des études pilotes, des études de la littérature grise et des analyses secondaires d'une étude principale pour laquelle les résultats pertinents ont été publiés avant 2012.
- ii. **Sources d'information et stratégie de recherche :** La stratégie de recherche a été développée en utilisant les MESH (Medical Subject Headings) et des mots liés aux essais randomisés en grappes, aux pays du Sud et aux interventions en santé publique. Nous avons développé la stratégie de recherche pour MEDLINE. Cette stratégie a été validée par une bibliothécaire et ensuite adaptée aux bases de données EMBASE, CINAHL, PubMed et EMB. Les études identifiées ont été téléchargées dans le logiciel EndNote et les doublons ont été éliminés.
- iii. **Sélection des études :** les études ciblées ont été sélectionnées en fonction des critères d'inclusion et des critères d'exclusion. Cette phase a été faite en deux temps par deux chercheurs avec une lecture indépendante. Nous avons inclus les essais randomisés en grappes qui ont le numéro d'enregistrement dans le résumé afin d'assurer la disponibilité du protocole. Les études sélectionnées ont été documentées dans un organigramme, et un tableau a été fourni pour décrire les études qui ont été exclues et les raisons de leur exclusion.

- iv. Extraction de données des études : Pour chaque étude incluse, nous avons comparé l'évaluation de la fidélité planifiée versus rapportée pour analyser leur concordance. Nous avons élaboré une grille pour enregistrer les informations, et nous avons évalué quelles sont les dimensions de la fidélité analysée et les méthodes de collecte de données utilisées pour évaluer chaque dimension (Annexe 1). Le cadre conceptuel de Carroll et coll (Carroll et al., 2007) modifié par Hasson (Hasson, 2010) a été utilisé.
- v. Priorisation des résultats : Le résultat principal de cette étude est la proportion de l'accord global entre le protocole et le rapport d'essai concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention. Les résultats secondaires sont détaillés dans le protocole publié.
- vi. Évaluation de la qualité des études incluses : Le guide Cochrane et d'autres sources clés permettant d'évaluer le risque de biais dans les essais randomisés en grappes ont été utilisés.
- vii. Synthèse des données : Les données ont été regroupées à l'aide de techniques quantitatives (pourcentage d'accord entre le protocole et l'étude publiée) et narratives en utilisant comme référence les dimensions de la fidélité du cadre conceptuel utilisé.

Pour ce volet de la thèse, nous avons trouvé 90 articles qui sont des interventions de santé publique évaluées par des essais randomisés en grappes dans les pays du Sud. Parmi ces études, 65 études ont rapporté des caractéristiques concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention soit dans le protocole de l'étude, dans l'article ou dans les deux. Étant donné que plusieurs revues ont des contraintes d'espace et que les auteurs de l'étude peuvent omettre des informations clés de l'évaluation de la fidélité dans l'article principal, nous avons accepté les essais randomisés en grappes rapportant l'évaluation de la fidélité de la mise en œuvre de l'intervention dans un document complémentaire. Cela nous a permis d'avoir une image complète du rapport actuel concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention. Notre révision a consisté à examiner 185 articles distribués dans 21 protocoles d'articles évalués par des pairs, 15 protocoles de littérature grise, 65 registres des essais, 65 articles évalués par des pairs, 6 rapports de littérature grise et 13 documents complémentaires. De façon synthétique, les caractéristiques

des études incluses (n=65) dans la revue systématique ont été 75% (n=49/65) des études en utilisant le modèle classique d'essai randomisé en groupe parallèle; 55% (n = 36/65) des études ont été menées en Afrique subsaharienne, 37% (n = 24/65) ont visé des interventions ciblant le VIH, la tuberculose, le paludisme et la dengue, et dans 72% (n = 47/65) des études ont choisi les services ou soins habituels comme groupe comparateur. Les résultats plus en détail de ce volet sont présentés dans un des articles (Article 1B) qui composent le Chapitre Résultats de cette thèse. Finalement, nous avons créé un outil d'évaluation de risque de biais, en nous guidant sur la base de l'outil de collaboration Cochrane et d'autres sources clés pertinentes pour évaluer le risque de biais dans les essais randomisés en grappes. En effet, l'outil a dû être créé, car le guide Cochrane est développé pour évaluer le biais dans les études randomisées individuellement, mais ne permet pas une évaluation complète de biais dans ce type d'étude.

2.1.3.1 Article 1a Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low-and middle-income countries: systematic review PROTOCOL

Myriam Cielo Pérez^{1,2†}; Nanor Minoyan^{1,2}; Valéry Ridde^{2,3}; Marie-Pierre Sylvestre^{1,2}; Mira Johri^{1,4†}

† These authors contributed equally.

STATUT: ARTICLE PUBLIÉ

Pérez, M. C., Minoyan, N., Ridde, V., Sylvestre, M. P., & Johri, M. (2016). Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low-and middle-income countries: systematic review protocol. *Systematic reviews*, 5(1), 177.

Author affiliations

1 Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal, Québec, Canada

2 Département de médecine sociale et préventive, École de santé publique (ESPUM), Université de Montréal, Montréal, Québec, Canada

3 Institut de Recherche en Santé Publique Université de Montréal (IRSPUM), Pavillon 7101 Avenue du Parc, P.O. Box 6128, Centre-ville Station, Montréal, Québec, H3C 3J7, Canada

4 Département de gestion, d'évaluation, et de politique de santé, École de santé publique, Université de Montréal, Montréal, Québec, Canada

Contribution des auteurs

À partir d'une idée originale de Dre Mira Johri, Myriam Cielo Pérez et Dre Mira Johri ont développé la conceptualisation du protocole ainsi que la rédaction de l'article. Dre Marie-Pierre Sylvestre a donnée des conseils statistiques, et avec Nanor Minoyan et Dr Valéry Ridde, ont apporté un regard critique et constructif à l'ensemble de l'article.

Tous les auteurs ont lu et approuvé le manuscrit final pour soumission.

Abstract

Background: Cluster-randomised trials (CRTs) are a key instrument to evaluate public health interventions, particularly in low- and middle-income countries (LMICs). Fidelity assessment examines study processes to gauge whether an intervention was delivered as initially planned. Evaluation of implementation fidelity (IF) is required to establish whether the measured effects of a trial are due to the intervention itself and may be particularly important for CRTs of complex interventions. Current CRT reporting guidelines offer no guidance on IF assessment. We will systematically review the scientific literature to study current practices concerning assessment of IF in CRTs of public health interventions in LMICs.

Methods: We will include CRTs of public health interventions in LMICs that planned or assessed IF in either the trial protocol or the main trial report (or an associated document). Search strategies use Medical Subject Headings (MESH) and text words related to CRTs, developing countries, and public health interventions. The electronic database search was developed first for MEDLINE and adapted for the following databases: EMBASE, CINAHL, PUBMED and EMB Reviews, to identify CRT reports in English, Spanish or French published on or after January 1, 2012. To ensure availability of a study protocol, we will include CRTs reporting a registration number in the abstract. For each included study, we will compare planned versus reported assessment of IF, and consider the dimensions of IF studied, and data collection methods used to evaluate each dimension. Data will be synthesised using quantitative and narrative techniques. Risk of bias for individual studies will be assessed using the Cochrane collaboration risk of bias tool criteria and additional criteria related to CRT methods. We will investigate possible sources of heterogeneity by performing subgroup analysis. This review was not eligible for inclusion in the PROSPERO registry

Discussion: Fidelity assessment may be a key tool for making studies more reliable, internally valid and externally generalizable. This review will provide a portrait of current practices related to assessment of intervention fidelity in CRTs and offer suggestions for improvement. Results will be relevant to researchers, those who finance health interventions, and for decision makers who seek the best evidence on public health interventions.

Keywords: cluster randomised trials; implementation fidelity; public health interventions; process evaluation; developing countries; systematic review protocol.

INTRODUCTION

As evidenced by their growing presence in the scientific literature (Bland, 2004; M. K. Campbell, Elbourne, Altman, & group, 2004), cluster-randomised trials (CRTs) have become a key instrument to evaluate public health interventions (Bland, 2004; M.K. Campbell, Mollison, Steen, Grimshaw, & Eccles, 2000; Isaakidis & Ioannidis, 2003; Moberg & Kramer, 2015; Osrin et al., 2009), particularly in low- and middle-income countries (LMICs) (Handlos, Chakraborty, & Sen, 2009; Isaakidis & Ioannidis, 2003). Randomized controlled trials (RCTs) are widely considered to provide the highest quality of evidence on the effectiveness of health interventions (M. Campbell et al., 2000; Craig et al., 2012; Evidence-Based, Medicine Working, & Group., 1992; Richards & Hallberg, 2015), and CRTs are a form of randomised trial in which clusters of individuals (such as families, villages, hospital services, or schools) rather than independent individuals are randomly allocated to intervention or control groups (2). Increasingly, public health researchers recognize the importance of developing health interventions that are directed not only to individuals but also to populations, communities and a wide range of social and environmental factors influencing health (Sanson-Fisher, D'Este, Carey, Noble, & Paul, 2014; Upshur, 2002). CRTs offer an appropriate design to assess such public health interventions, and also to measure the overall effect of an intervention at the population level (Handlos et al., 2009; R. J. Hayes, Alexander, Bennett, & Cousens, 2000; Isaakidis & Ioannidis, 2003; Moberg & Kramer, 2015; Sanson-Fisher et al., 2014), heterogeneity of impact among population subgroups, and equity (Morris, Ranson, Sinha, & Mills, 2007; Ranson, Sinha, Morris, & Mills, 2006).

Implementation fidelity in CRTs of public health interventions

Although the scientific debate is ongoing (Grossman & Mackenzie, 2005), randomised trials are generally viewed as the gold standard for establishing evidence of intervention effectiveness. Despite this, the use of CRTs to evaluate public health interventions raises unique methodological challenges. Recent systematic reviews of CRT methods have found evidence of improvements in the design and analysis of CRTs while noting deficiencies in trial implementation that may compromise their validity (Bonell, Oakley, Hargreaves, Strange, & Rees, 2006; Eldridge, Ashby, Bennett, Wakelin, & Feder, 2008). Previous systematic reviews have emphasised the importance of process evaluation to mitigate these methodological problems, which can affect the

internal and external validity of trial results (Brierley, Brabyn, Torgerson, & Watson, 2012; Eldridge et al., 2008; Isaakidis & Ioannidis, 2003; Oakley, Strange, Bonell, Allen, & Stephenson, 2006; Puffer, Torgerson, & Watson, 2003).

“Implementation fidelity” refers to the degree to which an intervention is delivered as initially planned (Carroll et al., 2007). Fidelity assessment is an aspect of process evaluation that aims to understand and measure to what extent the intervention is being implemented as intended, with a view to clarifying relationships between intervention and its intended outcomes and learning what specific reasons have caused the success or failure of the intervention (Carroll et al., 2007; Medical Research & Council., 2008; Richards & Hallberg, 2015). Evaluation of implementation fidelity within trials has multiple benefits, which may include increased confidence in scientific findings, increased power to control for confounding factors and detect intervention effects, and increased ability to evaluate the performance of an intervention based on theory (Borrelli, 2011). Several studies have found that interventions implemented with high fidelity achieved better results in comparison with low fidelity interventions (Bradley, Wiles, Kinmonth, Mant, & Gantley, 1999; Dane & Schneider, 1998; Dusenbury, Brannigan, & Falco, 2003; Hasson, Blomberg, & Dunér, 2012; Keith, Hopp, Subramanian, Wiitala, & Lowery, 2010; Spoth, Redmond, & Shin, 2001; Thom, 2017).

Fidelity assessment can improve internal and external validity of CRTs (Eldridge et al., 2008) by providing evidence that the trial results are due to the intervention itself rather than to confounding variables and facilitating generalization of results to contexts that may differ substantially from the original trial setting (Carroll et al., 2007; Eldridge et al., 2008). Fidelity assessment may be particularly important for trials of public health interventions, as these interventions tend to be complex and constituted by multiple components (M. Campbell et al., 2000; Hawe, 2015) that may act independently or interdependently (Moore et al., 2015), leading to a greater potential for variation during implementation (Hawe, 2015).

Framework for the evaluation of implementation fidelity used in this review

Table 1 outlines the conceptual framework for evaluation of implementation fidelity used in this review. The framework is based principally on work by Carroll et al. (2007) and includes

elements of implementation fidelity and moderating factors that may affect the delivery process (Carroll et al., 2007). The framework was further refined by Hasson, who expanded the list of moderating factors considered in the framework (Hasson, 2010). We selected this framework to guide the review because it provides a comprehensive synthesis of previous work on implementation fidelity and has been widely influential.

Table 1 Conceptual framework for implementation fidelity used in this review^{1,2}

Fidelity Components	
Content	Defined as an attempt to establish the “active ingredients” of the intervention, for example, in a theory of change or logic model, and assess whether they have been delivered as planned.
Coverage	Refers to the degree to which all persons who met study inclusion criteria received the intervention
Frequency	Refers to whether the intervention was delivered with the regularity or frequency planned by its designers.
Duration	Establishes whether the intervention was delivered with the duration planned by its designers.
Moderating factors	
Comprehensiveness of intervention description	Factors such as the degree of intervention complexity, and whether the intervention description is complete or incomplete, vague or clear, may influence the degree of implementation fidelity
Strategies to facilitate implementation	Several support strategies may be used to optimise and to standardise implementation fidelity
Quality of delivery	Concerns whether an intervention is delivered in a way that increases the likelihood of achieving the desired health outcomes
Participant responsiveness	Intervention uptake depends on its acceptance by and acceptability to those receiving it. Low participant involvement or responsiveness may negatively impact intervention fidelity.
Recruitment ²	Refers to procedures that were used to attract potential program participants.

Context ²	Refers to surrounding social systems, such as structures and cultures of organizations and groups, and historical and concurrent activities and events
----------------------	--

¹Adapted from Carrol et al. ² These components were added by Hasson, 2010.

Fidelity assessment in CRT reporting guidelines

The Consolidated Standards of Reporting Trials (CONSORT) group was created to provide guidance to improve the quality and transparency of reporting of RCTs (Moher, Hopewell, Schulz, Montori, Gøtzsche, et al., 2010). The CONSORT Statement offers a checklist of essential items that should be included in reporting a RCT (Moher, Hopewell, Schulz, Montori, Gøtzsche, et al., 2010). Due to the increasing use of CRT designs, the CONSORT group proposed a version of the CONSORT statement for the reporting of cluster randomised trials in 2004 and updated these guidelines in 2012 (M. K. Campbell et al., 2004; M. K. Campbell, Piaggio, Elbourne, Altman, & Group, 2012).

The CONSORT statement recognises that the trial protocol for a given study may not have been followed fully for some trial participants for a wide variety of reasons, including failure to receive the entire intervention as planned (Moher, Hopewell, Schulz, Montori, Gøtzsche, et al., 2010). Cases of protocol non-adherence may influence the interpretation and credibility of the results and thus the validity of the conclusions (Borrelli, 2011; Eldridge et al., 2008; Schneeweiss, 2006; Thabane et al., 2010). To preserve the ability to make strong inferences about the intervention effect, CONSORT offers recommendations on how issues of non-adherence should be handled at the level of analysis. Specifically, it recommends that all participants randomised be retained in the analysis and analysed according to their original assigned groups, an approach known as “intention-to-treat” or “ITT” analysis. This approach ignores noncompliance, protocol deviations and anything that occurs after randomisation. The rationale for the ITT approach is that random allocation procedures avoid bias when assigning interventions to trial participants and thus facilitate causal inference. Any exclusion of patients from the analysis risks compromising the randomisation and may lead to biased results. This ITT approach can be contrasted with a “per protocol” or “PP” analysis, which restricts the analysis to participants who fulfil the protocol in terms of eligibility, interventions, and outcome assessment (Borrelli, 2011; Eldridge et al., 2008; Schneeweiss, 2006; Thabane et al., 2010). According to CONSORT, although a PP analysis may

be appropriate in some instances, due to the exclusion of participants, it should be considered as a non-randomised, observational comparison.

The CONSORT guidance on handling protocol non-adherence has been primarily developed in relation to individually randomised parallel group trials. However, reasons for protocol non-adherence in individually randomised RCTs may differ from those in CRTs. In a clinical trial setting, non-adherence depends largely on the actions of the trial participant (e.g. failure to adhere to therapy) and the treatment provider (e.g. failure to follow treatment protocol), which may in turn be related to issues such as treatment side effects and safety. In CRTs of public health interventions, protocol non-adherence may occur because complex interventions that include multiple components are delivered with poor fidelity. However, despite the scientific importance of protocol non-adherence, the current CONSORT guidelines for individually randomized parallel group trials (Moher, Hopewell, Schulz, Montori, Gotzsche, et al., 2010), and CRTs (M. K. Campbell et al., 2004; M. K. Campbell et al., 2012) offer no advice on methods to assess its occurrence during the course of a trial.

Rationale for undertaking this review

LMIC governments and other development partners have strengthened research and intervention efforts to support the UN Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs) agenda. As the global community intensifies the search for the best evidence on public health interventions to improve health and development outcomes in LMICs, CRTs have become an essential tool. Policy makers are interested in using the best available evidence to make decisions about the effectiveness of specific interventions in LMICs facing considerable budget constraints. Although CRTs have been widely implemented to evaluate public health interventions in both high-income countries and LMICs, country context, interventions, approaches, and outcomes may differ substantially between settings. We therefore limit our focus to LMICs.

As earlier methodologically-oriented systematic reviews have demonstrated, CRTs of complex public health interventions may be particularly at risk of experiencing protocol deviations and non-adherence, and these may compromise the validity of their findings (Bonell et al., 2006;

Eldridge et al., 2008). Although process evaluation techniques such as evaluation of implementation fidelity can help to assess the extent of these problems and mitigate their negative effects, current reporting guidelines for CRTs offer no specific guidance on assessment of intervention fidelity within CRTs. Wide divergence in current practices is therefore likely. We will undertake a methodologically-oriented systematic review of current practices related to assessment of intervention fidelity within CRTs of public health interventions in LMICs, with a view to informing best practices for these CRTs. To our knowledge, no other systematic review has been conducted on this question.

Objective:

We will conduct a systematic review of the published scientific literature to study current practices concerning assessment of intervention fidelity in cluster randomised trials (CRTs) of public health interventions in LMICs.

This review will address the following research questions:

1. Based on information from the trial registry (and the published study protocol, if applicable): What proportion of recent CRTs of public health interventions in LMICs planned to assess implementation fidelity (IF)?
2. Based on information from the published trial report (or a complementary document such as a published article, a grey literature report, or an online appendix reporting assessment of IF), what proportion of recent CRTs of public health interventions in LMICs reported assessing IF?
3. For those studies that assessed IF, which fidelity components were examined, and which data collection methods were employed to assess each component?
4. Is there evidence of divergent practices between planned and reported studies, or of outcome reporting bias related to assessment of IF?
 - a. Based on comparison of the results of questions 1 and 2, what is the overall agreement between planned and reported assessment of IF?
 - b. Are trial reports with negative findings for the ITT analysis more likely to report a

PP analysis?

- c. For the subset of studies that included both ITT and PP analyses, what is the overall agreement between ITT and PP analyses concerning the intervention's effectiveness?
- d. Does the magnitude of the intervention effect differ for PP as compared to ITT analyses?

To answer our research questions, we will first identify all CRTs from 2012 onwards of public health interventions conducted in LMICs with an available study protocol registered in a public trial registry. A given CRT will be included in the review if the protocol, the trial report, or both, address IF. For each CRT included in the review, we will compare planned assessment methods for IF as described in the trial registry (and published study protocol, if applicable) with published methods and results from the main trial report (and related documents, if relevant). We will use a variety of measures to summarise results.

METHODS

We describe study methods in seven steps adapted from the 2015 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-P reporting guidelines for systematic reviews and meta-analysis protocols (Moher D, Liberati, Tetzlaff, Altman, & Group, 2009). The PRISMA-P checklist is provided as an additional file (see Additional file 1). As this review focuses on methodological issues rather than health-related outcomes, it was not eligible for inclusion in the PROSPERO registry (York. & dissemination, 2011). In the event of protocol amendments, we will provide the date of each amendment, a description of the change, and the rationale for the change.

Eligibility criteria

Studies will be selected from the peer-reviewed scientific literature according to the following study and report characteristics.

Study Characteristics

Study Designs

We will include all cluster-randomised trials (CRTs). For the purposes of this review, a CRT is defined as a trial in which intact social units or clusters of individuals, rather than independent individuals, are randomly allocated to intervention groups (M. K. Campbell et al., 2012). CRTs may include trials employing parallel group, stepped wedge, or factorial designs; cluster-randomised adaptive trials, and cluster-randomised pragmatic trials, among others. CRTs with an adaptive design allow modifications based on data accumulated following trial start, while preserving the integrity of the trial (Moher, Hopewell, Schulz, Montori, Gøtzsche, et al., 2010). Pragmatic trials are designed to evaluate the efficacy of an intervention in routine clinical practice in order to maximize applicability and generalizability the results of the study (Schwartz & Lellouch, 1967; Zwarenstein et al., 2008).

Participants

Study participants will be human adults or children living in LMICs. LMICs will be defined according to the 2016 World Bank country classifications (The World & Bank.).

Interventions

This review focuses on “public health interventions”. We employ the definition of “public health” proposed in the World Health Organization (WHO) health promotion glossary as “The science and art of promoting health, preventing disease, and prolonging life through the organized efforts of society” (Nutbeam, 1998). Adapting the definition proposed by Rychetnik and colleagues, we define a public health intervention as a disease prevention or health promotion intervention applied to many, most, or all members in a community, which aims to deliver a net benefit to the community or population, as well as benefits to individuals (Rychetnik, Frommer, Hawe, & Shiell, 2002; Rychetnik, Hawe, Waters, Barratt, & Frommer, 2004). Public health interventions are distinguished from clinical interventions aimed at preventing or treating diseases at the individual level (Rychetnik et al., 2002; Rychetnik et al., 2004).

In order to operationalize this definition and guide selection of specific studies, we will use the “Intervention Wheel”, a graphic model of population-based public health practice illustrated with specific examples, developed by the Minnesota Department of Health (Keller, Strohschein,

Lia-Hoagberg, & Schaffer, 2004). The intervention wheel provides 17 public health interventions, selected to meet five criteria. To be considered as public health interventions, interventions should: (i) focus on entire populations (or particular subgroups within a population) (ii) be grounded in an assessment of community health; (iii) consider the broad determinants of health; (iv) emphasize health promotion and prevention, and (v) intervene at multiple levels (Keller et al., 2004). We used these five criteria to aid in decisions concerning study inclusion.

According to Rychetnik and colleagues, public health interventions are inherently “complex, programmatic, and context dependent” and these characteristics raise challenges for their evaluation (Rychetnik et al., 2002). Assessment of intervention fidelity may be especially important for public health interventions, and this consideration underlies our choice to focus on them in this review.

Comparators

Comparators will be defined as planned per the original CRT. Given the nature of public health interventions and the pragmatic orientation of CRTs in LMICs, we anticipate that a large proportion of studies included in the review will define the comparison group as receiving “usual care”.

Outcomes

The focus of this methodologically-oriented review is on comparisons of planned and reported outcomes related to IF. For studies that assessed IF in either the trial protocol or the main trial report, we will include both the study protocol and the main trial report. Recognising that word limits for scientific journal articles are highly constrained and that current CONSORT reporting guidelines for CRTs do not require description of elements related to IF, we also decided to include CRTs reporting assessment of IF in a complementary document such as a published article, an online appendix to the main paper, or a grey literature report, in lieu of reporting assessment of IF in the main trial report. These elements will be verified by checking the bibliography for the main trial report.

For purposes of study selection, we considered that studies evaluated IF if they either proposed methods to assess or reported results related to evaluation of at least one of the four key fidelity components: (1) content, (2) coverage, (3) frequency, and (4) duration. For CRTs taking an adaptive approach we will consider if these trials respect pre-established decision rules regarding changes to their design. In addition, we will include all CRTs that reported a per-protocol analysis.

Report characteristics

Setting

Eligible studies will be implemented in LMICs as classified by the World Bank (The World & Bank.).

Availability of the study protocol

To ensure availability of a study protocol, we will include CRTs reporting a registration number in the abstract for any trial registry meeting WHO criteria (World Health & Organization., 2016). The WHO trial registration data set (TRDS) is an internationally agreed-upon set of items that provide information on the design, conduct and administration of clinical trials. The WHO International Clinical Trials Registry Platform (ICTRP) facilitates publication of the TRDS on a publicly accessible website, through a network of partner registries that have agreed to adopt the TRDS as a common standard. The TRDS will be used in this review to evaluate planned assessment of intervention fidelity, either alone, or in conjunction with a published study protocol. TRDS field 20 “Key secondary outcomes” is particularly pertinent for this assessment.

Publication Dates

We will include studies for which the main trial report was published on or after January 1, 2012. We chose this date because the last update of the CONSORT statement to improve reporting of CRTs was published in 2012 and we wanted to analyse current practices. No restriction will be applied to the publication date for the protocol.

Language

We will include studies published in English, Spanish or French, which are languages known by the research team.

Exclusion criteria

We will exclude studies that are (i) not cluster randomized trials, studies that (ii) do not plan or report assessment of IF (iii) are not public health interventions, (iv) conducted in a high income country as defined by the World Bank 2016 country classification (The World & Bank.), (v) are published before 2012; (vi) do not have a publicly available protocol for comparison, or (vii) for which only the protocol but not the main trial report has been published. Manuscripts will be also excluded if they are (viii) pilot studies, (ix) secondary reports of a main study for which the relevant findings were published prior to 2012, (x) studies published in a language other than English, Spanish or French, or (xi) studies from the grey literature.

Information sources and search strategy

Literature search strategies were developed in collaboration with an academic librarian experienced in conducting systematic review searches. Search strategies use Medical Subject Headings (MESH) and text words related to cluster randomised trials, developing countries, and public health interventions. The electronic database search was developed first for MEDLINE (ovid) (for the full search strategies see Additional file 2) and then adapted for the following electronic databases: EMBASE (ovid), CINAHL (ovid), PUBMED and EMB Reviews (ovid). Search terms are a combination of “cluster-randomized”, “cluster analysis”, “health program”, “public health service”, “health education”, “public health”, “health promotion”, “health behavior”, “health knowledge/ attitudes practice”, “Preventive health services”, “health care system”, “health education” and “developing countries”. The search strategy will span the time period from January 2012 to May 2016, and will be updated towards the end of the review. Searches will be filtered to articles concerning humans, written in English, French, or Spanish. To augment this list, we will add relevant studies suggested by members of the systematic review team. Identified records will be uploaded into EndNote reference management software (version X7.5.3, Thompson Reuters, 2016) and duplicates will be eliminated.

Study screening and selection

Study screening and selection will be done manually within EndNote based on the inclusion and exclusion criteria for this systematic review. To ensure the availability of study protocols, we will limit the search to CRTs that have the word stem “regist*” in the abstract and use these results to begin the process of screening and selection. We validated this procedure by examining a subset of excluded articles. Screening and selection will be done in two stages by two independent reviewers (MCP and NM). In the first stage, reviewers will independently screen the titles and abstracts of each identified reference against the inclusion criteria to eliminate irrelevant publications. In the second stage, we will screen the full text of all studies that appear to meet the inclusion criteria or for which there is uncertainty as to eligibility. For each study, we will identify additional articles of potential relevance, such as a published protocol or a process evaluation, by reviewing references from the main trial report, consulting the trial registry record, and searching the PubMed database for new publications by the lead trial author. To aid in article screening and selection, the team will develop and test a screening sheet for full text review. Any disagreement between reviewers will be resolved through discussion and, as necessary, through arbitration by a third author (MJ). The process of study selection will be documented in a flow diagram describing studies identified and excluded at each stage. We will also provide a summary table describing studies excluded at the stage of full text review, along with reasons for their exclusion.

Outcomes and prioritization

The search and selection process for this review is designed to identify two quantities required for calculation of outcomes based on proportions: (1) Numerator: These are studies that meet all inclusion and exclusion criteria. As for all systematic reviews, these studies are our principal focus and will be included in the review and given detailed analysis. (2) Denominator: This is the total “N” for the study, which we defined as all studies that satisfy all inclusion and exclusion criteria, with the exception of the outcome criterion (planned or reported IF assessment). It is essentially the universe of cluster-randomised trials of public health interventions in LMICs. Both quantities will be clearly indicated in the study flow diagram.

Primary Outcome

The primary outcome for this study will be the proportion of overall agreement between the protocol and trial report concerning occurrence of IF assessment. This corresponds to research question 4a.

Data will be summarised in a two-by-two table comparing assessment of intervention fidelity in the trial report to that in the protocol. “N” represents the set of recent CRTs of public health interventions in LMICs that have registered the study protocol in a publicly available trial registry. For each CRT in N, we will determine whether IF was assessed in the registered (or published) protocol or in the trial report (or associated documents). Studies judged to have assessed IF will be coded as “1”; others will be coded as “0”. Judgments will represent reviewer consensus (MCP and NM, with appeal to MJ in case of divergences). The proportion of overall agreement is defined as the proportion of eligible CRTs for which judgments concerning occurrence of implementation fidelity assessment agree in the protocol and in the trial report (i.e. both positive, or both negative). It will be computed as $(a+d) / N$.

	Protocol		
	+	-	
Trial Report	+ a	b	a + b
	- c	d	c + d
	a + c	b + d	N

Secondary Outcomes

To address research questions 1, 2, and 3, we will also calculate:

- The frequency and proportion of trial *protocols* reporting assessment of intervention fidelity, out of N

- The frequency and proportion of trial *reports* reporting assessment of intervention fidelity, out of N
- The proportion of positive agreement among those that agree, computed as $a/(a+d)$
- Frequency counts and percentages summarising fidelity components examined and data collection methods proposed or employed

To address research question 4b, for all studies included in the trial, we will also record authors' judgments as to whether the intervention was effective. Studies that concluded that the intervention is more effective than control will be coded as "1"; studies that were unable to reject the null hypothesis that there are no significant differences between groups will be coded as "0".

We will calculate:

- The conditional probability that a PP analysis is performed given that the ITT analysis shows no difference between groups.
- The conditional probability that a PP analysis is performed given that the ITT analysis shows a positive intervention effect

These measures will be calculated using a standard formula for conditional probabilities:

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

To address research questions 4c and 4d, we will examine the subset of trial reports containing both ITT and PP analyses. For studies comparing several interventions (e.g. factorial design), data on each intervention will be extracted separately.

To address research question 4c, we will study the proportion of overall agreement between the ITT and PP analyses concerning intervention effectiveness.

Data will be summarised in a two-by-two table comparing assessment of intervention effectiveness in the ITT analysis to that in the PP (intervention fidelity) analysis. “T” is the total number of included CRTs reporting both an ITT and PP analysis. Studies that concluded in favour of the intervention group will be coded as “1”; those that are unable to reject the null hypothesis that there is no significant difference between groups will be coded as “0”. Judgments will represent reviewer consensus (MCP and NM, with appeal to MJ in case of divergences). The proportion of overall agreement is defined as the proportion trial reports for which judgments concerning intervention effectiveness agree in ITT and PP analyses (i.e. both positing (favour the intervention group), or both negative (unable to reject the null hypothesis of no difference between groups)). It will be computed as $(w+z) / T$.

	ITT analysis		
	+	-	
PP	+ w	x	w + x
analysis	- y	z	y + z
	w + y	x + z	T

We will also calculate

- The frequency and proportion of ITT analyses that conclude in favour of the intervention, out of T
- The frequency and proportion of PP analyses that conclude in favour of the intervention, out of T

To address research question 4d, we will compare intervention effect sizes reported for ITT and PP analyses. Comparisons will be summarized as the percentage change in effect size, computed as the effect size for the PP analysis / effect size for the ITT analysis * 100

Risk of bias in individual studies

To assess possible risk of bias for included studies, we will use the Cochrane Collaboration tool to assess the risk of bias in randomised trials (Higgins et al., 2011) based on the following factors: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting and other sources of bias. Because the Cochrane Collaboration tool was developed for individually randomised studies whereas our study focuses on CRTs, we will also include several additional criteria specifically relevant to assessing risk of bias in CRTs, recommended by the Cochrane collaboration (Higgins et al., 2011) and other key sources (Donner & Klar, 2000; R. Hayes & Moulton, 2009; Higgins et al., 2011). These additional criteria will consider issues related to: recruitment bias (potential for participant self-selection to occur if individuals are recruited to the trial after the clusters have been randomized), baseline imbalances (because CRTs generally randomize a limited number of clusters, chance imbalances may affect comparability of intervention and control groups), loss of clusters (complete clusters may sometimes be lost from a trial and thus be omitted from the analysis; these missing data may lead to biased outcome assessments), and unit of analysis (failure to properly account for clustering in the analysis) (Higgins et al., 2011). For each domain or criterion of interest, we will assess each criterion as low risk, high risk or uncertain risk, and provide a sample text that illustrates the reasons for this judgment. This evaluation will be done independently by two reviewers (MCP and NM). Disagreements between reviewers will be resolved by consensus or, if consensus cannot be achieved, by consulting a third reviewer (MJ). Judgments related to risk of bias will be summarised graphically using RevMan 5.1 (Higgins et al., 2011). Risk of bias assessments will be used to create categories of high, uncertain, and low risk studies to be used in subgroup analyses.

Systematic reviews of health outcomes often assess the quality of a body of evidence using standardized tools such as the GRADE system (Guyatt et al., 2008). However, as this review focuses on methodological issues rather than health-related outcomes, we will not use this tool.

Data extraction and data items

Two review authors will extract data independently (MCP and NM). From each study protocol and trial report, reviewers will extract data on (i) study characteristics (study location, aims, intervention); (ii) all applicable descriptors of the CRT trial design (for example, parallel group, stepped wedge, factorial, adaptive, pragmatic); (iii) concepts related to assessment of IF (assessment of fidelity reported in protocol and/or main study, fidelity components and moderating factors evaluated, data collection methods, and any dimension used by the authors to evaluate intervention fidelity distinct from those proposed by Carroll and Hasson (Carroll et al., 2007; Dusenbury et al., 2003); (iv) whether events taking place in the control group were monitored, as these can influence the effectiveness of the intervention (Durlak & DuPre, 2008; Hasson et al., 2012; Ibrahim & Sidani, 2015); and (v) information for assessing the risk of bias of included studies. We will also extract (vi) statistical results concerning intervention effectiveness, and the authors' qualitative conclusions regarding the intervention effect for the primary (generally, ITT) analysis and one or more subgroup analyses relevant for intervention fidelity (generally, the PP analysis). If studies investigate more than one intervention, we will extract data relevant for each comparison. To reduce bias and errors in data extraction, reviewers will use a pre-defined template pilot tested on a subset of studies, and a guide for data extraction. To ensure consistency, reviewers will receive training prior to commencing extraction for the review and undertake calibration exercises. Reviewers will resolve disagreements by discussion, and by appeal to a third author (MJ) where necessary. All data extraction tools will be available as online supplementary documents.

Data synthesis

Results will be presented in accordance with the PRISMA Statement (Moher D et al., 2009). A narrative synthesis will be provided, with information presented in tables to summarise key data. The narrative synthesis will explore relationships and findings within and between the included studies. It will highlight the four key dimensions of intervention fidelity identified from the literature (content, coverage, frequency and duration), moderating factors for intervention fidelity (participant responsiveness, comprehensiveness of policy, strategies to facilitate implementation, quality of delivery, recruitment, and context), any new dimensions explored, and data collection

method used to evaluate each key dimension.

We will present quantitative data for all primary and secondary outcomes proposed. Where appropriate, data will be presented in tabular form.

We will investigate possible sources of heterogeneity by performing subgroup analysis. Specifically, we will recompute the main quantitative outcomes for subgroups of studies with high, uncertain and low risk of bias to better understand potential sources of variation in results. If the data permit, we will conduct a sensitivity analysis to explore whether studies at lower risk of bias undertake more comprehensive assessment of intervention fidelity. Because of the study question and the nature of the outcomes assessed, we do not intend to perform meta-analyses.

Planned assessment of meta-biases

We recognize that data may be biased due to non-study related processes and plan to assess specific meta-biases. This study compares results for protocols and published trial reports and is thus designed to address potential reporting bias and to investigate potential outcome bias. As our review focuses on methodological issues rather than outcomes assessment, we will not assess potential publication bias.

DISCUSSION

Development initiatives require high quality evaluations to determine whether the programs work or not and to know how to improve them (Banerjee & Duflo, 2011; Duflo, Glennerster, & Kremer, 2007). According to Rychetnik et al (Rychetnik et al., 2004), evaluation of public health interventions requires detailed information about the "design and implementation of an intervention; contextual circumstances in which the intervention was implemented; and how the intervention was received".

We will conduct a methodological systematic review to evaluate current practices for evaluating implementation fidelity in CRTs of public health interventions carried out in in low-

and middle-income countries (LMICs). Fidelity assessment may be a key tool for making studies more reliable, internally valid and externally generalizable (Bellg et al., 2004). In the absence of fidelity assessment, it may be difficult to determine if CRT results are due to the intervention design, to its implementation, or to unknown or external factors that may influence results. The rejection of effective interventions or acceptance of ineffective interventions incurs incalculable costs, due to use of financial and scientific resources and the inability of the authors to extrapolate the results (Borrelli, 2011). Improved assessment and reporting of intervention fidelity may be important for researchers, for those who finance health interventions, and for decision makers who seek the best evidence on public health interventions to promote health, prevent disease and reduce health inequalities.

List of abbreviations

CONSORT: Consolidated Standards of Reporting Trials

CRTs: Cluster randomization trials

ICTRP: International Clinical Trials Registry Platform

IF : intervention fidelity

ITT: intention-to-treat

LMICs: low- and middle-income countries

MESH: Medical Subject Headings

PP: per protocol

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCTs: Randomized controlled trials

TRDS: trial registration data set

WHO: World Health Organization

DECLARATIONS

Ethical Approval and Consent to participate

This research does not involve human subjects. It is exempt from research ethics board review as it relies exclusively on publicly available information for which there is no reasonable expectation of privacy.

Consent for publication

Not applicable

Availability of supporting data

Not applicable

Competing interests

The authors declare that they have no competing interests

Funding

IC-IMPACTS (the India-Canada Centre for Innovative Multidisciplinary Partnerships to Accelerate Community Transformation and Sustainability) provided funding for this study in the form of doctoral scholarships for MCP and NM. V. Ridde holds a CIHR-funded Research Chair in Applied Public Health (CPP-137901) and is an IC-IMPACTS co-investigator. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

MJ and MCP developed the idea for the systematic review and lead development of the study protocol. MCP and MJ prepared the first draft of the protocol; MPS provided statistical guidance. NM, VR and MPS reviewed the protocol for important intellectual content and provided feedback. All authors read and approved the final manuscript for submission. MJ assumes overall responsibility for the scientific integrity of the work.

Acknowledgements

We would like to acknowledge Daniela Ziegler, Librarian at the University of Montreal Hospital (CHUM) for her help with the database search strategy, and Professor Christina Zarowsky, University of Montreal, for helpful comments on an earlier manuscript draft.

Author's Information

Not applicable

Additional file 1

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols)

2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Reported on page #
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	NA
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	10
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	24,25
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	NA
Support:			
Sources	5a	Indicate sources of financial or other support for the review	24
Sponsor	5b	Provide name for the review funder and/or sponsor	24

Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	25
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	4-8
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	8-9
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	10-14
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	14-15
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Additional file 2
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	15
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	15-16

Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	15-16
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	20-21
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	16-19
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	19-20
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	21-22
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I ² , Kendall's τ)	21-22
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	21-22
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	21-22
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	22

Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	20
-----------------------------------	----	--	----

* It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647

Additional file 2

The search strategy for MEDLINE

Database: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R)

Daily and Ovid MEDLINE(R) 1946 to Present Interface: OvidSP

Research date: May 13th, 2016

Filters 2012 onwards, English, French, Spanish

1	Cluster Analysis/	49336
2	Cluster Analys*.tw,kw,sh.	59132
3	(communit* adj2 random* adj2 trial*).tw,kw,sh.	630
4	(group adj2 random* adj2 trial*).tw,kw,sh.	2007
5	(cluster adj2 random*).tw,kw,sh.	7944
6	cluster-random*.tw,kw,sh.	6686
7	or/1-6	67832
8	Afghanistan/ or Libya/ or Albania/ or Macedonia Republic/ or Algeria/ or Madagascar/ or American Samoa/ or Malawi/ or Angola/ or Malaysia/ or Argentina/ or Indian Ocean Islands/ or Armenia/ or Mali/ or Azerbaijan/ or Micronesia/ or Bangladesh/ or Mauritania/ or Republic of Belarus/ or Mauritius/ or Belize/ or Mexico/ or Benin/ or Bhutan/ or Moldova/ or Bolivia/ or Mongolia/ or Bosnia-Herzegovina/ or Montenegro/ or Botswana/ or Morocco/ or Brazil/ or Mozambique/ or Bulgaria/ or Myanmar/ or Burkina Faso/ or Namibia/ or Burundi/ or Nepal/ or Nicaragua/ or Cambodia/ or Niger/ or Cameroon/ or Nigeria/ or Central African Republic/ or Pakistan/ or Chad/ or Palau/ or China/ or Panama/ or Colombia/ or Papua New Guinea/ or Comoros/ or Paraguay/ or Democratic Republic of the Congo/ or Congo/ or Peru/ or Philippines/ or Costa Rica/ or Romania/ or Cote d'Ivoire/ or Rwanda/ or Cuba/ or Samoa/ or Djibouti/ or Atlantic Islands/ or Dominica/ or Senegal/ or Dominican Republic/ or Serbia/ or Ecuador/	823062

	<p>or Seychelles/ or Egypt/ or Sierra Leone/ or El Salvador/ or Melanesia/ or Eritrea/ or Somalia/ or Ethiopia/ or South Africa/ or Fiji/ or Sudan/ or Gabon/ or Sri Lanka/ or Gambia/ or Saint Lucia/ or Georgia/ or "Saint Vincent and the Grenadines"/ or Ghana/ or Grenada/ or Suriname/ or Guatemala/ or Swaziland/ or Guinea/ or Syria/ or Guinea-Bissau/ or Tajikistan/ or Guyana/ or Tanzania/ or Haiti/ or Thailand/ or Honduras/ or East Timor/ or Hungary/ or Togo/ or India/ or Tonga/ or Indonesia/ or Tunisia/ or Iran/ or Turkey/ or Iraq/ or Turkmenistan/ or Jamaica/ or Micronesia/ or Jordan/ or Uganda/ or Kazakhstan/ or Ukraine/ or Kenya/ or Uzbekistan/ or Vanuatu/ or Korea/ or Venezuela/ or Kosovo/ or Vietnam/ or Kyrgyzstan/ or Middle East/ or Yemen/ or Lebanon/ or Lesotho/ or Zambia/ or Zimbabwe/ or Liberia/ or Developing Countries/ or Poverty/ or Rural Population/</p>	
9	<p>(Afghanistan or Albania or Algeria or American Samoa or Angola or Argentina or Armenia or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bhutan or Bolivia or Bosnia Herzegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burundi or Cabo Verde or Cambodia or Cameroon or Central African Republic or Chad or China or Colombia or Comoros or Congo or Costa Rica or Cote d'Ivoire or Cuba or Djibouti or Dominica or Dominican Republic or Ecuador or Egypt or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gambia or Georgia or Ghana or Grenada or Guatemala or Guinea or Guinea-Bissau or Guyana or Haiti or Honduras or Hungary or India or Indonesia or Iran or Iraq or Jamaica or Jordan or Kazakhstan or Kenya or Kiribati or Korea or Kosovo or Kyrgyz or Lao PDR or Lebanon or Lesotho or Liberia or Libya or Macedonia or Madagascar or Malawi or Malaysia or Maldives or Mali or Marshall Islands or Mauritania or Mauritius or Mexico or Micronesia or Moldova or Mongolia or Montenegro or Morocco or Mozambique or Myanmar or Namibia or Nepal or Nicaragua or Niger or Nigeria or Pakistan or Palau or Panama or Papua New Guinea or Paraguay or Peru or</p>	792287

	Philippines or Romania or Rwanda or Samoa or Sao Tome or Senegal or Serbia or Seychelles or Sierra Leone or Solomon Islands or Somalia or South Africa or South Sudan or Sri Lanka or St Lucia or Sudan or Suriname or Swaziland or Syrian Arab Republic or Tajikistan or Tanzania or Thailand or Timor Leste or Togo or Tonga or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Vanuatu or Venezuela or Vietnam or Gaza or Yemen or Zambia or Zimbabwe or Developing Countr* or middle income countr* or low income countr*).tw,kw.	
10	8 or 9	1172790
11	((public adj2 health) or (poverty adj2 economic*) or interven* or program* or approach*).tw,kw,sh.	2459336
12	Public Health/ or Community Health Services/ or Community Health Workers/ or Community Health Planning/ or Community Networks/	107616
13	11 or 12	2515327
14	7 and 10 and 13	4068
15	remove duplicates from 14	3964
16	limit 15 to yr="2012 -Current"	2069
17	limit 16 to (english or french or spanish)	2024

REFERENCES

- Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*: Public Affairs.
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Treatment Fidelity Workgroup of the, N. I. H. B. C. C. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*, *23*(5), 443-451. doi:10.1037/0278-6133.23.5.443
- Bland, J. M. (2004). Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol*, *4*(1), 21. doi:10.1186/1471-2288-4-21
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V., & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ*, *333*(7563), 346-349. doi:10.1136/bmj.333.7563.346
- Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent*, *71*(s1), S52-63.
- Bradley, F., Wiles, R., Kinmonth, A.-L., Mant, D., & Gantley, M. (1999). Development and evaluation of complex interventions in health services research: case study of the Southampton heart integrated care project (SHIP). *BMJ*, *318*(7185), 711-715.
- Brierley, G., Brabyn, S., Torgerson, D., & Watson, J. (2012). Bias in recruitment to cluster randomized trials: a review of recent publications. *J Eval Clin Pract.* , *18*(4), 878-886.
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ*, *321*(7262), 694-696. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10987780>
- Campbell, M. K., Elbourne, D. R., Altman, D. G., & group, C. (2004). CONSORT statement: extension to cluster randomised trials. *BMJ*, *328*(7441), 702-708. doi:10.1136/bmj.328.7441.702
- Campbell, M. K., Mollison, J., Steen, N., Grimshaw, J. M., & Eccles, M. (2000). Analysis of cluster randomized trials in primary care: a practical approach. *Fam Pract*, *17*, 192-196.
- Campbell, M. K., Piaggio, G., Elbourne, D. R., Altman, D. G., & Group, C. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ*, *345*, e5661. doi:10.1136/bmj.e5661
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci*, *2*, 40. doi:10.1186/1748-5908-2-40

- Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., & Thompson, S. (2012). Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health*, *66*(12), 1182–1186.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical psychology review*, *18*(1), 23-45.
- Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, *4*, 3895-3962.
- Durlak, J. A., & DuPre, E. P. J. A. j. o. c. p. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *41*(3-4), 327-350.
- Dusenbury, L., Brannigan, R., & Falco, M., Hansen WB. . (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res*, *18*(2), 237–256.
- Eldridge, S., Ashby, D., Bennett, C., Wakelin, M., & Feder, G. J. B. (2008). Internal and external validity of cluster randomised trials: systematic review of recent trials. *336*(7649), 876-880.
- Evidence-Based, Medicine Working, & Group. (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*, *268*(17), 2420.
- Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard? *Perspectives in biology and medicine*, *48*(4), 516-534.
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, *336*(7650), 924-926.
- Handlos, L. N., Chakraborty, H., & Sen, P. K. (2009). Evaluation of cluster-randomized trials on maternal and child health research in developing countries. *Trop Med Int Health*, *14*(8), 947–956.
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci*, *5*, 67. doi:10.1186/1748-5908-5-67
- Hasson, H., Blomberg, S., & Dunér, A. (2012). Fidelity and moderating factors in complex interventions: a case study of a continuum of care program for frail elderly people in health and social care. *Implement Sci*, *7*(23), 1–11.

- Hawe, P. (2015). Lessons from complex interventions to improve health. *Annu Rev Public Health*, 36, 307–323.
- Hayes, R., & Moulton, L. (2009). *Cluster randomised trials*.
- Hayes, R. J., Alexander, N. D., Bennett, S., & Cousens, S. N. (2000). Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Methods Med Res*, 9(2), 95–116.
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., & Sterne, J. A. (2011). The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *BMJ*, 343, d5928.
- Ibrahim, S., & Sidani, S. J. H. (2015). Fidelity of intervention implementation: a review of instruments. 7(12), 1687.
- Isaakidis, P., & Ioannidis, J. P. (2003). Evaluation of cluster randomized controlled trials in sub-Saharan Africa. *Am J Epidemiol*, 158(9), 921–926.
- Keith, R. E., Hopp, F. P., Subramanian, U., Wiitala, W., & Lowery, J. C. (2010). Fidelity of implementation: development and testing of a measure. *Implementation Science*, 5(1), 99.
- Keller, L. O., Strohschein, S., Lia-Hoagberg, B., & Schaffer, M. A. (2004). Population-based public health interventions: practice-based and evidence-supported. *Part I. Public Health Nurs*, 21(5), 453–468.
- Medical Research, & Council. (2008). *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council.
- Moherg, J., & Kramer, M. (2015). A brief history of the cluster randomised trial design. *J R Soc Med*, 108(5), 192–198. doi:10.1177/0141076815582303
- Moher D, Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, 6 (7), e1000097.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., & D.G., A. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*, 63(8), e1–e37.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c869. doi:10.1136/bmj.c869
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350, h1258.

- Morris, S. S., Ranson, M. K., Sinha, T., & Mills, A. J. (2007). Measuring improved targeting of health interventions to the poor in the context of a community-randomised trial in rural India. *Contemp Clin Trials*, 28(4), 382–390.
- Nutbeam, D. (1998). Health promotion glossary. *Health promotion international*, 13(4), 349-364.
- Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ*, 332(7583), 413–416.
- Osrin, D., Azad, K., Fernandez, A., Manandhar, D. S., Mwansambo, C. W., Tripathy, P., & Costello, A. M. (2009). Ethical challenges in cluster randomized controlled trials: experiences from public health interventions in Africa and Asia. *Bull World Health Org*, 87(10), 772–779.
- Puffer, S., Torgerson, D., & Watson, J. (2003). Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ*, 327(7418), 785–789.
- Ranson, M. K., Sinha, T., Morris, S. S., & Mills, A. J. (2006). CRTs—cluster randomized trials or “courting real troubles”: challenges of running a CRT in rural Gujarat, India. *Can J Public Health*, 97(1), 72.
- Richards, D. A., & Hallberg, I. R. (2015). Complex interventions in health: an overview of research methods. *London & New York: Routledge*.
- Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*, 56, 119–127. doi:doi:10.1136/jech.56.2.119
- Rychetnik, L., Hawe, P., Waters, E., Barratt, A., & Frommer, M. (2004). A glossary for evidence based public health. *J Epidemiol Community Health*, 58(7), 538–545.
- Sanson-Fisher, R. W., D’Este, C. A., Carey, M. L., Noble, N., & Paul, C. L. (2014). Evaluation of systems-oriented public health interventions: alternative research designs. *Annu Rev Public Health*, 35, 9–27.
- Schneeweiss, S. (2006). Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*, 15(5), 291–303.
- Schwartz, D., & Lellouch, J. (1967). Explanatory and pragmatic attitudes in therapeutical trials. *Journal of chronic diseases*, 20(8), 637-648.
- Spoth, R. L., Redmond, C., & Shin, C. (2001). Randomized trial of brief family interventions for general populations: adolescent substance use outcomes 4 years following baseline. *Journal of consulting and clinical psychology*, 69(4), 627.

- Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., . . . Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol*, *10*, 1. doi:10.1186/1471-2288-10-1
- The World, & Bank. Country and lending groups. Data & statistics: country classification (2016). In. Washington, D.C: . Available at: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519> Accessed 3 July,2016.
- Thom, B. (2017). Good practice in school based alcohol education programmes. *Patient education and counseling*, *100*, S17-S23.
- Upshur, R. E. (2002). Principles for the justification of public health intervention. *Can J Public Health*, *93*(2), 101–103.
- World Health, & Organization. (2016). *WHO international clinical trials registry platform (ICRP). New standards for registration of human medical research. Trial registration data set (version 1.2.1)*. Retrieved from Available at <http://www.who.int/ictrp/network/trds/en/>
- York., U. o., & dissemination, C. f. r. (2011). PROSPERO: international prospective register of systematic reviews. Retrieved from www.crd.york.ac.uk/PROSPERO.
- Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., . . . Moher, D. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*, *337*, a2390.

2.2 Deuxième volet : Étude de cas

2.2.1 Démarche scientifique

Pour ce deuxième volet, nous avons privilégié une étude de cas unique à trois unités d'analyse imbriquées selon une approche mixte concomitante (R. Yin, 2009). Une étude de cas est une stratégie de recherche synthétique, qui examine un phénomène contemporain au sein de son contexte réel (Eisenhardt, 1989). Elle vise la compréhension en profondeur des dynamiques présentes dans leur contexte, en combinant des méthodes de collecte et des sources de données (Eisenhardt, 1989; Muse & Gagnon, 2012). L'étude de cas peut fournir des informations riches et significatives sur le fonctionnement de l'intervention, et améliorer la compréhension sur le phénomène étudié, offrant ainsi une excellente validité interne de ses résultats (Contandriopoulos, Champagne F, Potvin L, Denis JL, & P, 2005). Selon Brown (2008, p.10), « les études de cas fournissent une compréhension humaniste et holistique des situations complexes, et, en tant que telles, sont de précieux outils de recherche » (P. A. Brown, 2008). Par ailleurs, l'étude de cas peut s'appuyer sur une épistémologie post-positiviste, et recourir à des méthodes mixtes pour la considérer comme une stratégie de recherche valide, pour étudier l'intervention à évaluer qui prend place dans un contexte et à un moment précis pour déterminer la démarche méthodologique à suivre (Creswell & Poth, 2016; Stake, 2008). Malgré que la portée de l'étude de cas unique soit limitée et que la question de la transférabilité de résultats associée à la validité externe demeure problématique, la qualité des liens entre ces choix méthodologiques et la qualité du raisonnement théorique sont essentielles pour améliorer la compréhension théorique d'un phénomène, qui permettra la transférabilité possible à un contexte particulier et à la théorie existante (Ayerbe & Missonier, 2007; Laperrière, 1997).

Le choix du cas est une étape fondamentale qui est guidé par les concepts théoriques sur le sujet à étudier. Notre cas est l'intervention Tika Vaani, qui est évaluée à travers une étude pilote sous le design d'un essai randomisé en grappes. Cette intervention est mise en place dans des villages de la zone rurale d'Hardoi, Uttar Pradesh en Inde. La littérature documente que ce type d'intervention peut être dépendante du contexte, et peut varier pendant leur mise en œuvre, ce qui le rend favorable à faire l'objet d'une évaluation de fidélité. L'étude de cas unique avec différentes

unités d'analyses permet d'expliquer un phénomène à partir des différents points de vue, ce qui permet souvent de mieux répondre à la question de recherche (Contandriopoulos et al., 2005). En ce qui concerne les unités d'analyse dans notre cas, ils sont représentés par les intervenants-terrain, la communauté et les travailleurs de la santé.

Quant à la conception de l'approche mixte concomitante (convergente), des approches quantitatives et qualitatives sont complémentaires, soit tout au long, soit à certaines étapes du processus de recherche telles que la collecte ou l'analyse des données, ou les deux. Puis, les deux ensembles de données sont comparés pour voir s'il existe une convergence, des différences ou une combinaison des deux (Creswell & Clark, 2017). Généralement, la priorité de chaque approche est égale, mais dans la pratique, la priorité peut être donnée à l'approche quantitative ou qualitative (P. Pluye, Bengoechea, & Granikov, 2018). Les résultats des deux méthodes sont intégrés lors de la phase d'interprétation. Cette interprétation peut soit noter la convergence des résultats comme un moyen de renforcer les revendications des connaissances de l'étude, soit expliquer tout manque de convergence qui pourrait en résulter (Creswell & Clark, 2017). Pour ce volet, la question de la recherche est abordée sur le même sujet (soit quel est le niveau de fidélité obtenu et quels facteurs expliquent ce niveau atteint), et l'objectif est généralement de trianguler les deux ensembles de données qui ont le même poids dans la recherche pour vérifier et mieux comprendre les résultats obtenus.

Afin de répondre à un besoin des concepteurs de l'intervention Tika Vaani pour évaluer la fidélité de la mise en œuvre de l'intervention, cette recherche évaluative cherche à fournir des résultats utiles à l'amélioration de l'intervention avant de l'appliquer à grande échelle. Ce volet de notre projet de recherche s'inscrit dans le champ de l'évaluation de la mise en œuvre. Selon Patton (2008), l'évaluation de l'intervention qui vise la mise en œuvre s'intéresse à fournir l'information des conditions qui façonnent et influencent les résultats, les variations dans la mise en œuvre et les effets apportés (M. Patton, 2008). Donc, cette évaluation permet de préciser le fonctionnement du programme et le processus qui entoure sa mise en œuvre. Pour Weiss (1997) et Chen (2004), cette évaluation aide à comprendre pourquoi et comment l'intervention produit les effets en vue de l'améliorer (H.-T. Chen, 2004; Weiss, 1997). Donc, dans un contexte d'évaluation formative, cette évaluation favorise le transfert de connaissances pour les concepteurs de l'intervention, afin de

l'améliorer et de s'assurer que l'intervention atteigne les résultats visés (Dehar et al., 1993; Nebot et al., 2011; M. Patton, 2008).

2.2.2 Objectif de recherche

L'objectif de ce volet de recherche consiste à évaluer la plausibilité et la fidélité de la mise en œuvre de l'intervention Tika Vaani mise en place dans le contexte d'un essai randomisé en grappes dans la zone rurale d'Hardoi (Uttar Pradesh) en Inde. Cette évaluation vise trois objectifs spécifiques suivants, auxquels répondent deux articles :

- a. Comprendre la conception de l'intervention à évaluer afin d'analyser leur plausibilité
- b. Mesurer le degré de fidélité de la mise en œuvre de l'intervention Tika Vaani
- c. Comprendre le rôle des facteurs modérateurs dans le degré d'hétérogénéité et de la fidélité de la mise en œuvre de l'intervention Tika Vaani.

Le but ultime de cette évaluation est de contribuer à l'amélioration de l'intervention si nécessaire, avant de la mettre en place à grande échelle.

2.2.3 Méthode de recherche

La démarche méthodologique pour ce volet de la recherche est présentée en détail dans deux des articles (Article 2 et Article 3) que l'on retrouve au Chapitre Résultats. De façon synthétique, nous présentons ici la démarche méthodologique pour chacun de nos objectifs spécifiques :

- a. Comprendre la conception de l'intervention à évaluer afin d'analyser leur plausibilité

Afin de mieux comprendre la conception de l'intervention à évaluer (Article 2), nous avons fait une analyse documentaire à partir des registres de l'intervention pour mieux comprendre son origine et son développement. Par la suite, nous avons fait des échanges avec les concepteurs de l'intervention pour valider les informations et pour valider la compréhension de l'intervention. Étant donné que l'intervention cible deux stratégies, soit la mobilisation communautaire, et soit l'utilisation de la technologie mobile comme ressources pour transmettre le contenu de l'information, nous nous sommes inspirés du guide mERA (anglais: mHealth evidence reporting and assessment) pour la documenter. Cette guide est un outil développé par le groupe d'examen des preuves techniques mHealth de l'Organisation Mondiale de la Santé pour améliorer la transparence des rapports sur les interventions de santé, qui utilisent la technologie mobile dans ses interventions (Agarwal et al., 2016). Donc, ce guide a permis de fournir des informations concernant le processus du développement, de la mise en œuvre et des leçons de l'intervention de Tika Vaani pour sa mise en place à grande échelle. Cette étude contribue à la littérature pour améliorer les rapports de ce type d'interventions connues sous le nom de « santé mobile ». De même, et comme un complément de l'évaluation de l'intervention, cet article permet de fournir aux chercheurs des points clés à retenir pendant le développement des interventions, qui visent des communautés difficiles à atteindre pour améliorer les résultats de santé.

D'autre part, pour évaluer et analyser la plausibilité d'une intervention, il faut d'abord définir et établir leur modèle logique (H.-t. Chen & Rossi, 1992). Selon Chen et Rossi (1992), l'analyse logique vise à tester la plausibilité de la théorie d'un programme en définissant les actions à accomplir, et en identifiant des alternatives pour atteindre les effets souhaités (H.-t. Chen & Rossi, 1992). Cette analyse se base sur des connaissances scientifiques ou des experts, pour évaluer la validité de la théorie de l'intervention (Brousselle et al., 2011; Rossi, Lipsey, & Henry, 2018). L'analyse logique cible deux modèles : le modèle logique théorique qui évalue si l'intervention repose sur une théorie adéquate quant à ses mécanismes causaux, et le modèle logique opérationnel qui vise à évaluer, si des ressources et des activités sont de qualité et quantité suffisantes pour atteindre les résultats escomptés (Brousselle et al., 2011).

À cet effet, l'analyse logique a été l'approche utilisée pour analyser la plausibilité de la théorie d'intervention. Pour ce faire, nous avons fait d'abord la révision des documents développés

par les concepteurs (le projet, les outils, les rapports et les documents de travail), afin d'identifier les différentes composantes de l'intervention. Ensuite, nous avons fait une révision de la littérature (des revues systématiques et des guides directifs de l'OMS) concernant l'utilisation d'une plateforme mobile, la théorie du changement du comportement et la vaccination des enfants, en raison de leur pertinence quant aux objectifs de l'intervention, afin de développer un modèle opérationnel pour représenter le lien entre les objectifs, les activités et les ressources mises en place. Enfin, des rencontres avec les intervenants-terrain et les concepteurs de l'intervention ont été menées pour réviser et valider le modèle logique développé (Article 3). Ce modèle logique a servi de base de référence pour évaluer la fidélité de la mise en œuvre de l'intervention.

- a. Mesurer le degré de fidélité de la mise en œuvre de l'intervention Tika Vaani
- b. Comprendre le rôle des facteurs modérateurs dans le degré d'hétérogénéité et de la fidélité de la mise en œuvre de l'intervention Tika Vaani.

La démarche méthodologique de cette phase est présentée en détail dans un des articles scientifiques (Article 3) du Chapitre Résultats de cette thèse. Nous le présentons ici de manière synthétique. Afin de mesurer le degré de fidélité de la mise en œuvre de l'intervention, dans le contexte d'un essai randomisé en grappes et d'un projet de recherche, nous émettons l'hypothèse que le degré de la fidélité de la mise en œuvre de l'intervention Tika Vaani sera élevé (supérieur à 80%) (B. Borrelli, 2011; Lorencatto et al., 2016).

Dans l'étude de cas, tenir compte d'un cadre conceptuel peut être un outil clé pour guider le chercheur à élaborer les outils de collecte de données, et délimiter la collecte de données par rapport à l'objet étudié (Miles & Huberman, 2003). Donc, afin de mesurer la fidélité de la mise en œuvre de l'intervention, et pour mieux comprendre les facteurs qui peuvent influencer la mise en œuvre de l'intervention telle que prévue par les concepteurs, et aussi pour mieux comprendre leur niveau de fidélité atteint, nous avons utilisé le cadre conceptuel de Carrol (2007), modifié par Hasson (2010), car selon l'état des connaissances, ce cadre conceptuel est plus complet pour analyser les dimensions de la fidélité (adhésion + facteurs modérateurs), selon une approche multidimensionnelle qui va au-delà de la dimension classique limitée à l'adhésion (contenu, fréquence, durée et couverture) (Carroll et al., 2007; Hasson, 2010). En effet, ils proposent

d'évaluer des facteurs modérateurs tels que la description détaillée de l'intervention, les stratégies pour faciliter la mise en œuvre, la qualité de l'intervention, la réponse des participants, le recrutement et le contexte, afin d'avoir une compréhension plus complète de la fidélité d'intervention. Ces dimensions permettent à leur tour d'explorer l'acceptabilité de l'intervention avant son passage à grande échelle. Ce cadre a été testé par d'autres auteurs, en soulignant que c'est un outil très utile pour évaluer la fidélité des interventions complexes telles que les essais randomisés en grappes (Lawton et al., 2017; M. C. Perez, Minoyan, Ridde, Sylvestre, & Johri, 2018), ainsi qu'un outil pour comprendre des adaptations qui pourraient se produire pendant la mise en œuvre (von Thiele Schwarz et al., 2015). Dans le contexte d'un essai randomisé en grappes, il faut aussi évaluer si des événements se déroulent dans le groupe contrôle, car ils peuvent influencer l'efficacité de l'intervention (Graham et al., 2016; Hasson et al., 2012; Ibrahim & Sidani, 2015). Ainsi, cette dimension a été aussi évaluée pour mieux comprendre le degré d'hétérogénéité et de la fidélité de la mise en œuvre et avoir un portrait complet de l'évaluation.

Lors d'une étude de cas, les sources d'information doivent être multiples afin de permettre l'analyse d'une variété de données, de renforcer la validité de construit, d'assurer la cohérence et de démontrer la fiabilité des données (Creswell & Clark, 2017; Gagnon, 2005; R. Yin, 2009). Pour notre cas, les sources d'information proviennent de trois groupes de la population à l'étude: (i) au niveau communautaire, la communauté de villages du site d'intervention qui veulent y participer et les parents des enfants ayant moins de deux ans, qui est la population cible de l'intervention; (ii) au niveau des ressources humaines de la santé, les travailleurs de santé communautaire qui travaillent dans le village touché par l'intervention; et (iii) au niveau des intervenants, l'équipe de terrain en charge de la mise en œuvre de l'intervention.

Les données utilisées pour rendre compte de ces objectifs spécifiques ont été collectées à travers différentes stratégies qui visent une approche qualitative telle que l'analyse documentaire, des entretiens qualitatifs semi dirigés, des rencontres communautaires, des focus groups, un journal de bord et des observations semi structurées, ainsi qu'une approche quantitative en recourant à des listes de contrôle et à des documents administratifs de l'intervention et des rapports de la plateforme mobile.

En ce qui concerne la sélection des participants, elle est faite en fonction de l'information qu'ils ont à partager (M. Q. Patton, 2002). Donc, nous avons utilisé la méthode d'échantillonnage par choix raisonné, afin d'augmenter l'utilité de l'information et pour bien cerner la compréhension approfondie de la fidélité de la mise en œuvre de l'intervention (M. Q. Patton, 2002). Un échantillon raisonné, basé sur les caractéristiques du groupe cible de l'intervention, les parents d'enfants de moins de 2 ans et les travailleurs de santé qui travaillent dans les villages touchés par l'intervention, indépendamment de la participation, et qui étaient disponibles pour assister à un focus group, a été utilisé. Cependant, tous les membres de la communauté ont été invités à participer pour donner leur avis sur l'intervention. Les membres de la communauté ont été recrutés dans les endroits où se développent les activités communautaires, et par des visites à la maison à travers les intervenants-terrain. Concernant les travailleurs de la santé, tous ceux qui travaillent dans les villages ciblés par l'intervention et qui étaient disponibles pour assister aux entretiens ont été utilisés. Ces participants ont été aussi contactés par les intervenants-terrain. Le recrutement de ces participants a cessé une fois la saturation empirique atteinte (Contandriopoulos et al., 2005; Creswell & Poth, 2016). Enfin, pour les intervenants-terrain, tous ont été invités à participer à travers le superviseur-terrain, une fois la mise en place de l'intervention terminée. Cette invitation est lancée pour faire des rencontres de groupe et des entretiens dans leur bureau de travail.

Au niveau communautaire, plusieurs réunions ont eu lieu dans chaque village pour encourager la participation et pour faciliter la discussion. Pendant la période d'intervention, des *rencontres communautaires* ont été menées pour connaître l'expérience des participants avec la technologie mobile, et pour connaître la perception des activités mises en place et de l'information reçue. Pour cette phase, un outil de collecte de données prédéfini, comprenant des réponses sous forme fermée et un espace pour les commentaires ouverts afin d'enregistrer les commentaires des participants, a été utilisé (Annexe 2). Une fois la période d'intervention terminée, nous avons menés des *focus groups* dans les 15 villages ciblés par l'intervention avec des parents d'enfants de moins de 2 ans (indépendamment de la participation, et qui étaient disponibles pour assister à un focus group) et des membres de la communauté. Pour ces focus groups, nous avons utilisé un guide d'entrevue semi structurée pour guider les rencontres (Annexe 3). Les focus groups visaient à comprendre les perceptions, la compréhension et l'acceptation des activités par les participants. Les *rencontres communautaires* ont eu lieu dans un endroit facilement accessible aux participants

et ont duré environ 60 à 90 minutes. Des périodes de *focus groups* de 45 à 60 minutes ont été abordées dans un endroit facilement accessible aux participants, et ces focus groups ont été formés de façon homogène selon le genre, pour faciliter la participation de la communauté. Ces focus groups ont été animés par les intervenants à l'aide d'une guide qui a été développé et qui a été expliqué avant les rencontres. Toute l'information obtenue a été enregistrée et traduite en anglais.

Pour les travailleurs de la santé, des entretiens semi-dirigés ont été menés afin d'approfondir la perception et le vécu des participants concernant l'intervention, par des questions guidées par toutes les dimensions de notre cadre conceptuel (Annexe 4). Elles ont été d'une durée d'environ 45-60 minutes. Les entretiens ont été menés dans la langue locale pour des intervenants. Ces entretiens ont été enregistrés avec l'autorisation des participants. Ensuite, les entretiens ont été traduits en anglais.

Pour les intervenants-terrain, des entretiens semi-dirigés ont été menés afin d'approfondir la perception et le vécu des participants concernant la mise en place de l'intervention, par des questions ouvertes mais guidées par toutes les dimensions de notre cadre conceptuel (Annexe 5). Elles ont été d'une durée d'environ 45-60 minutes. Les entretiens ont été menés dans la langue locale par une personne qui a été entraînée pour le faire, et qui ne faisait pas partie de l'équipe de recherche. Ces entretiens ont été enregistrés avec l'autorisation des participants. Ensuite, les entretiens ont été traduits en anglais pour une personne externe à l'équipe de recherche. De même, des focus groups ont été abordés avec les intervenants dans un premier temps, pour valider la compréhension du cadre logique de l'intervention, puis, à la fin, pour explorer la perception globale de la mise en place de l'intervention. Donc, une approche participative (entretien + focus groups) a créé un consensus sur la conception théorique de l'intervention, et a permis d'explorer les facteurs qui ont eu une incidence positive ou négative sur la fidélité de la mise en œuvre de l'intervention.

Nous avons mené une analyse documentaire en nous appuyant sur les documents administratifs de l'intervention (le projet, les outils, les registres d'activités rapportés par le suivi de l'intervention, les documents de rétroaction et les documents de travail développés par les concepteurs) pour nous aider à mieux comprendre les effets du contexte sur l'intervention, faciliter l'interprétation et corroborer l'information recueillie par les autres sources de données. Pour des

raisons logistiques et des limitations dans le temps de l'étudiante-chercheuse, des entrevues n'ont pas été menées dans des groupes contrôles. Cette dimension a été évaluée par les groupes des chercheurs de l'étude qui évaluent la faisabilité de l'intervention. Donc, la révision des documents administratifs a été consultée afin d'évaluer si des événements, qui se sont déroulés dans le groupe contrôle, pouvaient influencer l'efficacité de l'intervention Tika Vaani.

Enfin, un journal de bord a été tenu tout au long de la collecte des données par l'étudiante-chercheuse, afin de consigner des données descriptives aidant à mieux comprendre le contexte socioculturel de l'étude, ainsi que de consigner les réflexions et les analyses personnelles. Les deux visites sur le terrain, ainsi que les multiples rencontres avec la chercheuse en charge du projet et le superviseur-terrain, ont permis une meilleure compréhension du contexte et de l'intervention même pour la collecte, l'analyse et l'interprétation des données.

Pour la collecte de données quantitatives, des listes de contrôle concernant les différentes activités proposées dans l'intervention ont été utilisées auprès des intervenants-terrain pour mesurer le niveau de fidélité de l'intervention concernant l'adhésion (couverture, fréquence, contenu et durée). Nous nous sommes basés sur la stratégie proposée par Rebchoock et al (2004) qui a déjà été utilisée et testée par d'autres auteurs (D. Perez et al., 2016; Valéry Ridde, Druetz, Poppy, Kouanda, & Haddad, 2013), afin de construire notre grille pour évaluer la fidélité de la mise en œuvre de l'intervention (Annexe 6). Pour chaque activité planifiée, nous avons évalué si elle a été mise en œuvre comme prévue (I), modifiée (M) ou non mise en œuvre (N) (Rebchoock, Kegeles, Huebner, & Team, 2006).

La vérification de la qualité des grilles pour le focus group, les entretiens et la liste de contrôle ont été faites en deux temps avant leur utilisation formelle. D'abord, la révision a été faite par la directrice et le codirecteur de ce projet de recherche, ensuite, auprès de certains participants ciblés pour l'intervention, afin de vérifier la compréhension des questions et de valider leur formulation.

Pour l'analyse de données, les données qualitatives ont été analysées de façon concomitante. Nous avons utilisé l'analyse de contenu sous l'approche d'analyse de type «

Framework », qui consiste à organiser les données en fonction des thèmes, des concepts et des catégories émergentes (Ritchie & Spencer, 2002). La force de cette approche est qu'en suivant une procédure bien définie, il est possible de retravailler les idées, car le processus analytique a été documenté et donc accessible (Gale, Heath, Cameron, Rashid, & Redwood, 2013; Ritchie & Spencer, 2002). Cette phase a été réalisée en cinq étapes : (i) les transcriptions et la traduction de verbatim des participants ont été faites le plus fidèlement possible en anglais par une personne indépendante de la recherche, et ont été validées et saisies dans le logiciel Word. Une lecture globale du matériel transcrit a été faite afin de se familiariser avec les données; (ii) une codification initiale de ces matériaux a été élaborée par l'identification des thèmes ou des concepts clés; et (iii) les données ont été regroupées, selon notre cadre de référence, afin d'établir une réduction des données. La codification et l'analyse des données ont été exécutées en utilisant le logiciel QDA Miner, et nous avons suivi une double approche : déductive et inductive.

L'approche déductive consistait à identifier, à partir des dimensions de la fidélité de notre cadre conceptuel, des éléments d'intérêts manifestes dans les données recueillies. L'approche inductive consistait à prendre en considération les données pour identifier d'autres informations émergentes, qui pourraient nous aider à comprendre le rôle des facteurs contextuels et modérateurs dans le degré de l'hétérogénéité et de la fidélité de la mise en œuvre de l'intervention, informations qui ne seraient pas perçues dans le modèle retenu de notre cadre conceptuel ; (iv) une matrice a été utilisée pour résumer les données par catégorie de chaque transcription; et, finalement, (v) l'interprétation des données a été faite en identifiant les caractéristiques et les différences entre les données, et en établissant des liens entre les catégories pour mesurer et comprendre la fidélité de la mise en œuvre de l'intervention. En ce qui concerne les données quantitatives, nous avons utilisé l'analyse statistique descriptive en nous appuyant sur les moyennes, les pourcentages et les diagrammes pour représenter les données. Afin de mesurer le niveau de la fidélité, nous avons quantifié la proportion des activités qui ont été proposées par l'ensemble de l'intervention. Chaque activité a été codée par lettres, afin de savoir si elle a été, soit mise en œuvre comme prévue (I), modifiée (M) ou non mise en œuvre (N). Ensuite, nous avons calculé le pourcentage d'accord entre ce qui a été planifié et ce qui a été mis en œuvre.

Finalement, lors de l'intégration de résultats, nous avons triangulé les données qualitatives et quantitatives de façon complémentaire. Bien que la majorité des sources de données soient de nature qualitative, les deux types de données avaient la même priorité dans cette étude. Les résultats ont été présentés avec une synthèse narrative en suivant le cadre de référence utilisé, et en nous appuyant sur les données quantitatives, dans le but de mieux comprendre le degré de fidélité de l'intervention. Nous nous sommes inspirés de la liste de contrôle GRAMMS (O'cathain, Murphy, & Nicholl, 2008), laquelle est un guide pour améliorer le bon rapport des études sur les méthodes mixtes, pour fournir des informations sur l'évaluation de la fidélité de la mise en œuvre de l'intervention Tika Vaani, à travers une conception de triangulation concurrente à méthodes mixtes (Annexe 7). Le tableau 2 synthétise pour chaque volet de cette recherche évaluative, le devis, les objectifs, les outils et les sources de données, ainsi que les participants dans ce processus évaluatif.

Tableau 2. – Synthèse sur les données collectées dans le cadre de la thèse doctorale

Volet de recherche	Devis de recherche	Objectif de l'évaluation	Outils pour la collecte, l'analyse et le rapport de données	Sources de données à l'étude		Participants
Première volet	Revue systématique	Étudier comment l'évaluation de la fidélité des interventions en santé publique, utilisant des essais randomisés en grappes (ERG) dans les pays du Sud, a été rapportée	-Bases de données : MEDLINE, EMBASE, CINAHL, PubMed et EMB -Guide Cochrane et sources supplémentaires -Logiciel de gestion bibliographique EndNote et logiciel Excel -Cadre conceptuel pour évaluer la fidélité -Guide PRISMA pour le rapport de l'étude	-Articles publiés évalués par des pairs		n=65 articles
				-Registre du protocole de l'article évalué		n=65 registres
				Documents annexés	Protocoles publiés évalués par des pairs	n=21 protocoles
					Protocoles dans la littérature grise	n=15 doc. PDF
Rapports complémentaires	n=19 doc. PDF					
Deuxième Volet	Étude de cas unique à trois unités	Comprendre la conception de l'intervention à évaluer afin	Analyses documentaires et guide mERA pour le rapport de l'étude	-Documents préparés avant la mise en œuvre : PDFs sur l'évaluation formative et la structure de la plateforme mobile		n=48 doc. PDF
				Rétroaction structurée	-Révision de la littérature	-Concepteurs de l'intervention

	d'analyse imbriquées selon une approche mixte concomitante	d'analyser leur plausibilité	Focus groups	Intervenants-terrain	n=7
		* Mesurer le degré de fidélité de l'intervention Tika Vaani * Comprendre le rôle des facteurs modérateurs dans le degré d'hétérogénéité et de la fidélité de la mise en œuvre de l'intervention Tika Vaani	Observation structurée	-Intervenants-terrain	n=7
			Entretiens semi-structurés	Intervenants-terrain	n=7
				Travailleurs de santé	n=17
			Focus groups	Focus groups avec les intervenants-terrain	n=7
				25 focus groups avec la communauté	n=290
			Journal de bord	La chercheuse évaluatrice	n=1
			Analyses documentaires	Registres des enquêtes de villages ciblés par l'intervention	n=13 villages
				-Registres des rencontres communautaires :	n=1698 (GR) participants
				Grande rencontre (GR) Petite rencontre (PR) -Registres de la plateforme mobile	n=1428 (PR) participants
Guide GRAMMS pour le rapport de l'étude Cadre conceptuel pour évaluer la fidélité Logiciel stata, Excel et QDA Miner					

2.3 Rôle de l'étudiante

Dans le cadre de cette thèse doctorale, le sujet de thèse « évaluer la fidélité des interventions en santé publique dans le cadre des essais randomisés en grappes » a été suggéré par la directrice de recherche.

L'étudiante-chercheuse a été invitée à participer au sein de l'équipe de recherche de la directrice, ce qui a favorisé un environnement optimal en termes d'orientation scientifique, d'expertise professionnelle et du terrain, ainsi que des ressources humaines, matérielles et économiques pour mener l'ensemble de cette recherche doctorale. Donc, avec l'idée d'évaluer la fidélité de la mise en œuvre des interventions en santé publique, l'étudiante-chercheuse a pris en considération l'initiative en se documentant en profondeur sur le sujet. À partir de cette suggestion et de la révision de la littérature, le premier volet de la thèse a surgi et a fait l'objet de deux articles publiés : le protocole et les résultats de la révision systématique (articles 1A et 1B). Le rôle de l'étudiante-chercheuse pour ce premier volet a été décrit dans les articles publiés.

En ce qui concerne le deuxième volet de recherche, l'étude de cas, il s'inscrit dans le projet de recherche de la directrice de recherche. Au sein de l'équipe de recherche, l'étudiante-chercheuse a été invitée dans un premier temps à participer pour contribuer au développement du contenu de l'intervention. Elle a participé dans plusieurs rencontres avec l'équipe de recherche, ce qui lui a permis de mieux comprendre la théorie de l'intervention, le contexte et le besoin de la mise en place de l'intervention. La participation de l'étudiante-chercheuse et son interaction avec l'équipe de recherche dans cette phase a facilité la compréhension de la théorie et du développement de l'intervention pour l'élaboration d'un des articles (Article 2), et a permis de mieux envisager, dans un deuxième temps, la phase évaluative concernant la fidélité de la mise en œuvre de l'intervention qui est rapportée dans un des articles (Article 3).

L'évaluation de la fidélité de la mise en œuvre de l'intervention Tika Vaani découle d'un besoin de concepteurs de l'intervention et constitue le mandat précis pour l'étudiante-chercheuse pour mener à bien l'évaluation. L'intervention évaluée à travers un essai randomisé en grappes, mise en place dans sa phase pilote, fait l'objet d'une évaluation formative pour savoir d'abord, si

elle a été implémentée telle que conçue, et, ensuite, pour faire des recommandations si nécessaire avant sa mise en grande échelle. Pour répondre à ce besoin, l'étudiante-chercheuse a adopté une position collaborative dans son rôle d'évaluatrice, mais, tout de même, indépendante pour assurer l'objectivité dans le processus évaluatif.

En effet, dans la pratique évaluative, la position de l'évaluateur découle du point de vue de l'objet à évaluer, ainsi que des attributs propres à l'évaluateur (Brousselle et al., 2011) Donc, le rôle de l'étudiante-chercheuse fait partie fondamentale de la pratique évaluative et doit être expliqué (Brousselle et al., 2011). Par conséquent, le rôle a été précisé au départ pour faciliter la compréhension des enjeux concernant la prise de décision au sein de l'équipe de recherche. La participation au sein de l'équipe a permis de mieux connaître le milieu, savoir à qui s'adresser pour mettre sur pied le projet d'évaluation et mieux adapter les outils pour la collecte de données. Toutefois, l'étudiante-chercheuse a gardé une certaine distance auprès de l'équipe pour s'assurer d'accomplir le mandat avec une vision plus large, pour ne pas y travailler avec des idées préconçues et afin de mieux comprendre la réalité à évaluer. Un regard critique et réfléchi sur la pratique évaluative a été tenu tout au long de la recherche, en articulant les connaissances acquises et les compétences pour faire face à des obstacles durant ce processus.

Bien que la directrice à charge du projet ait supervisé l'étudiante-chercheuse pendant tout le processus, cette dernière a été responsable de développer tout le projet évaluatif concernant la conceptualisation, la méthodologie, les instruments pour la collecte des données, l'analyse des données et l'interprétation des résultats concernant l'évaluation de la fidélité du projet. Donc, le rôle des parties prenantes a été un rôle de collaboration et d'orientation pour faciliter l'information, et un moindre rôle dans la construction technique de l'évaluation (Brousselle et al., 2011). Enfin, l'étudiante-chercheuse a tenu compte des différentes perspectives des participants pour augmenter l'objectivité et la représentativité du jugement fait. Ainsi, la validité de l'évaluation a été consolidée par les différentes sources de données, par la triangulation de l'information, par la position de la chercheuse-évaluatrice avec un regard collaboratif et réfléchi, mais indépendant de l'objet évalué et des parties prenantes. Le fait de participer dans la compréhension du développement de l'intervention, et non pas dans leur implémentation, a facilité également l'interprétation des résultats avec un regard plus critique, contextualisé et objectif de l'évaluation faite.

2.4 Considérations éthiques

Le premier volet de cette thèse concernant la revue systématique n'a pas fait l'objet d'évaluation éthique, car elle ne vise pas des participants humains, mais plutôt des écrits scientifiques. Pour le deuxième volet de cette thèse concernant l'étude de cas, l'évaluation de la fidélité de la mise en œuvre de l'intervention s'inscrit au projet principal de ma directrice de recherche Dre Mira Johri, qui a déjà reçu l'approbation par le Comité d'éthique du CHUM et par le comité d'éthique en Inde (Annexe 8). Pour le projet principal, la démarche de consentement pour les participants de la communauté ciblée par l'intervention a été menée à deux niveaux. D'abord la permission a été obtenue par la communauté et ensuite par les personnes interrogées. Spécifiquement, le processus pour obtenir des permis a été menée comme suit: (a) au niveau de l'État -gouvernement à travers le programme de santé rurale et le programme de vaccination; (b) au niveau du District par le magistrat de district et son équipe; (c) au niveau du village par le représentant élu qui est le maire du village appelé le « sarpanch » et s'il n'était pas disponible le processus était fait par un autre membre élu du maire appelé le « panchayat » et finalement (d) au niveau individuel ou de la famille, spécifiquement par les personnes interrogées.

Des enjeux liés au niveau du projet et de la pratique sur le terrain pourraient surgir et seraient à considérer (Guillemin & Gillam, 2004; Olivier de Sardan, 1995). Des pistes de solutions ont été envisagées. Concernant le projet: (i) un formulaire de consentement éclairé, rédigé dans la langue locale, a été adressé aux différents participants de la recherche. Le formulaire spécifie que la participation de chacun est volontaire, et que chaque participant reste libre de refuser en tout temps de répondre à une question ou de se retirer sans justification. Particulièrement, pour les intervenants-terrain, l'objectif de l'évaluation a été précisé, ainsi que la position de l'étudiante-chercheuse qui agissait à titre d'étudiante du doctorat, et non à titre de responsable de l'intervention, pour éviter un effet de désirabilité sociale (Althubaiti, 2016). De même, cette précision a été faite pour réduire la perception de conflit d'intérêts, qui pourrait émerger en fonction du statut de membre intégrant de l'équipe de recherche. Également, lors de l'entretien, ce groupe de participants a été informé que ce qui était recherché n'était pas la bonne réponse théorique, mais plutôt une compréhension de ce qu'ils ont vécu et perçu de l'intervention mise en place; (ii) il n'y a pas eu de mesures incitatives pour les participants; (iii) les noms du personnel acceptant de

participer à l'étude restent confidentiels. Afin de préserver l'anonymat de chaque participant, un code a été attribué à chaque entrevue et les noms de tous les participants ont été supprimés avant la retranscription des données, et n'ont pas été mentionnés dans tout autre texte ou communication scientifique; (iv) les données sont conservées sous clé pour une période de sept ans, dans des fichiers électroniques à accès protégé (sur le serveur du CRCHUM) et dans un classeur verrouillé (au CRCHUM). À la fin de ce délai, les documents comportant des renseignements personnels seront détruits.

Des pistes de solutions envisagées pour la pratique sur le terrain ; (v) une position réfléchie et de savoir-faire a été privilégiée, pour éviter des biais dans la production des données dans ce contexte d'aide internationale et d'évaluation, qui exige un regard analytique et critique sur les participants et sur moi, ainsi que sur mon rôle dans le cadre de la recherche mentionnée auparavant. À cet effet, le journal de bord a été un outil clé pour documenter le processus d'apprentissage, pour favoriser la réflexion personnelle sur la pratique, et pour retracer les sentiments survenus lors de moments marquants tout au long de cette recherche.

Chapitre 3 – Résultats

Ce chapitre de la thèse présente les résultats de la recherche doctorale. Il est constitué de trois articles:

Article 1B Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review

Article 2 Development of an mHealth Behavior Change Communication Strategy: A case-study from rural Uttar Pradesh in India

Article 3 Implementation fidelity and acceptability of an intervention to improve vaccination uptake and child health in rural India: A mixed methods evaluation of a pilot cluster randomized controlled trial

3.1 Article 1B Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low-and middle-income countries: systematic review

Myriam Cielo Pérez^{1,2} ; Nanor Minoyan^{1,2}; Valéry Ridde^{3,4}, Marie-Pierre Sylvestre^{1,2}; Mira Johri^{1,5}

STATUT : ARTICLE PUBLIÉ

Perez, M. C., Minoyan, N., Ridde, V., Sylvestre, M. P., & Johri, M. (2018). Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low-and middle-income countries: systematic review. *Trials*, 19(1), 410.

Author affiliations

¹ Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM),
900, rue Saint-Denis, Pavillon R Montréal (Québec) H2X 0A9, Canada

² Département de médecine sociale et préventive, École de santé publique (ESPUM), Université de Montréal, 7101, avenue du Parc, 3e étage Montréal (Québec) H3N 1X9, Canada

³ Institut de Recherche en Santé Publique Université de Montréal (IRSPUM), Pavillon 7101 Avenue du Parc, P.O. Box 6128, Centre-ville Station, Montréal, Québec, H3C 3J7, Canada

⁴ Institut de recherche pour le développement (IRD). Le Sextant 44, bd de Dunkerque, CS 90009 13572 Marseille cedex 02, France.

⁵ Département de gestion, d'évaluation, et de politique de santé, École de santé publique, Université de Montréal, 7101, avenue du Parc, 3e étage Montréal (Québec) H3N 1X9, Canada

Contribution des auteurs :

À partir d'une idée originale de Dre Mira Johri, et sous sa supervision, Myriam Cielo Pérez a développé la stratégie de recherche documentaire en collaboration étroite avec une spécialiste de l'information scientifique (bibliothécaire), a élaboré les outils de collecte de données, la sélection des articles, l'abstraction et l'analyse des données, l'interprétation des résultats et la rédaction de l'article. Nanor Minoyan a participé dans l'abstraction et l'analyse des données, et avec Dr Valéry Ridde et Dre Marie-Pierre Sylvestre, ont contribué à l'interprétation des résultats, et ont apporté un regard critique et constructif à l'article.

Dre Mira Johri a conçu l'étude, supervisé l'ensemble des étapes et a contribué à la sélection des articles, l'interprétation des résultats et a apporté un regard critique et constructif aux outils d'analyse, aux conclusions des analyses et à l'ensemble de l'article.

Tous les auteurs ont lu et approuvé le manuscrit final pour soumission.

ABSTRACT

Background: Cluster randomised trials (CRTs) are a key instrument to evaluate public health interventions. Fidelity assessment examines study processes to gauge whether an intervention was delivered as initially planned. Evaluation of Implementation Fidelity (IF) is required to establish whether the measured effects of a trial are due to the intervention itself and may be particularly important for CRTs of complex interventions conducted in low- and middle-income countries (LMICs). However, current CRT reporting guidelines offer no guidance on IF assessment. The objective of this review was to study current practices concerning the assessment of IF in CRTs of public health interventions in LMICs.

Methods: CRTs of public health interventions in LMICs that planned or reported IF assessment in either the trial protocol or the main trial report were included. The MEDLINE/PubMed, CINAHL, and EMBASE databases were queried from January 2012 to May 2016. To ensure availability of a study protocol, CRTs reporting a registration number in the abstract were included. Relevant data were extracted from each study protocol and trial report by two researchers using a predefined screening sheet. Risk of bias for individual studies was assessed.

Results: We identified 90 CRTs of public health interventions in LMICs with a study protocol in a publicly available trial registry published from January 2012 to May 2016. Among these 90 studies, 25 (28%) did not plan or report assessing IF; the remaining 65 studies (72%) addressed at least one IF dimension. IF assessment was planned in 40% (36/90) of trial protocols and reported in 71.1% (64/90) of trial reports. The proportion of overall agreement between the trial protocol and trial report concerning occurrence of IF assessment was 66.7% (60/90). Most studies had low to moderate risk of bias.

Conclusions: Fidelity assessment is not currently a systematic practice in CRTs of public health interventions carried out in LMICs. In the absence of fidelity assessment, it may be difficult to determine if CRT results are due to the intervention design, to its implementation, or to unknown

or external factors that may influence results. CRT reporting guidelines should promote IF assessment.

Registration: Protocol published and available at: [10.1186/s13643-016-0351-0](https://doi.org/10.1186/s13643-016-0351-0)

Keywords: Cluster randomised trials, Implementation fidelity, Public health interventions, Process evaluation, Developing countries, Systematic review

BACKGROUND

Cluster randomised trials (CRTs) have become a key instrument to evaluate public health interventions (Bland, 2004; M.J. Campbell, A. Donner, & N. Klar, 2007; M. K. Campbell, Elbourne, Altman, & group, 2004; M.K. Campbell, Mollison, Steen, Grimshaw, & Eccles, 2000; Isaakidis & Ioannidis, 2003; Moberg & Kramer, 2015; Osrin et al., 2009), particularly in low- and middle-income countries (LMICs) (Handlos, Chakraborty, & Sen, 2009; Isaakidis & Ioannidis, 2003). Randomised controlled trials (RCTs) are widely considered to provide the highest quality of evidence on the effectiveness of health interventions (M. Campbell et al., 2000; Craig et al., 2012; Evidence-Based, Medicine Working, & Group., 1992; Richards & Hallberg, 2015), and CRTs are a form of randomised trial in which clusters of individuals (such as families, villages, hospital services, or schools) rather than independent individuals are randomly allocated to intervention or control groups (M. K. Campbell et al., 2004). Increasingly, public health researchers recognize the importance of developing health interventions that target populations rather than individuals and can address the wide range of social and environmental factors influencing health (Sanson-Fisher, D'Este, Carey, Noble, & Paul, 2014; Upshur, 2002). CRTs offer an appropriate design to assess such public health interventions as well as to measure the overall effect of an intervention at the population level (Handlos et al., 2009; R. J. Hayes, Alexander, Bennett, & Cousens, 2000; Isaakidis & Ioannidis, 2003; Moberg & Kramer, 2015; Sanson-Fisher et al., 2014) heterogeneity of impact among population subgroups, and equity (Morris, Ranson, Sinha, & Mills, 2007; Ranson, Sinha, Morris, & Mills, 2006).

Implementation fidelity in CRTs of public health interventions

Recent methodologically oriented systematic reviews have found evidence of improvements in the design and analysis of CRTs, while also noting deficiencies in trial implementation (Bonell, Oakley, Hargreaves, Strange, & Rees, 2006; S. Eldridge, Ashby, Bennett, Wakelin, & Feder, 2008). Previous systematic reviews have emphasised the potential importance of process evaluation to mitigate these methodological problems, which can affect the internal and external validity of trial results (Brierley, Brabyn, Torgerson, & Watson, 2012; S. Eldridge et al., 2008; Isaakidis & Ioannidis, 2003; Oakley, Strange, Bonell, Allen, & Stephenson, 2006; Puffer,

Torgerson, & Watson, 2003).

“Implementation fidelity” refers to the degree to which an intervention is delivered as initially planned (Carroll et al., 2007). Fidelity assessment is the practice of evaluating implementation fidelity. Considered an aspect of process evaluation, fidelity assessment aims to understand and measure to what extent the intervention is being implemented as intended and identify specific reasons for the success or failure of the intervention (Carroll et al., 2007; Medical Research & Council., 2008; Richards & Hallberg, 2015). The benefits of evaluating implementation fidelity include increased confidence in scientific findings, increased power to control for confounding factors and increased ability to evaluate the performance of theory-based interventions (Borrelli, 2011). Fidelity assessment can improve internal and external validity of CRTs (S. Eldridge et al., 2008) by providing evidence that trial results are due to interventions themselves rather than to external factors, and facilitating generalization of results to contexts that may differ substantially from the original trial setting (Carroll et al., 2007; Richards & Hallberg, 2015). Fidelity assessment may be particularly important for public health interventions, as these interventions tend to be complex and constituted by multiple components (M. Campbell et al., 2000; Hawe, 2015) that act independently or interdependently (Moore et al., 2015), allowing greater potential for variation during implementation (Carroll et al., 2007).

Framework for the evaluation of implementation fidelity

Table 1 outlines the conceptual framework used in this review, based on the work of Carroll et al. (2007) (Carroll et al., 2007) as refined by Hasson (Hasson, 2010). We selected this framework to guide the review because it provides a comprehensive synthesis of previous work on implementation fidelity and has been widely influential.

Table 1. Conceptual framework for implementation fidelity used in this review*

Fidelity Components	
Content	Defined as an attempt to establish the “active ingredients” of the intervention.
Coverage	Refers to the degree to which all persons who met study inclusion criteria received the intervention
Frequency	Refers to whether the intervention was delivered with the regularity or frequency planned by its designers.
Duration	Establishes whether the intervention was delivered with the duration planned by its designers.
Moderating factors	
Comprehensiveness of intervention description	Factors such as the degree of intervention complexity, and whether the intervention description is complete or incomplete, vague or clear, may influence the degree of implementation fidelity
Strategies to facilitate implementation	Several support strategies may be used to optimise and to standardise implementation fidelity
Quality of delivery	Concerns whether an intervention is delivered in a way that increases the likelihood of achieving the desired health outcomes
Participant responsiveness	Refers whether researchers established strategies to increase acceptance by and acceptability to those receiving an intervention.
Recruitment ^a	Refers to procedures used to attract potential program participants.
Context ^a	Refers to surrounding social systems, such as structures and cultures of organizations and groups, and historical and concurrent activities and events, that may influence study activities

*Adapted from Carroll et al. 2007. ^a These components were added by Hasson, 2010.

Fidelity assessment in CRT reporting guidelines

The Consolidated Standards of Reporting Trials (CONSORT) group was created to provide guidance to improve the quality and transparency of RCT reporting (Moher et al., 2010) including CRTs (M. K. Campbell et al., 2004; M. K. Campbell, Piaggio, Elbourne, Altman, & Group, 2012). The CONSORT statement recognises that the trial protocol for a given study may not have been followed fully by some trial participants for a wide variety of reasons, including failure to receive the entire intervention as planned (Moher et al., 2010). Cases of protocol non-adherence may influence interpretation and credibility of the results and thus the validity of the conclusions (Borrelli, 2011; S. Eldridge et al., 2008; Schneeweiss, 2006; Thabane et al., 2013). To preserve the ability to make strong inferences about intervention effects, CONSORT recommends that all participants randomised be retained in the analysis and analysed according to their original assigned groups, an approach known as “intention-to-treat” or “ITT” analysis.

Reasons for protocol non-adherence in individually randomised RCTs may differ from those in CRTs. In a clinical trial setting, non-adherence depends largely on the actions of the trial participant and the treatment provider, which may in turn be related to issues such as treatment side effects and safety. In CRTs of public health interventions, protocol non-adherence may occur because complex interventions that include multiple components are delivered with poor fidelity. However, both for individually randomized parallel group trials (Moher et al., 2010) and CRTs (M. K. Campbell et al., 2004; M. K. Campbell et al., 2012) current CONSORT guidelines offer no advice on methods to assess protocol non-adherence.

Rationale for undertaking this review

As earlier methodologically oriented systematic reviews have demonstrated, CRTs of complex public health interventions may be particularly at risk of experiencing protocol deviations and non-adherence, and these may compromise the validity of their findings (Bonell et al., 2006; S. Eldridge et al., 2008). Although process evaluation techniques such as evaluation of implementation fidelity can help to identify these problems and mitigate their negative effects, current reporting guidelines

for CRTs offer no specific guidance on intervention fidelity assessment. Wide divergence in current practices is therefore likely. We undertook a methodologically oriented systematic review of current practices related to assessment of intervention fidelity within CRTs of public health interventions in LMICs, so as to inform best practices for CRTs. To our knowledge, no other systematic review addresses this question. Although CRTs have been widely implemented to evaluate public health interventions in both high-income countries and LMIC contexts, interventions, approaches, and outcomes may differ substantially between settings. Given the limitations of the resources, these challenges may be even more important in CRTs of public health interventions in LMICs. We therefore limit our focus to LMICs.

Objective

We conducted a methodologically oriented systematic review to address the following research questions:

5. What proportion of recent CRTs of public health interventions in LMICs planned to assess implementation fidelity (IF)?
6. What proportion of recent CRTs of public health interventions in LMICs reported assessing IF?
7. Which fidelity components were examined, and which data collection methods were employed to assess each component?
8. Is there evidence of divergent practices between planned and reported studies?

METHODS

We report results according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher D, Liberati, Tetzlaff, Altman, & Group, 2009) ; the PRIMSA checklist is available in an online supplement (Additional file 1). For a detailed description of the methods, see the published protocol (M. C. Perez, Minoyan, Ridde, Sylvestre, & Johri, 2016) available at: [10.1186/s13643-016-0351-0](https://doi.org/10.1186/s13643-016-0351-0). Some technical modifications were made while implementing the review. We report the date and the rationale for each change in italics.

Study eligibility

Study Designs:

We included all cluster-randomised trials (CRTs). For the purposes of this review, a CRT is defined as a trial in which “intact social units or clusters of individuals, rather than independent individuals, are randomly allocated to intervention groups” (M. K. Campbell et al., 2012). CRTs may include trials employing parallel group, stepped wedge, factorial, adaptive or pragmatic designs, among others.

Participants:

Study participants were humans living in LMICs according to the World Bank country classifications (The World & Bank.).

Interventions:

This review focuses on “public health interventions”. Following Rychetnik and colleagues, we define a public health intervention as a disease prevention or health promotion intervention applied to many, most, or all members in a community, which aims to deliver a net benefit to the community or population, as well as benefits to individuals (Rychetnik, Frommer, Hawe, & Shiell, 2002; Rychetnik, Hawe, Waters, Barratt, & Frommer, 2004). Assessment of intervention fidelity may be especially important for public health interventions, which are inherently “complex, programmatic, and context dependent” (Rychetnik et al., 2002) In order to operationalize this definition and guide study inclusion decisions, we used the Minnesota Department of Health “Intervention Wheel” (Keller, Strohschein, Lia-Hoagberg, & Schaffer, 2004).

Comparators:

Comparators were defined as per the original CRT.

Outcomes:

We included CRTs that assessed IF in the trial protocol, the main trial report, or both. We also included CRTs reporting assessment of IF in a complementary document such as a published article, an online appendix to the main paper, or a grey literature report, in lieu of reporting assessment of IF in the main trial report only. We considered studies as evaluating IF if they proposed or implemented methods to assess at least one of the four key fidelity components: (1) content, (2) coverage, (3) frequency, and (4) duration.

Report characteristics**Availability of the study protocol:**

To ensure availability of a study protocol, we included CRTs reporting a registration number in the abstract for any registry meeting WHO trial registration data set (TRDS) criteria (World Health & Organization., 2016). The TRDS was used in this review to evaluate planned assessment of intervention fidelity, either alone, or in conjunction with a published study protocol.

Exclusion criteria:

We excluded studies that (i) were not cluster randomized trials, (ii) did not plan or report assessment of IF (iii) were not public health interventions, (iv) were conducted in high income countries (The World & Bank.) (v) were published before 2012; (vi) did not have a publicly available protocol, or (vii) for which only the protocol but not the main trial report has been published. Manuscripts were also excluded if they were (viii) pilot studies, (ix) secondary reports of a main study for which the relevant findings were published prior to 2012, date of the last update to the CONSORT Statement to improve reporting of CRTs, (x) studies published in a language other than English, Spanish or French, or (xi) CRT trials published only in the grey literature.

Information sources and search strategy

Literature search strategies were developed in collaboration with an academic librarian. The search strategy was limited to the period from January 2012 to May 2016 and, due to resource constraints, *was not updated towards the end of the review (12-07-2017)*. Search strategies used Medical Subject Headings (MESH) and text words related to cluster randomised trials, developing countries, and public health interventions. The electronic database search was developed first for MEDLINE (ovid) and then adapted for the following electronic databases: EMBASE (ovid), CINAHL (ovid), PUBMED and EMB Reviews (ovid). The full search strategy is available in an online supplement (Additional file 2). Searches were filtered to articles concerning humans, written in English, French, or Spanish. We added relevant studies suggested by the review team. Identified records were uploaded into EndNote software and duplicates were eliminated.

Study screening and selection

Study screening and selection were done manually within EndNote. To ensure the availability of study protocols, we limited the search to CRTs that have the word “regist*” in the abstract and used these results to begin the process of screening and selection. We validated this procedure by examining a subset of excluded articles. Screening and selection were done in two stages by two independent reviewers (MCP and NM). In the first stage, reviewers independently screened the titles and abstracts of each identified reference against the inclusion criteria to eliminate irrelevant publications. In the second stage, each reviewer independently screened the full text of all studies that appeared to meet the inclusion criteria, or for which eligibility was uncertain. For each study, reviewers identified additional articles of potential relevance, by reviewing references from the main trial report, consulting the trial registry record, and searching the PubMed database for new publications by the lead trial author or using trial identifiers. To aid in article screening and selection, the team developed and tested a full text screening sheet. Any disagreements between reviewers were resolved through discussion and, as necessary, through arbitration by a third author (MJ). A summary table describing studies excluded with reasons is available in an online supplement (Additional file 3).

Data extraction and data items

According to our protocol, two review authors would extract data independently. Given the volume and complexity of materials to review, it was not practical for both reviewers to abstract data independently. We adopted the following procedure (20-09-2016): Included studies were randomly assigned to each of the two reviewers. Relevant data were extracted from each study protocol and trial report by one reviewer (MCP or NM) and crosschecked by the other reviewer (MCP or NM). To ensure high quality data extraction and to reach consensus, frequent meetings were held by the two reviewers (MCP and NM) and included a third author (MJ) when necessary.

Outcomes and prioritization

The search and selection process for this review was designed to identify two quantities required for calculation of outcomes based on proportions: (1) Numerator: These are studies that meet all inclusion and exclusion criteria. As for all systematic reviews, these studies were our principal focus and were included in the review. (2) Denominator: This is the total of studies for this review (N), which we defined as all studies that satisfy all inclusion and exclusion criteria, with the exception of the outcome criterion (planned or report IF assessment).

Primary Outcome

The primary outcome for this study was the proportion of overall agreement between the protocol and trial report concerning occurrence of IF assessment. This corresponds to research question 4. Data were summarised in a two-by-two table comparing assessment of intervention fidelity in the trial report to that in the protocol. “N” represents the set of recently published CRTs of public health interventions in LMICs that have registered the study protocol in a publicly available trial registry. For each CRT in N, we determined whether IF was assessed in the registered (or published) protocol or in the trial report (or associated documents). Studies judged to have assessed IF were coded as “1”; others were coded as “0”. The proportion of overall agreement is defined as the proportion of eligible CRTs for which the protocol and the trial report agree in the

practice of implementation fidelity assessment (i.e. both positive, or both negative). It was computed as $(a+d) / N$.

	Protocol		
	+	-	
Trial Report	+ a	b	a + b
	- c	d	c + d
	a + c	b + d	N

Secondary Outcomes

To address research questions 1, 2, and 3, we also calculated:

- The frequency and proportion of trial *protocols* reporting assessment of intervention fidelity, out of N
- The frequency and proportion of trial *reports* reporting assessment of intervention fidelity, out of N
- The proportion of positive agreement among those that agree, computed as $a/(a+d)$
- Frequency counts and percentages summarising fidelity components examined and data collection methods proposed or employed

Risk of bias in individual studies

We used the Cochrane Collaboration tool to assess the risk of bias in individually randomised trials (Higgins et al., 2011) and included additional criteria relevant to assessing risk of bias in CRTs (Donner & Klar, 2000; R. Hayes & Moulton, 2009; Higgins et al., 2011) which are available in an online supplement (Additional file 4). We assessed each criterion as low, high, or uncertain risk, and provided a sample text that illustrates the reasons for this judgment. The initial evaluation was done by the primary reviewer (MCP or NM) and verified by the other

reviewer (MCP or NM). In case of doubt a third reviewer (MJ) was consulted. To investigate whether study quality influences IF practices, we stratified the sample to create categories of low-risk studies (7 or more of the 10 criteria were evaluated as low risk, n=36 studies), versus uncertain or high-risk studies (6 or fewer of the 10 criteria were evaluated as low risk, n=29 studies). Cut-points were chosen empirically to create groups of roughly equal size.

Data synthesis

A narrative synthesis was used to explore relationships and findings within and between the included studies. As our review focuses on methodological issues, we did not assess potential publication bias.

RESULTS

Study selection

The electronic search identified 6876 publications, of which 226 potentially relevant studies were assessed in full text. Stage 1 of the full text review was used to define the study denominator: of the 226 articles screened, 90 were CRTs of public health interventions in LMICs with a study protocol in a publicly available trial registry published from January 2012 to May 2016. In Stage 2 of the full text review, we reviewed these 90 studies to identify those that in addition considered IF in either the protocol, the trial report, or both. We excluded 25 studies (28%) that did not plan or report assessing IF. The remaining 65 studies addressed at least one IF dimension. Figure 1 shows the PRISMA flow chart.

Study characteristics

The 65 studies were published in English between January 2012 and May 2016. The included studies are very diverse in terms of sample size and target population, ranging from 190 to more than 280,000 eligible participants per study. Table 2 presents general characteristics of the

studies included in the review. Characteristics of the individual studies included are summarised in an online supplement (Additional file 5). Interventions reviewed addressed diverse public health aims; a detailed classification is available in an online supplement (Additional file 6).

Outcome:

We identified 90 CRTs of public health interventions in LMICs with a study protocol in a publicly available trial registry published from January 2012 to May 2016. IF assessment was planned in 40% (36/90) of trial protocols and reported in 71.1% (64/90) of trial reports. Notably, 27.8% (25/90) of studies did not assess any IF dimension in either the trial or protocol. Practices concerning IF assessment diverged between trial protocols and reports (Table 3). The proportion of overall agreement between the trial protocol and trial report concerning occurrence of IF assessment was 66.7% (60/90). The proportion of positive agreement among those that agree was 58.3% (35/60). We were unable to compute outcomes 4b, 4c, and 4d as specified in the original protocol (M. C. Perez et al., 2016).

Practices of intervention fidelity assessment in protocols and trial reports

Of 65 studies assessing IF, 71% (46/65) described fidelity assessment only in the main trial report, 20% (13/65) of studies published a process evaluation, and 9% (6/65) of studies detailed IF in a complementary document. We found that fidelity components and moderating factors were not systematically assessed in CRTs (Table 4). In addition, only 28% (18/65) of studies monitored events taking place in control groups.

Fig 1 Flow diagram summarising study selection process

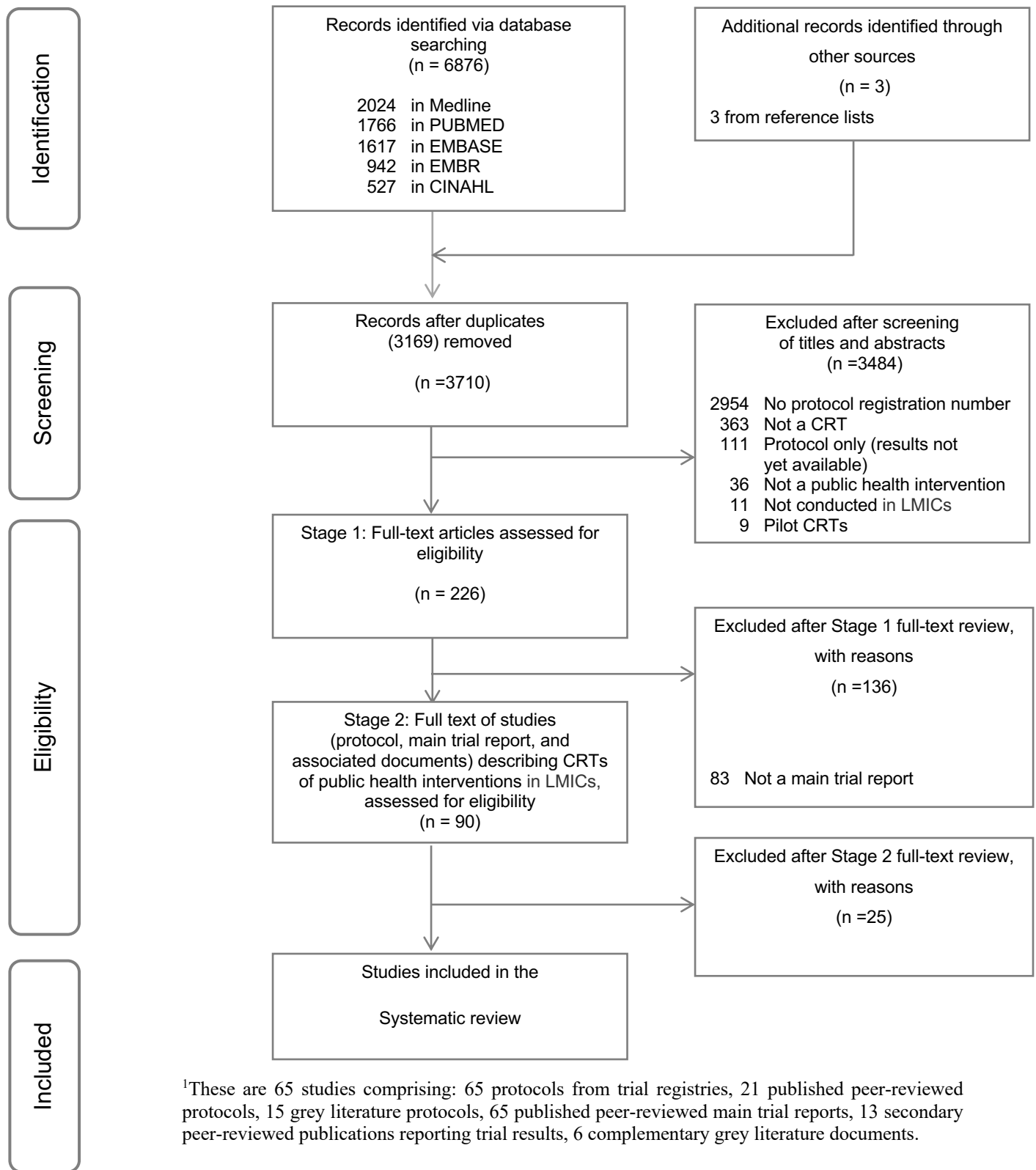


Table 2 General characteristics of included studies (n=65)

Characteristics	(n)	(%)
Study designs		
2-arm parallel CRT	49	75
2 × 2 factorial design CRTs	9	14
Three-arm parallel CRTs	5	8
Stepped wedge designs	2	3
Region		
Sub-Saharan Africa	36	55
South Asia	14	22
Latin America & Caribbean	6	9
East Asia & Pacific	6	9
Multiple regions	3	5
Intervention category		
HIV, tuberculosis, malaria, dengue	24	37
Child health conditions	19	29
Maternal and newborn health	14	22
Non-communicable diseases	8	12
Comparator (control group)		
Receipt of usual services or care	47	72
Applied a simplified version of the intervention	15	23
Used placebo	3	5
Intervention Fidelity assessment		
Protocol		
4 dimensions	2	3
3 dimensions	5	8
2 dimensions	11	17
1 dimension	18	28
Not done	29	44
Trial		
4 dimensions	5	8
3 dimensions	18	28
2 dimensions	21	32
1 dimension	20	31
Not done	1	1

Table 3 Intervention fidelity assessment among CRTs of public health interventions in LMICs (n=90)

	Trial Report		Total	
	Positive	Negative		
Trial Protocol	Positive	35 (a)	1 (c)	36 a + c
	Negative	29 (b)	25 (d)	54 b + d
		64 a + b	26 c + d	90 N (a+ b+ c+ d)

Table 4 Fidelity components and moderating factors identified* (n=65)

Trial Protocol			
Fidelity components	Studies		Study ID
	(n)	(%)	
Content	34	52,3	43-76
Coverage	13	20	43,44,50,53,57,62,64,69,73-75,77,78
Frequency	12	18,5	43-45,49, 50, 52,57-59,62, 70,73
Duration	4	6,2	50,63, 70,73
Trial Report			
Fidelity components	Studies		Study ID
	(n)	(%)	
Content	60	92,3	43-60,62-69,71, 73-105
Coverage	34	52,3	43-48, 50,52,53, 57-59, 62-64, 66,68, 71-75,77,80,84, 85,90,93,94,97,98,101,103,105

Frequency	33	50,8	43-45,48-50,52,57-59,62,63,65,68-70,72-74,78,80-82, 84,85,90,94,99,100,101,104-106
Duration	9	13,8	43, 47,50, 52,53, 68,101, 106,107

Protocol & Trial

Moderating factors	Studies		Study ID
	(n)	(%)	
Comprehensiveness of intervention description	20	30,8	43, 48,50,52,55,59,62,63,66-70,73,74,79,87,90,96, 106,
Strategies to facilitate implementation	60	92,3	43-74,76-85,88-93,95-102,104-107
Quality of delivery	22	33,8	43-45, 52,53,59,62-64,66,67,70,71,73, 80, 85,89 , 91,96, 100, 102,106
Participant responsiveness	37	56,9	46,48,49,50-53, 55,57,58,62,64-66, 69,72-75, 78-81, 84, 87-91,94,96,97,99,102, 104,106,107
Recruitment	5	7,7	48,52, 73,80, 81,
Context	8	12,3	45,52, 65,66, 69, 89, 94,106

*with methods proposed or used to assess key fidelity components and moderating factors

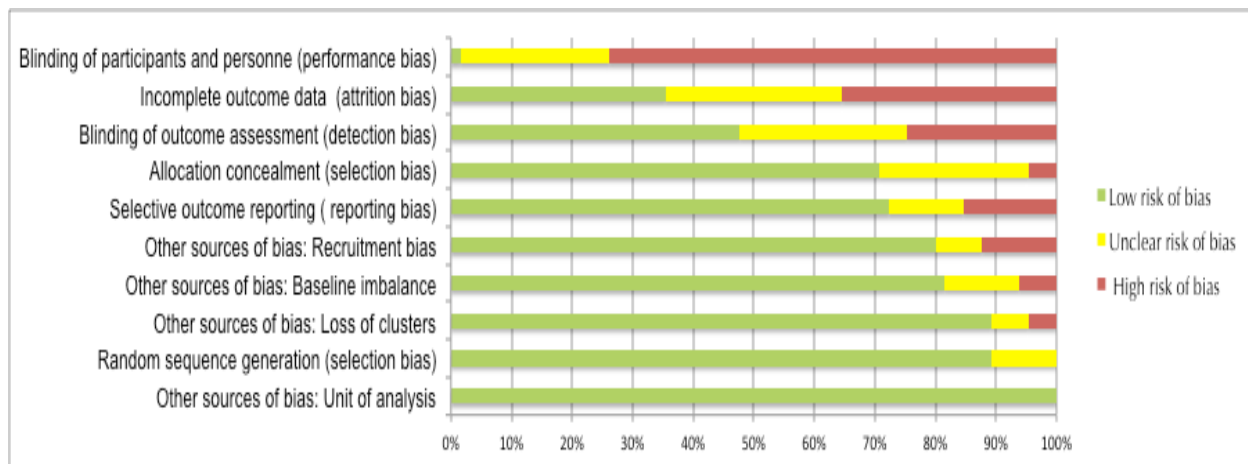
Methods employed to assess fidelity components and moderating factors

Among the 65 the CRTs reviewed, a variety of procedures were used to assess, maintain and monitor intervention fidelity, including: (i) training in intervention delivery, (ii) consultation of manuals and standardized guidelines , (iii) regular meetings between personnel involved in intervention delivery and researchers to provide feedback, support and monitoring during implementation, (iv) logistical and management support for trial coordinators, (v) support teams composed of local area and site managers, (vi) supervision and monitoring to ensure data quality and (viii) field notes, self-reported evaluation or checklists to verify implementation progress. Researchers often reported the strategies used to promote implementation fidelity but failed to describe methods used to assess IF (Dusenbury, Brannigan, & Falco, 2003).

Risk of bias within studies

All studies reported accounting for clustering in the analysis. Due to the nature of the interventions, 74% (48/65) of the studies presented a high risk of bias in the category of "blinding of participants and personnel". Only one study (1,5%) had low risk of bias for all dimensions evaluated (Boisson et al., 2013). The risk of bias graph for individual studies is available in an online supplement (Additional file 7). The overall risk of bias summary is reported in Figure 2.

Figure 2 Summary of risk of bias in included RCTs



Subgroup analyses

To better understand heterogeneity of results concerning IF assessment, we compared studies at low risk of bias versus those at uncertain or high risk of bias (combined). We found no clear differences in IF assessment between subgroups. Results are available in an online supplement (Additional file 8).

DISCUSSION

Fidelity assessment may contribute to making studies more reliable, internally valid, and externally generalizable [(Bellg et al., 2004)]. Our review shows that fidelity assessment is not currently a systematic practice in CRTs of public health interventions carried out in LMICs. In addition, the amount of detail given and data collection methods used for each component were variable between studies. In the absence of fidelity assessment, it may be difficult to determine if CRTs results are due to the intervention design, to its implementation, or to unknown or external factors that may influence results.

Strengths and limitations of the study

The strengths of this review include publication of the study protocol, adherence to the PRISMA guidelines, the use of five bibliographic databases providing comprehensive information on randomised trials (Gasparyan, Ayvazyan, & Kitao, 2013) and, recognizing that word limits for scientific journal articles are highly constrained and that the current CONSORT reporting guidelines for CRTs do not require a description of elements related to IF, extending the search for IF assessment to associated documents beyond the main trial report.

Our study also has limitations. First, we included only CRTs that have the word ‘‘regist*’’ in the abstract. Therefore, we may not have identified all CRTs. However, we reviewed sixteen excluded studies to assess this risk but none met our inclusion criteria. The effect of this limitation is likely to be small. Second, according to the original protocol (M. C. Perez et al., 2016) we were unable to compute outcomes for research questions 4b (Are trial reports with negative findings for the ITT analysis more likely to report a per-protocol (‘‘PP’’) analysis?, 4c (what is the overall agreement between ITT and PP analyses concerning intervention effectiveness?, and 4d (Does the magnitude of the intervention effect differ for PP as compared to ITT analyses?) for two main reasons: (i) It was not always possible to determine whether articles had performed an ITT analysis to assess the intervention effect. Although the CONSORT statement requires that authors indicate whether analysis were performed on an intention-to-treat basis, 16 (25%) studies did not explicitly declare the analysis as ITT or PP. These findings confirm that ITT analysis is ‘‘often inadequately described and inadequately applied’’ (Hollis & Campbell, 1999) in relation to missing outcome

data in randomised controls trials (Alshurafa et al., 2012; Hollis & Campbell, 1999); (ii) many articles reported more than one primary outcome. 58% (38/65) of the studies had more than one primary outcome, while 42% (27/65) of studies have results for one primary outcome. Within studies, conclusions regarding the effectiveness of interventions sometimes diverged. 55% (36/65) of studies reported at least one outcome effective.

Strengths and weaknesses in relation to other studies

This is the first published systematic review focussing on methodological issues related to fidelity assessment in CRTs. Previous reviews have shown that the information required to ensure internal and external validity is often poorly reported in CRTs (S. Eldridge et al., 2008; Ivers et al., 2011). This study complements previous reviews by documenting considerable variation in the practice of fidelity assessment for CRTs of public health interventions in LMICs and suggesting the potential need for methodological improvement in this area.

Interpretation of findings

This review demonstrates that fidelity assessment is not currently a systematic practice in CRTs. Public health interventions are complex and may be affected by multiple contextual factors affecting the internal and external validity of the results. In CRTs contextual differences at the cluster level could favour local adaptation of the intervention and differences in participants' exposure to intervention components (Ibrahim & Sidani, 2015; O'Donnell, 2008). This variability can lead to the heterogeneity of fidelity, reducing power to detect significant effects where they might be present. Fidelity assessment could help to fill this knowledge gap to help researchers attribute outcomes to the actual intervention and identify contextual factors that influence on intervention in CRTs (S. Eldridge et al., 2008; Moore et al., 2015; Oakley et al., 2006). However, current reporting guidelines for CRTs offer no specific guidance on intervention fidelity assessment.

Failure to consider fidelity assessment in CRTs can be explained by several factors. First, approaches to assess intervention fidelity within randomised trials have not yet been made accessible to practitioners. Carroll and colleagues synthesise previous work (Bellg et al., 2004; Borrelli, 2011; Carroll et al., 2007; Hasson, 2010; D. Perez, Van der Stuyft, Zabala, Castro, & Lefevre, 2016), establishing a conceptual framework to evaluate intervention fidelity (Hasson, Blomberg, & Dunér, 2012; Morello et al., 2017; D. Perez et al., 2016; Schinckus, Van den Broucke, Housiaux, & Diabetes Literacy, 2014). However, this framework does not provide guidance on methods to assess specific fidelity dimensions in practice. Second, fidelity assessment is not required at present by the CONSORT guidelines for CRT reporting. This does not encourage researchers to think about the importance of assessing intervention fidelity, nor to report their evaluations if conducted. Third, scientific journals limit the number of words in the final report, which may restrict authors' ability to report their evaluations.

CONCLUSIONS

Randomised trials are generally viewed as the gold standard for establishing evidence of intervention effectiveness. Public health interventions require high-quality evaluations to determine whether the programs work or not and to know how to improve them (Banerjee & Duflo, 2011; Duflo, Glennerster, & Kremer, 2007). Results of this methodologically oriented systematic review demonstrate considerable heterogeneity in the practice of fidelity assessment in CRTs of public health interventions carried out in LMICs. Failure to assess implementation fidelity may weaken the internal and external validity of a study. We offer two recommendations to enhance the quality of future CRTs.

First, CRT reporting guidelines should promote fidelity assessment as a means to facilitate valid causal inference and replication of results. Guidelines such as those produced by the Cochrane network and by CONSORT play an important role in helping authors to strengthen study methods and reporting and thereby in improving the quality of health research (M. J. Campbell, A. Donner, & N. Klar, 2007; M. K. Campbell et al., 2012; Higgins et al., 2011; J. P. Higgins et al., 2016; Hoffmann et al., 2014; Moher et al., 2010). The most recent Cochrane risk of bias guidelines for the first time include consideration of the potential biases arising from deviations from intended

interventions (S.; Eldridge et al., 2016; J. Higgins et al., 2016). Based on the results of this review, we make two suggestions to strengthen CRT methods and reporting. CRT guidelines such as those produced by CONSORT should promote fidelity assessment and reporting to improve the transparency of reports and to help draw firm conclusions about the outcomes of interventions. By the same principle, scientific journals should encourage researchers to conduct and report IF assessment in CRTs.

Second, a practical tool is required to guide trialists involved in the design and implementation of CRTs on how to assess IF. The TIDieR checklist encourages researchers to improve the reporting of interventions and suggests within its 12 items that authors describe how the fidelity of the intervention was assessed (Hoffmann et al., 2014). However, to our knowledge, neither the TIDieR checklist nor any other resource offers researchers a practical guide on how to approach implementation fidelity. This is a fundamental gap that should be remedied.

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials

CRTs: Cluster randomization trials

ICTRP: International Clinical Trials Registry Platform

IF: intervention fidelity

ITT: intention-to-treat

LMICs: low- and middle-income countries

MESH: Medical Subject Headings

PP: per protocol

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RCTs: Randomized controlled trials

TRDS: trial registration data set

WHO: World Health Organization

DECLARATIONS

Funding

IC-IMPACTS (the India-Canada Centre for Innovative Multidisciplinary Partnerships to Accelerate Community Transformation and Sustainability) provided funding for this study in the form of doctoral scholarships for MCP and NM. V. Ridde holds a CIHR-funded Research Chair in Applied Public Health (CPP-137901). MP Sylvestre holds a Research Scholars – Junior 1 award from the Fonds de recherche du Québec (FRSQ). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

All data from this study—including literature searches and data extraction forms- are available from the corresponding author.

Authors' contributors

MJ and MCP conceived the study. MCP, NM, and MJ performed the data abstraction and the data analysis. VR and MPS contributed to study design and interpretation of the results. MCP drafted the manuscript in collaboration with MJ, and all authors contributed to the critical revision and final approval of the manuscript. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. MCP is guarantor.

Competing interests

The authors declare that they no competing interests.

Consent for publication

Not applicable

Ethics approval and consent to participate

This research does not involve human subjects. It is exempted from research ethics board review as it relies exclusively on publicly available information for which there is no reasonable expectation of privacy.

Acknowledgements

Not applicable

Additional File 1

PRISMA Checklist: Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review.

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3,4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	7,8
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	9
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	9,10

Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	11
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	11 (appendix 2)
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	11-12 (appendix 3)
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	11-12
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	12-14
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	14
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	12-14
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	NA

Additional File 2 Full search strategy for MEDLINE

Database: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R)

Daily and Ovid MEDLINE(R) 1946 to Present Interface: OvidSP

Research date: May 13th, 2016

Filters 2012 onwards, English, French, Spanish

#	Search strategy	Résultats
1	Cluster Analysis/	49336
2	Cluster Analys*.tw,kw,sh.	59132
3	(communit* adj2 random* adj2 trial*).tw,kw,sh.	630
4	(group adj2 random* adj2 trial*).tw,kw,sh.	2007
5	(cluster adj2 random*).tw,kw,sh.	7944
6	cluster-random*.tw,kw,sh.	6686
7	or/1-6	67832
8	Afghanistan/ or Libya/ or Albania/ or Macedonia Republic/ or Algeria/ or Madagascar/ or American Samoa/ or Malawi/ or Angola/ or Malaysia/ or Argentina/ or Indian Ocean Islands/ or Armenia/ or Mali/ or Azerbaijan/ or Micronesia/ or Bangladesh/ or Mauritania/ or Republic of Belarus/ or Mauritius/ or Belize/ or Mexico/ or Benin/ or Bhutan/ or Moldova/ or Bolivia/ or Mongolia/ or Bosnia-Herzegovina/ or Montenegro/ or Botswana/ or Morocco/ or Brazil/ or Mozambique/ or Bulgaria/ or Myanmar/ or Burkina Faso/ or Namibia/ or Burundi/ or Nepal/ or Nicaragua/ or Cambodia/ or Niger/ or Cameroon/ or Nigeria/ or Central African Republic/ or Pakistan/ or Chad/ or Palau/ or China/ or Panama/ or Colombia/ or Papua New Guinea/ or Comoros/ or Paraguay/ or Democratic Republic of the Congo/ or Congo/ or Peru/ or Philippines/ or Costa Rica/ or Romania/ or Cote d'Ivoire/ or Rwanda/ or Cuba/ or Samoa/	823062

	<p>or Djibouti/ or Atlantic Islands/ or Dominica/ or Senegal/ or Dominican Republic/ or Serbia/ or Ecuador/ or Seychelles/ or Egypt/ or Sierra Leone/ or El Salvador/ or Melanesia/ or Eritrea/ or Somalia/ or Ethiopia/ or South Africa/ or Fiji/ or Sudan/ or Gabon/ or Sri Lanka/ or Gambia/ or Saint Lucia/ or Georgia/ or "Saint Vincent and the Grenadines"/ or Ghana/ or Grenada/ or Suriname/ or Guatemala/ or Swaziland/ or Guinea/ or Syria/ or Guinea-Bissau/ or Tajikistan/ or Guyana/ or Tanzania/ or Haiti/ or Thailand/ or Honduras/ or East Timor/ or Hungary/ or Togo/ or India/ or Tonga/ or Indonesia/ or Tunisia/ or Iran/ or Turkey/ or Iraq/ or Turkmenistan/ or Jamaica/ or Micronesia/ or Jordan/ or Uganda/ or Kazakhstan/ or Ukraine/ or Kenya/ or Uzbekistan/ or Vanuatu/ or Korea/ or Venezuela/ or Kosovo/ or Vietnam/ or Kyrgyzstan/ or Middle East/ or Yemen/ or Lebanon/ or Lesotho/ or Zambia/ or Zimbabwe/ or Liberia/ or Developing Countries/ or Poverty/ or Rural Population/</p>	
9	<p>(Afghanistan or Albania or Algeria or American Samoa or Angola or Argentina or Armenia or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bhutan or Bolivia or Bosnia Herzegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burundi or Cabo Verde or Cambodia or Cameroon or Central African Republic or Chad or China or Colombia or Comoros or Congo or Costa Rica or Cote d'Ivoire or Cuba or Djibouti or Dominica or Dominican Republic or Ecuador or Egypt or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gambia or Georgia or Ghana or Grenada or Guatemala or Guinea or Guinea-Bissau or Guyana or Haiti or Honduras or Hungary or India or Indonesia or Iran or Iraq or Jamaica or Jordan or Kazakhstan or Kenya or Kiribati or Korea or Kosovo or Kyrgyz or Lao PDR or Lebanon or Lesotho or Liberia or Libya or Macedonia or Madagascar or Malawi or Malaysia or Maldives or Mali or Marshall Islands or Mauritania or Mauritius or Mexico or Micronesia or Moldova or Mongolia or Montenegro or Morocco or Mozambique or Myanmar or Namibia or Nepal or Nicaragua or Niger or Nigeria or Pakistan or Palau or Panama or Papua New Guinea or Paraguay or Peru or Philippines or</p>	792287

	Romania or Rwanda or Samoa or Sao Tome or Senegal or Serbia or Seychelles or Sierra Leone or Solomon Islands or Somalia or South Africa or South Sudan or Sri Lanka or St Lucia or Sudan or Suriname or Swaziland or Syrian Arab Republic or Tajikistan or Tanzania or Thailand or Timor Leste or Togo or Tonga or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Vanuatu or Venezuela or Vietnam or Gaza or Yemen or Zambia or Zimbabwe or Developing Countr* or middle income countr* or low income countr*).tw,kw.	
10	8 or 9	1172790
11	((public adj2 health) or (poverty adj2 economic*) or interven* or program* or approach*).tw,kw,sh.	2459336
12	Public Health/ or Community Health Services/ or Community Health Workers/ or Community Health Planning/ or Community Networks/	107616
13	11 or 12	2515327
14	7 and 10 and 13	4068
15	remove duplicates from 14	3964
16	limit 15 to yr="2012 -Current"	2069
17	limit 16 to (english or french or spanish)	2024

Additional file 3 Articles excluded with reasons

➤ **Stage 1: Excluded after full-text review, with reasons (n=136)**

Study	Reason for exclusion
Adi, A. E., et al. (2015)	Not a CRTs
Aliyu, M. H., et al. (2016)	Article not found
Allison, J. J., et al. (2016)	Not a main trial report
Amza, A., et al. (2012).	Not a CRTs
Anchala, R., et al. (2015)	Not a public health intervention
Anchala, R., et al. (2012).	Not a main trial report
Araya, R., et al. (2013).	Not conducted in LMICs
Arifeen, S. E., et al. (2012)	Not a public health intervention
Arnold, B. F., et al. (2013).	Not a main trial report
Aung, M. N., et al. (2012)	Not a main trial report
Baiden, F., et al. (2016)	Not a public health intervention
Banda, H. T., et al. (2015)	Not a main trial report
Barbosa Filho, V. C., et al. (2015)	Not a main trial report
Barnabas, R. V., et al. (2016)	Not a CRTs
Beattie, T. S., et al. (2015)	Not a main trial report
Bernabe-Ortiz, A., et al. (2014)	Not a main trial report
Berwanger, O., et al. (2012)	Not a public health intervention
Beune, E. J., et al. (2014)	Not conducted in LMICs
Birungi, N., et al. (2015)	Not a main trial report

Blaya, J. A., et al. (2014)	Not a public health intervention
Boisson, S., et al. (2014)	Not a main trial report
Bowen, A., et al. (2012)	Not a main trial report
Browning, C., et al. (2016)	Not a public health intervention
Cavalcanti, A. B., et al. (2016).	Not a public health intervention
Chan, A., et al. (2015).	Not a main trial report
Chandir, S., et al. (2014)	Not a CRTs
Chen, S., et al. (2015)	Not a public health intervention
Chen, Y., et al. (2014)	Not a public health intervention
Chen, Y., et al. (2015).	Not a main trial report
Chibanda, D., et al. (2015)	Not a main trial report
Chinbuah, M. A., et al. (2013)	Not a public health intervention
Christian, P., et al. (2013)	Not a main trial report
Coulibaly, Y. I., et al. (2013)	Not a public health intervention
Crea, T. M., et al. (2015).	Not a main trial report
Cundill, B., et al. (2015).	Not a public health intervention
Drummond, J., et al. (2014)	Not conducted in LMICs
Dryden-Peterson, S., et al. (2015).	Not a public health intervention
Du, X., et al. (2014)	Not a public health intervention
Dumont, A., et al. (2013).	Not a public health intervention
Durovni, B., et al. (2014).	Not a public health intervention
Edmond, K., et al. (2012)	Not a main trial report

Ellard, D., et al. (2012)	Not a main trial report
Engebretsen, I. M., et al. (2014)	Not a main trial report
Eriksson, L., et al. (2016)	Not a main trial report
Falb, K. L., et al. (2016)	Not a main trial report
Gallegos-Carrillo, K., et al. (2014)	Not a main trial report
Gannon, B., et al. (2014)	Not a public health intervention
Gibbs, L., et al. (2015)	Not conducted in LMICs
Halliday, K. E., et al. (2014)	Not a main trial report
Hargreaves, J. R., et al. (2015)	Not a main trial report
Henao-Restrepo, A. M., et al. (2015)	Not a main trial report
Hess, S. Y., et al. (2015)	Not a public health intervention
Hirschhorn, L. R., et al. (2015)	Not a main trial report
Investigators, C.-I., et al. (2015)	Not a main trial report
Jemmott, J. B., 3rd, et al. (2014).	Not a main trial report
Jenkins, R., et al. (2013)	Not a public health intervention
Jiamjariyaporn, T., et al. (2014)	Not a main trial report
Kajula, L., et al. (2016)	Not a main trial report
Kalyango, J. N., et al. (2013)	Not a main trial report
Kalyango, J. N., et al. (2013)	Not a main trial report
Kent, L. M., et al. (2015).	Not a main trial report
Kestler, E., et al. (2013)	Not a main trial report
Krebs, N. F., et al. (2012)	Not a public health intervention

Kyriacos, U., et al. (2015)	Not a public health intervention
Laezer, K. L., et al. (2013)	Not a main trial report
Lakshminarayana, R., et al. (2013).	Not a public health intervention
Lawal, I. U., et al. (2015).	Not a main trial report
Li, W. H., et al. (2015)	Not a main trial report
Lignou, S., et al. (2016)	Not a main trial report
Lin, Q., et al. (2015)	Not a main trial report
Lippman, S. A., et al. (2016)	Not a main trial report
Liu, J. Y. and C. Lai (2014)	Not a main trial report
Liu, X., et al. (2015)	Not a public health intervention
Machado, F., et al. (2015)	Not a main trial report
MacIntyre, C. R., et al. (2015)	Not a public health intervention
Martin, R. M., et al. (2013)	Not a public health intervention
Maselko, J., et al. (2015)	Article not found
Mash, B., et al. (2012)	Not a main trial report
Mazumder, S., et al. (2014)	Not a main trial report
Mbacham, W. F., et al. (2014)	Not a public health intervention
McClure, E. M., et al. (2014)	Not a main trial report
McNairy, M. L., et al. (2015)	Not a main trial report
Menya, D., et al. (2013)	Not a main trial report
Monse, B., et al. (2012)	Not a public health intervention
Nabulsi, M., et al. (2014)	Not a main trial report

Nery, S. V., et al. (2015)	Not a main trial report
Nsibande, D., et al. (2013)	Not a main trial report
Noa Noatina, B., et al. (2013)	Not a CRTs
Ogedegbe, G., et al. (2014)	Not a main trial report
Ogedegbe, G., et al. (2014)	Not conducted in LMICs
Ojha, S., et al. (2014)	Not a main trial report
Oken, E., et al. (2013).	Not a main trial report
Oluoch, T., et al. (2016)	Not a public health intervention
Operario, D., et al. (2016)	Not a public health intervention
Orne-Gliemann, J., et al. (2015)	Not a main trial report
Ota, M. O. C., et al. (2012)	Not a main trial report
Overgaard, H. J., et al. (2012)	Not a main trial report
Peltzer, K., et al. (2013)	Not a public health intervention
Penfold, S., et al. (2014).	Not a main trial report
Pettifor, A., et al. (2015)	Not a main trial report
Prado, E. L., et al. (2016)	Not a public health intervention
Prado, E. L., et al. (2012).	Not a public health intervention
Qu, C., et al. (2014)	Not a main trial report
Rabbani, F., et al. (2014)	Not a main trial report
Ramsey, K., et al. (2013)	Not a main trial report
Rasoamanamihaja, C. F., et al. (2016)	Not a CRTs
Robertson, L., et al. (2012)	Not a main trial report

Rustagi, A. S., et al. (2016)	Not a public health intervention
Sando, D., et al. (2014)	Not a main trial report
Sathish, T., et al. (2013)	Not a main trial report
Scott, S., et al. (2014)	Not a main trial report
Senarathna, L., et al. (2013)	Not a public health intervention
Shafique, S., et al. (2016)	Not a public health intervention
Shah, L., et al. (2015)	Not a main trial report
Shah More, N., et al. (2013)	Not a main trial report
Shelley, D., et al. (2015)	Not a main trial report
Sherr, K., et al. (2014)	Not a main trial report
Shet, A. S., et al. (2015)	Not a main trial report
Solomon, E., et al. (2014)	Not conducted in LMICs
Some, J. W., et al. (2015)	Not a public health intervention
Soofi, S., et al. (2012)	Not a public health intervention
Stinson, K., et al. (2014)	Not a main trial report
Sunguya, B. F., et al. (2013)	Not a main trial report
Taneja, S., et al. (2015)	Not a main trial report
Tine, R. C., et al. (2014)	Not a main trial report
Tiono, A. B., et al. (2015)	Not a main trial report
Tomlinson, M., et al. (2015)	Not a main trial report
Turan, J. M., et al. (2012)	Not a main trial report
Vijayakumar, L., et al. (2013)	Not a main trial report

Wagner, G. J., et al. (2014)	Not a main trial report
West, K. P., Jr., et al. (2014)	Not a public health intervention
Widyawati, W., et al. (2014)	Not a main trial report
Xu, F., et al. (2014)	Not a main trial report
Xu, W. H., et al. (2014)	Not a main trial report
Yore, J., et al. (2016)	Not a main trial report
Zlotkin, S., et al. (2013)	Not a public health intervention

➤ **Stage 2: Excluded after full-text review of protocol, main trial report, and associated publications, with reasons (n =25)**

Study	Reason for exclusion
Abdullah, F., et al. (2013).	Do not assess implementation fidelity
Arrossi, S., et al. (2015).	Do not assess implementation fidelity
Baird, S. J., et al. (2012)	Do not assess implementation fidelity
Cunha, D. B., et al. (2013).	Do not assess implementation fidelity
Deribew, A., et al. (2012).	Do not assess implementation fidelity
Doherty, T., et al. (2013).	Do not assess implementation fidelity
Ezeanolue, E. E., et al. (2015).	Do not assess implementation fidelity
Garcia, P. J., et al. (2012).	Do not assess implementation fidelity
Gebre, T., et al. (2012).	Do not assess implementation fidelity
Hong, X. C., et al. (2013)	Do not assess implementation fidelity
Joos, O., et al. (2016).	Do not assess implementation fidelity

Kim, D. A., et al. (2015).	Do not assess implementation fidelity
Labhardt, N. D., et al. (2014).	Do not assess implementation fidelity
MacPherson, P., et al. (2014)	Do not assess implementation fidelity
Maust, A., et al. (2015)	Do not assess implementation fidelity
Miller, G., et al. (2012).	Do not assess implementation fidelity
Mori, R., et al. (2015).	Do not assess implementation fidelity
Pickering, A. J., et al. (2015).	Do not assess implementation fidelity
Puchalski Ritchie, L. M., et al. (2015)	Do not assess implementation fidelity
Qadri, F., et al. (2015).	Do not assess implementation fidelity
Robertson, L., et al. (2013)	Do not assess implementation fidelity
Tol, W. A., et al. (2014).	Do not assess implementation fidelity
Tola, H. H., et al. (2016)	Do not assess implementation fidelity
Xu, F., et al. (2015).	Do not assess implementation fidelity
Yektatalab, S., et al. (2016).	Do not assess implementation fidelity

Additional file 4a Tool for assessing risk of bias^(a)

Domain	Support for judgement
Selection bias	
Random sequence generation. Was the allocation sequence adequately generated?	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups. <ul style="list-style-type: none">• The use of a random component should be sufficient for adequate sequence generation.• Sometimes restricted randomization is used to generate a sequence to ensure particular allocation ratios to the intervention groups (e.g. 1:1). Blocked randomization (random permuted blocks) is a common form of restricted randomization. Also common is stratified randomization, in which restricted randomization is performed separately within strata.
Allocation concealment. Was the allocation sequence adequately concealed?	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment <ul style="list-style-type: none">• Defined as “the process used to ensure that the person deciding to enter a participant into a randomised controlled trial does not know the comparison group into which that individual will be allocated” Among the different methods used to conceal allocation, central randomization by a third party is perhaps the most desirable. Methods using envelopes are more susceptible to manipulation than other approaches (Schulz 1995b). If investigators use envelopes, they should develop and monitor the allocation process to preserve concealment. In addition to use of sequentially numbered, opaque, sealed envelopes, they should ensure that the envelopes are opened sequentially, and only after the envelope has been irreversibly assigned to the participant.

Performance bias

Blinding of participants and personnel Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.

Was blinding of participants and personnel adequately addressed?

- After enrolment into the study, **blinding** (or masking) **of study participants and personnel** may reduce the risk that knowledge of which intervention was received, rather than the intervention itself, affects outcomes. Effective blinding can also ensure that the compared groups receive a similar amount of attention, ancillary treatment and diagnostic investigations
- Performance bias concerning the fidelity of the interventions, and quality of the information regarding who received what interventions, including blinding of participants and healthcare providers.

Detection bias

Blinding of outcome assessment Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective

Were assessors adequately blinded to what intervention a participant received?

- Different types of people can be blinded in a clinical trial (Gøtzsche 1996, Haahr 2006):
 1. participants (e.g. patients or healthy people);
 2. healthcare providers (e.g. the doctors or nurses responsible for care);
 3. **outcome assessors, including primary data collectors (e.g. interview staff responsible for measurement or collection of outcome data) and any secondary assessors (e.g. external outcome adjudication committees);**
 4. data analysts (e.g. statisticians); and
 5. manuscript writers.

The first two type of people are addressed in the tool by the item ‘Blinding of participants and personnel’. **The third is addressed by the item ‘Blinding of outcome assessment’**. The last two are not explicitly covered by the tool.

Attrition bias

Incomplete outcome data

Data on incomplete outcomes reported?

Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.

- We scored:

Overall:

- ✓ Score low risk if <10% of outcomes data is missing at the individual level AND, for trials reporting some missing data, reasons for missing data are given OR if a sensitivity analysis conducted and conclusion of bias unlikely
 - ✓ Score unclear risk
 - o if the amount of missing outcomes data is unclear
 - o if <10% outcomes data is missing but, for trials reporting some missing data, reasons for missing data are NOT given
 - o if more than >10% but less than <20% of outcomes data is missing : a sensitivity analysis (b) for missing data was conducted and the results demonstrated that missing data were unlikely to bias results
-

✓ Score high risk

- o if more than >10% but less than <20% of outcomes data are missing and no sensitivity analysis conducted OR sensitivity analysis conducted & results indicate risk of bias
- o if more than >20% of outcomes data are missing

Intervention vs. control:

- ✓ Score low risk if there are no statistically significant differences in missing outcomes data between individuals in intervention and control groups $p < 0.05^{(c)}$ at endline
- ✓ Score unclear risk
 - o if the potential for differences in missing outcomes data between individuals in intervention and control groups is not evaluated/ reported
 - o if there are statistically significant differences in missing outcomes data between individuals in intervention and control groups $p < 0.05^{(c)}$ addressed through imputation or sensitivity analysis
- ✓ Score high risk if there are statistically significant differences in missing outcomes data between individuals in intervention and control groups $p < 0.05^{(c)}$ & these are not adequately addressed through missing values (imputation or sensitivity analysis) analysis.

« High rates of missing outcome data are present in the majority of CRTs, yet handling missing data in practice remains suboptimal. Researchers and applied statisticians should carry out appropriate

missing data methods, which are valid under plausible assumptions in order to increase statistical power in trials and reduce the possibility of bias. Sensitivity analysis should be performed, with weakened assumptions regarding the missing data mechanism to explore the robustness of results reported in the primary analysis » (Fiero et al, 2016)

Scoring decision: Score high if evaluated as high risk by *either* overall or by group criteria; unclear if unclear for *both* overall and by group criteria, OR unclear in one and low in the other; low risk if low risk in both overall and by group criteria

Reporting bias

Selective reporting

State how the possibility of selective outcome reporting was examined by the review authors, and what was found.

Was the study free from selective outcome reporting?

If the protocol is available, then outcomes in the protocol and published report can be compared. If not, then outcomes listed in the methods section of an article can be compared with those whose results are reported. Some differences between protocol and publication may be explained by legitimate changes to the protocol.

- ✓ We have analyzed the changes made in the primary outcome. If change made it was High risk. If a secondary outcome is not reported, we ignore it. A possible exception might be when the change is declared in advanced of looking at the data. Unclear risk if the language was refined (prevalence/incidence), but the main outcome is the same

Other bias*

Recruitment bias:

Recruitment bias can occur when participants are recruited to the trial after the clusters have been randomly allocated. All eligible participants ideally must be identified before clusters are randomised

Are all eligible participants (or a random sample thereof) identified before clusters are randomised?

« Differential individual recruitment or consent—This is when different proportions of participants are recruited to the different arms of the trial. If recruitment rates differ between groups this may lead to the risk of bias » (Puffer 2003)

CRTs can have a cohort design (same people at baseline and endline) or be based on parallel cross-sections (different groups at each time period).

- ✓ We scored high risk if randomization of clusters was done and individuals were recruited afterwards
- ✓ We scored low risk if the study design was a survey pre randomization and a survey post randomization, and not the same individuals followed up.

Baseline imbalance

Was baseline comparability of clusters reported, or did the authors report statistical adjustments or randomization methods to render baseline characteristics comparable (e.g. stratified or pair-matched

Although not a form of bias as such, the risk of baseline differences can be reduced by using stratified or pair-matched randomization of clusters. Reporting of the baseline comparability of clusters, or statistical adjustment for baseline characteristics, can help reduce concern about the effects of baseline imbalance.

randomization of clusters)?

Loss of clusters

In CRTs, participants lost to follow-up can occur at the level of the cluster (cluster withdrawal or lost to follow-up, inactive cluster, merging of clusters) or the individual (participant withdrawal or lost to follow-up, or transfer from one cluster to another) and have to be omitted from the analysis this may lead to bias. Take into account any cluster / participant who withdrew or was lost to follow-

Were any clusters lost to follow up during the trial, and if so, were ad hoc missing data methods used to take this into account?

up by using ad hoc missing data methods

We score :

- ✓ Low risk if less than 10% of clusters lost or if no clusters lost
- ✓ Unclear risk
 - if the number of clusters missing was not clear
 - if more than 10% of clusters lost & [(a sensitivity analysis for missing data was conducted) & (the results demonstrated that missing data were unlikely to bias results)]
- ✓ Score high risk if more than 10% of clusters lost & [(no sensitivity analysis conducted) OR (sensitivity analysis conducted & results indicate risk of bias)]

“Some trials lost complete clusters after randomisation. However, the proportion of clusters lost ...was relatively low and therefore would be unlikely to introduce bias” (Puffer 2003)

« Take into account any cluster that withdrew or was lost to follow-up by using ad hoc missing data methods (+ITT) for participants included in these clusters » (Giraudeau B, Ravaud P (2009) ^(d)

Unit of analysis

Many cluster-randomized trials are analysed by incorrect statistical methods, not taking the clustering into account. If a cluster-randomised trial is analysed without recognition of the clustering, the analysis will ignore the possible correlation between members of the same cluster. When

Does the statistical analysis positively correlated observations are treated as independent, the result may be standard errors which consider clustering, or, are too small, confidence intervals which are too narrow, and P values which are too small, leading alternatively, is analysis conducted to conclusions which may be false. One way to avoid unit-of-analysis errors in cluster-randomized trials is to conduct the analysis at the same level as the allocation, however, this might reduce the power of the study, depending on the number and size of the clusters. The ideal information to extract from a cluster-randomized trial is a direct estimate of the required effect measure (for example, an odds ratio with its confidence interval) from an analysis that properly accounts for the cluster design

^(a) Synthesis from The Cochrane Handbook for Systematic Reviews of Interventions (2007).; Hayes, R. J., & Moutlon, L. H. (2009). Cluster randomised trials (Chapman & Hall/CRC Interdisciplinary Statistics) ; Donner, A., & Klar, N. (2000). Design and analysis of cluster randomization trials in health research (pp.6-10)

^(b)Fiero, M. H., Huang, S., Oren, E., & Bell, M. L. (2016). Statistical analysis and handling of missing data in cluster randomized trials: a systematic review. *Trials*, 17(1), 72.

^(c) Puffer, S., Torgerson, D., & Watson, J. (2003). Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *Bmj*, 327(7418), 785-789.

^(d) Giraudeau B, Ravaud P (2009) Preventing Bias in Cluster Randomised Trials. *PLoS Med* 6(5): e1000065. doi:10.1371/journal.pmed.1000065

Additional file 4b Criteria for judging risk of bias in the ‘Risk of bias’s assessment Cochrane Collaboration tool

RANDOM SEQUENCE GENERATION

Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.

Criteria for a judgement of ‘Low risk’ of bias. The investigators describe a random component in the sequence generation process such as:

- Referring to a random number table;
- Using a computer random number generator;
- Coin tossing;
- Shuffling cards or envelopes;
- Throwing dice;
- Drawing of lots;
- Minimization*.

*Minimization may be implemented without a random element, and this is considered to be equivalent to being random.

Criteria for the judgement of ‘High risk’ of bias. The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example:

- i. Sequence generated by odd or even date of birth;
- ii. Sequence generated by some rule based on date (or day) of admission;
- iii. Sequence generated by some rule based on hospital or clinic record number.

Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-random categorization of participants, for example:

- Allocation by judgement of the clinician;

- Allocation by preference of the participant;
- Allocation based on the results of a laboratory test or a series of tests;
- Allocation by availability of the intervention.

Criteria for the judgement of 'Unclear risk' of bias. Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'.

ALLOCATION CONCEALMENT

Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.

Criteria for a judgement of 'Low risk' of bias. Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation:

- Central allocation (including telephone, web-based and pharmacy-controlled randomization);
- Sequentially numbered drug containers of identical appearance;
- Sequentially numbered, opaque, sealed envelopes.

Criteria for the judgement of 'High risk' of bias. Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on:

- Using an open random allocation schedule (e.g. a list of random numbers);
- Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or nonopaque or not sequentially numbered);
- Alternation or rotation;
- Date of birth;
- Case record number;
- Any other explicitly unconcealed procedure.

Criteria for the judgement of 'Unclear risk' of bias. Insufficient information to permit judgement of 'Low risk' or 'High risk'. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement – for example if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque and sealed.

BLINDING OF PARTICIPANTS AND PERSONNEL

Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.

Criteria for a judgement of 'Low risk' of bias. Any one of the following:

- No blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding;
- Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken.

Criteria for the judgement of 'High risk' of bias. Any one of the following:

- No blinding or incomplete blinding, and the outcome is likely to be influenced by lack of blinding;
- Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken, and the outcome is likely to be influenced by lack of blinding.

Criteria for the judgement of 'Unclear risk' of bias. Any one of the following:

- Insufficient information to permit judgement of 'Low risk' or 'High risk';
- The study did not address this outcome.

BLINDING OF OUTCOME ASSESSMENT

Detection bias due to knowledge of the allocated interventions by outcome assessors.

Criteria for a judgement of ‘Low risk’ of bias. Any one of the following:

- No blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding;
- Blinding of outcome assessment ensured, and unlikely that the blinding could have been broken.

Criteria for the judgement of ‘High risk’ of bias. Any one of the following:

- No blinding of outcome assessment, and the outcome measurement is likely to be influenced by lack of blinding;
- Blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome measurement is likely to be influenced by lack of blinding.

Criteria for the judgement of ‘Unclear risk’ of bias. Any one of the following:

- Insufficient information to permit judgement of ‘Low risk’ or ‘High risk’;
- The study did not address this outcome.

INCOMPLETE OUTCOME DATA

Attrition bias due to amount, nature or handling of incomplete outcome data.

Criteria for a judgement of ‘Low risk’ of bias. Any one of the following:

- No missing outcome data;
- Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias);
- Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups;
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate;

- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size;
- Missing data have been imputed using appropriate methods.

Criteria for the judgement of

‘High risk’ of bias.

Any one of the following:

- Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups;
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate;
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size;
- ‘As-treated’ analysis done with substantial departure of the intervention received from that assigned at randomization;
- Potentially inappropriate application of imputation.

Criteria for the judgement

of ‘Unclear risk’ of bias.

Any one of the following:

- Insufficient reporting of attrition/exclusions to permit judgement of ‘Low risk’ or ‘High risk’ (e.g. number randomized not stated, no reasons for missing data provided);
- The study did not address this outcome.

SELECTIVE REPORTING

Reporting bias due to selective outcome reporting.

Criteria for a judgement of ‘Low

risk’ of bias.

Any of the following:

- The study protocol is available and all of the study’s pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way;

<p>Criteria for the judgement of 'High risk' of bias.</p>	<ul style="list-style-type: none"> • The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon). <p>Any one of the following:</p> <ul style="list-style-type: none"> • Not all of the study's pre-specified primary outcomes have been reported; • One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; • One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); • One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; • The study report fails to include results for a key outcome that would be expected to have been reported for such a study.
<p>Criteria for the judgement of 'Unclear risk' of bias.</p>	<p>Insufficient information to permit judgement of 'Low risk' or 'High risk'. It is likely that the majority of studies will fall into this category.</p>
<p>OTHER BIAS</p>	
<p>Bias due to problems not covered elsewhere in the table.</p>	
<p>Criteria for a judgement of 'Low risk' of bias.</p>	<p>The study appears to be free of other sources of bias.</p>
<p>Criteria for the judgement of 'High risk' of bias.</p>	<p>There is at least one important risk of bias. For example, the study:</p> <ul style="list-style-type: none"> • Had a potential source of bias related to the specific study design used; or • Has been claimed to have been fraudulent; or • Had some other problem.

Criteria for the judgement of There may be a risk of bias, but there is either:

'Unclear risk' of bias.

- Insufficient information to assess whether an important risk of bias exists; or
- Insufficient rationale or evidence that an identified problem will introduce bias.

Additional file 5 Characteristics of included studies

Study ID	Study Design	Country	Intervention Name	Aims	IF components	
					Protocol	Trial reports ¹
Abrams ky, T., et al. (2014)	2-arm parallel	Uganda	SASA!	To assess the community level impacts of SASA to prevent violence against women and reduce HIV-risk behaviours.	CONT	CONT
					COV	COV
					FREQ	FREQ
						DUR
Althabe , F., et al. (2015)	2-arm parallel	Argentina, Zambia, Guatemala, India, Pakistan and Kenya	ACT	To assess the feasibility, effectiveness, and safety of an intervention to increase antenatal corticosteroid use for the reduction of neonatal mortality	CONT	CONT
					COV	COV
					FREQ	FREQ

Andersson, N., et al. (2015)	2-arm parallel	Nicaragua and Mexico	Camino Verde	To test whether community mobilization based on feedback for dengue immunological status and entomological indicators adds effectiveness to conventional dengue control	--	CONT
Andrade, S., et al. (2014)	2-arm parallel	Ecuador	ACTIVITAL	To evaluate the impact of a school-based health promotion intervention for adolescents on physical fitness indicators	--	CONT COV FREQ
Ansah, E. K., et al. (2015)	2-arm parallel	Ghana	--	To examine the impact of providing rapid diagnostic tests for malaria on antimalarial dispensing, with the aim of reducing current over diagnosis and overtreatment of malaria	--	CONT FREQ
Attanasio, O. P., et al. (2014)	2 x 2 factorial	Colombia	--	To assess the effectiveness of an integrated early child development intervention (stimulation and micronutrient supplementation), on children's development, growth, and haemoglobin levels.	--	CONT FREQ

Aung, T., et al. (2014)	2-arm parallel r	Myanma	--	To test the impact of a social franchising program on uptake of oral rehydration solution plus zinc for childhood diarrhoea	COV	CONT COV
Awasthi, S., et al. (2013)	2 x 2 factorial	India	Deworming and Enhanced Vitamin A Supplementation DEVTA Project	To assess the effects of a widely practicable periodic deworming regimen (albendazole) and Vitamin A (retinol) on mortality at ages 1·0–6·0 years	CONT FREQ	CONT COV FREQ
Ayles, H., et al. (2013)	2 x 2 factorial	Zambia and South Africa	ZAMSTAR	To reduce the burden of tuberculosis by facilitating either rapid sputum diagnosis or integrating tuberculosis and HIV services within the community	--	CONT
Baker-Henningsham, et al. (2012)	2-arm parallel	Jamaica	The Incredible Years Teacher	To determine the effects of a universal pre-school-based intervention on child conduct problems and social skills at school and at home.	--	CONT COV FREQ

		Training Programme					
Bhanda ri, N., et al. (2012)	2-arm parallel	India	Training health workers to implement Integrated Managemen t of Neonatal and Childhood Illness (IMNCI) programme	To evaluate the IMNCI programme, which integrates improved treatment of illness for children with home visits for newborn care, to inform its scale-up	--	CONT COV FREQ	
Bird, C., et al. (2014)	2-arm parallel	Zanzibar	--	To test whether footwear reduces prevalence and intensity of hookworm infection in school-aged children	--	CONT	
Boisson , S., et	2-arm parallel	India	--	To evaluate the effect of promotion and free distribution of sodium dichloroisocyanurate (NaDCC) disinfection	CONT	CONT COV	

al. (2013)				tablets for household water treatment in preventing diarrhoea prevalence among children under five.		
Boone, P., et al. (2016)	2-arm parallel	Guinea-Bissau	The EPICS trial: Enabling Parents to Increase Child Survival	To assess whether a community-based intervention package in the absence of health system strengthening activities could generate a rapid and cost-effective reduction in under-5 mortality in these regions.	CONT	CONT COV DUR
Bousema, T., et al. (2016)	2-arm parallel	Kenya	--	To measure the effect of hotspot-targeted interventions with larviciding, distribution of long-lasting insecticide-treated nets, indoor residual spraying, and focal mass drug administration in evaluation zones surrounding malaria hotspots.	CONT	CONT COV FREQ
Burnham, N. H., et al. (2015)	2-arm parallel	South Africa	--	To test the effectiveness of Team Awareness (TA), a programme aimed at reducing the risky use of alcohol and alcohol-related HIV risk and increasing help-seeking behaviour among a sample of employees.	--	CONT

Chen-Hussey, V. et al. (2013)	2-arm parallel	Lao PDR	--	To determine whether using a topical 15% DEET repellent, would reduce malaria incidence against exophagic vectors amongst rural populations using LLINs.	CONT FREQ	CONT FREQ
Clasen, T., et al. (2014)	2-arm parallel	India	--	To assess the effectiveness of a rural household sanitation intervention (provision of household latrines) to prevent diarrhoea, soil-transmitted helminth infection, and child malnutrition.	CONT COV FREQ DUR	CONT COV FREQ DUR
Coates, T. J., et al. (2014)	2-arm parallel	South Africa, Tanzania, Zimbabwe and Thailand	--	To investigate whether a social and behavioural prevention strategy could reduce HIV incidence, increase HIV testing, reduce HIV risk behaviour, and change social and behavioural norms.	CONT	CONT
Colbourn, T., et al. (2013)	2 x 2 factorial	Malawi	--	To investigate whether (1) a rural participatory women's group community intervention, and (2) a quality improvement intervention at health centres, are effective in reducing deaths in three districts of Malawi.	CONT FREQ	CONT COV FREQ

							DUR
Corbel, V., et al. (2012)	2 x 2 factorial	Benin		--	To investigate whether the combination of long-lasting insecticidal mosquito nets (LLINs) with indoor residual spraying (IRS) or carbamate-treated plastic sheeting (CTPS) conferred enhanced protection against malaria and better management of pyrethroid-resistance in vectors than did LLINs alone	CONT COV	CONT COV DUR
Dave, P. V., et al. (2016)	2-arm parallel	India		--	Non-inferiority trial to assess the effect of using family members as DOT providers on treatment success rates among children with newly diagnosed TB.	--	CONT
Deressa, W., et al. (2014).	2-arm parallel	Ethiopia		--	To determine the effect of combining community-based mosquito repellent with LLINs in the reduction of malaria	CONT	CONT
Devries, K. M., et al. (2015)	2-arm parallel	Uganda	The Good Schools Study		To determine whether the Good School Toolkit could reduce physical violence from school staff to students	CONT	CONT

Durovni, B., et al. (2013)	step wedge	Brazil	--	To assess the effect of widespread use of isoniazid preventive therapy on rates of tuberculosis and death in people with HIV in Brazil.		CONT
Fairall, L., et al. (2012)	2-arm parallel	South Africa	Streamlining Tasks and Roles to Expand Treatment and Care for HIV (STRETCH) programme	To assess the effects on mortality, viral suppression, and other health outcomes and quality indicators of the STRETCH programme, which provides educational outreach training of nurses to initiate and represcribe ART, and to decentralise care.	--	FREQ DUR
Feikin, D. R., et al. (2014)	2-arm parallel	Kenya	--	To investigate whether providing zinc at home for future diarrhoea treatment had an impact at the community level on the reduction of diarrhoea incidence, reduction of respiratory disease incidence, and increased diarrhoea episodes treated with zinc, as compared to having zinc available in the clinic only	CONT	CONT

Fottrell, E., et al. (2013)	2-arm parallel	Bangladesh	--	To assess the effect of a participatory women's group intervention involving participatory learning and action cycles with higher population coverage on neonatal mortality in Bangladesh	--	CONT COV FREQ
Fylkesnes, K., et al. (2013)	2-arm parallel	Zambia	--	To assess the impact of a home-based voluntary human immunodeficiency virus (HIV) counselling and testing model on acceptance, equity in uptake, and negative life events, compared to standard testing services.	--	CONT
Go, V. F., et al. (2015)	2 x 2 factorial	China	--	To develop and evaluate a behavioural intervention for HIV-infected PWID involving (1) individual-level post test counselling and skill-building support groups, (2) a structural-level community stigma reduction program, (3) both individual and structural level activities, as compared to (4) standard of care HIV testing and counselling.	CONT COV FREQ	CONT COV FREQ
Gong, J., et al. (2015)	2-arm parallel	China	Keep Moving toward Healthy Heart and	To evaluate the efficacy of the KM2H2 program in encouraging physical activities for the prevention of heart attack and stroke among hypertensive patients enrolled in	--	CONT

			Healthy Brain (KM2H2)	the Community-Based Hypertension Control Program (CBHCP).		
Gunawar dena, N., et al. (2016)	2-arm parallel	Sri Lanka	--	To examine the effect of a school-based intervention to enable school children to act as change agents on weight, physical activity and diet of their mothers	--	DUR
Halliday, K. E., et al. (2014)	2 x 2 factorial	Kenya	--	To investigate the effect of (1) intermittent screening and treatment (IST) for malaria alone; (2) a literacy intervention alone; (3) both interventions combined; or (4) control group where neither intervention was implemented, on the health and education of school children in an area of low to moderate malaria transmission.	CONT FREQ	CONT COV FREQ
Hanson, C., et al. (2015)	2-arm parallel	Tanzania	--	To develop, implement and evaluate the effectiveness and cost of interventions at community level (focussed on a community-based health worker) and of health system strengthening on newborn survival in rural southern Tanzania	CONT FREQ	CONT COV FREQ

He, F. J., et al. (2015)	2-arm parallel	China	School-EduSalt	To determine whether an education programme targeted at schoolchildren can lower salt intake in children and their families.	CONT	CONT
Huybre gts, L., et al. (2012)	2-arm parallel	Chad	Prevent Acute Child Malnutrition (PREAMA)	To examine the effect of a ready-to-use supplementary food (RUSF) on prevention of wasting in 6- to 36-mo-old children within the context of a general food distribution program	CONT	--
Jack, S. J., et al. (2012)	2-arm parallel	Cambodian	--	To evaluate the effectiveness of micronutrient Sprinkles alongside infant and young child feeding (IYCF) education compared with IYCF education alone on anaemia, deficiencies in iron, vitamin A, and zinc, and infant growth.	--	CONT COV
Kirkwood, B. R., et al. (2013)	2-arm parallel	Ghana	Newborn home visits neonatal mortality trial (NEWHINTS)	To assess the effect of home visits by community-based surveillance volunteers to pregnant women and their babies on all-cause neonatal mortality rate and essential newborn-care practices.	CONT COV FREQ	CONT COV FREQ

Lewyck a, S., et al. (2013)	2 x 2 facto- rial	Malawi	MaiMwana	To assess whether community mobilisation through women's groups, and health education through female volunteer peer counsellors affect rates of infant care, feeding, morbidity, and mortality	CONT DUR	CONT COV FREQ
Lutge, E., et al. (2013)	2-arm parallel	South Africa	--	To test the feasibility and effectiveness of delivering economic support (vouchers) to improve pulmonary tuberculosis treatment outcomes	--	CONT COV FREQ
Mengist ie, B., et al. (2013)	2-arm parallel	Ethiopia	--	To assess the effectiveness of household water chlorination in reducing incidence of diarrhoea among children <5 years of age.	--	CONT
More, N. S., et al. (2012)	2-arm parallel	Mumbai	--	To test whether slum-dweller women's groups discussing perinatal health, lead to improved knowledge through peer learning, and development and implementation of local strategies to improve perinatal care and outcomes.	CONT COV	CONT COV
Nikiem a, L., et al.	3-arm parallel	Burkina Faso	LUCOMA	To compare the effectiveness of weekly child-centred counselling (CCC), with an improved cornsoy blend	CONT	CONT

al. (2014)				[corn-soy blend with added micronutrients (CSB++)] or a locally produced ready-to-use supplementary food (RUSF), in treating moderate acute malnutrition through first-line rural health services.			FREQ
Olney, D. K., et al. (2015)	3-arm parallel	Burkina Faso	EHFP (enhanced-homestead food production program)	To assess the impact of EHFP, Helen Keller International's integrated agriculture and nutrition and health behaviour change communication program.	--		CONT
Ononge, S., et al. (2015)	step wedge	Uganda	--	To determine whether antenatal distribution of misoprostol for pregnant women to self-administer during home births reduces postpartum haemorrhage.	--		CONT COV
Onwuje kwe, O., et al. (2015)	3-arm parallel	Nigeria	--	To evaluate the effectiveness of provider training alone or in combination with a school-based community intervention in improving adherence to malaria treatment guidelines	CONT		CONT COV

Pasha, O., et al. (2013)	2-arm parallel	India, Pakistan, Kenya, Zambia, Guatemala and Argentina	EmONC	To evaluate the impact of a comprehensive intervention of community mobilization, birth attendant training and improvement of quality of care in health facilities on perinatal mortality	CONT	CONT
Patil, S. R., et al. (2014)	2-arm parallel	India	Total Sanitation Campaign (TSC)	To measure the effect of the TSC implemented with capacity building support from the World Bank's Water and Sanitation Program on availability of individual household latrines, defecation behaviours, and child health.	--	CONT COV
Pinder, M., et al. (2015)	2-arm parallel	Gambian	--	To assess whether the addition of DDT indoor residual spraying to long-lasting insecticide-treated nets provides a significantly different level of protection against clinical malaria in children or against house entry by vector mosquitoes.	CONT	CONT COV FREQ DUR

Raj, A., et al. (2016)	2-arm parallel	India	CHARM	To evaluate the impact of a gender-equity focused, male-centered family planning intervention on marital contraceptive use and incident pregnancy.	CONT COV	CONT FREQ
Rothera m-Borus, M. J., et al. (2014)	2-arm parallel	South Africa	Project Masihambisane	To evaluate the effect of clinic-based support by HIV-positive Peer Mentors, in addition to standard clinic care, on maternal and infant well being among Women Living with HIV.	CONT FREQ DUR	FREQ
Sangoro , O., et al. (2014)	2-arm parallel	Tanzania	--	To assess whether 15% DEET topical repellent in combination with long-lasting insecticide-treated nets (LLINs) can prevent greater malaria transmission, compared with placebo and LLINs	--	CONT FREQ
Singla, D. R., et al. (2015)	2-arm parallel	Uganda	--	To assess the effects of an integrated, community-based parenting intervention based on social-cognitive learning theory that targeted child development and maternal wellbeing	--	CONT FREQ

Soofi, S., et al. (2012)	2-arm parallel	Pakistan	--	To establish whether community case identification and management of severe pneumonia by oral amoxicillin delivered through community health workers has the potential to reduce the number of infants dying at home.	CONT	CONT COV
Sorensen, G., et al. (2016)	2-arm parallel	India	Mumbai Worksite Tobacco Control Study	To test the effectiveness of a worksite in increasing tobacco use cessation among workers in the manufacturing sector.	--	CONT COV FREQ DUR
Suchdev, P. S., et al. (2012)	2-arm parallel	Kenya	--	To evaluate the effect of community based marketing and distribution of Sprinkles micronutrient powder (MNP) on childhood rates of anaemia and iron and vitamin A deficiency.	COV	CONT FREQ
Tian, M., et al. (2015)	2-arm parallel	China and India	SimCard	To develop and evaluate a simplified cardiovascular management program delivered by community health workers with the aid of a smartphone-based electronic decision support system on patient-reported antihypertensive medication use.	CONT	COV FREQ
Tripathy, P., et al.	2-arm parallel	India	JOHAR	To assess the impact of a community mobilization intervention with women's groups facilitated by	CONT COV	CONT COV

al. (2016)				Accredited Social Health Activists (ASHAs) to improve maternal and newborn health outcomes	FREQ DUR	FREQ
Waiswa , P., et al. (2015)	2-arm parallel	Uganda	Uganda Newborn Survival Study	To assess the effect of a community health worker home visit strategy combined with health facility strengthening on uptake of newborn care seeking, practices and services, and to link the results to national policy and scale-up.	CONT COV	CONT COV FREQ
Wang, P. C., et al. (2015)	3-arm parallel	Southern Zambia	--	To assess the integration of early infant HIV diagnosis with the expanded programme for immunization to determine whether infant and postpartum maternal HIV testing rates increase without harming immunization uptake.	CONT COV	CONT COV
Weiss, S. M., et al. (2015)	2-arm parallel	Zambia	Spear and Shield Intervention	To test the effect of a community-level intervention designed to increase demand for voluntary medical male circumcision among hard-to-reach men.	--	CONT
West, P. A., et al. (2014)	2-arm parallel	Tanzania	--	To investigate whether the combination of indoor residual spraying with bendiocarb and long lasting Insecticide treated mosquito nets (ITNs) provides added protection compared to ITNs alone.	--	CONT COV

Yotebie ng, M., et al. (2015)	3-arm parallel	DR Congo	--	To assess the effect of a short-cut implementation of the Ten Steps to Successful Breastfeeding programme, a key component of the Baby- Friendly Hospital Initiative.	CONT	CONT
Young, S. D., et al. (2015)	2-arm parallel	Peru	Harnessing Online Peer Education (HOPE)	To examine the efficacy of the HOPE social media intervention to increase HIV testing among men who have sex with men (MSM).	--	CONT FREQ
Yousafzai, A. K., et al. (2014)	2 x 2 factorial	Pakistan	--	To investigate the feasibility and effectiveness of the integration of interventions to enhance child development and growth outcomes in the Lady Health Worker (LHW) programme.	--	CONT COV FREQ

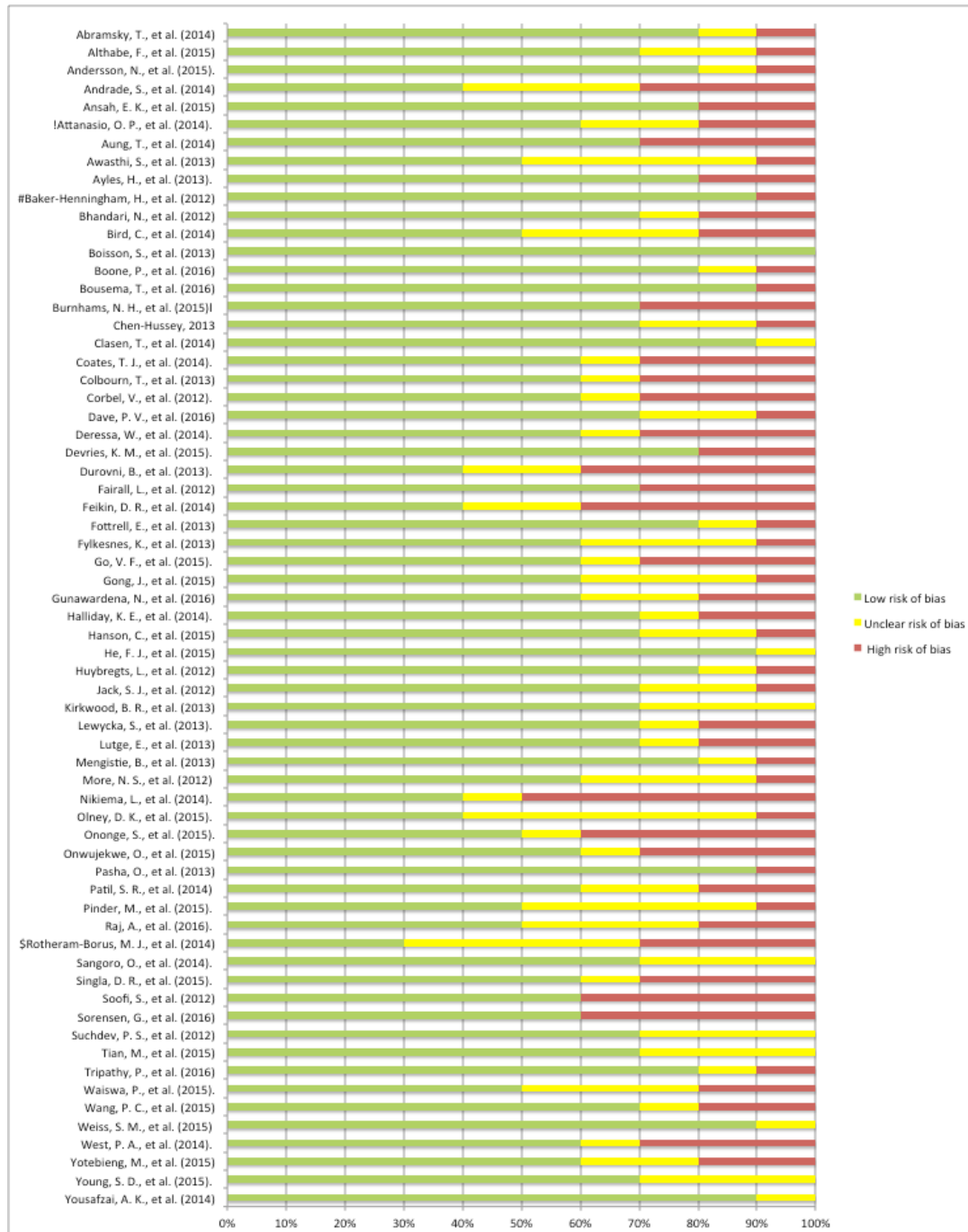
¹ Trial reports include the main study report, and relevant results from peer-reviewed and grey literature if applicable

Abbreviations: CONT – Content, COV –Coverage, FREQ – Frequency, DUR – Duration

Additional file 6 Classification of Public Health Interventions*

Study ID	Health teaching	Case management	Social marketing	Community organizing	Delegated Functions	Counseling	Screening	Referral and follow-up	Consultation	Case finding	Outreach	Collaboration	Advocacy	Coalition building	Policy enforcement	Surveillance	Access to priority interventions
Abramsky, T., et al. (2014)	■		■									■					
Althabe, F., et al. (2015)	■																■
Andersson, N., et al. (2015)				■													
Andrade, S., et al. (2014)	■		■														
Ansah, E. K., et al. (2015)		■			■												
Attanasio, O. P., et al. (2014)	■										■						■
Aung, T., et al. (2014)	■	■						■									
Awasthi, S., et al. (2013)		■														■	■
Ayles, H., et al. (2013)	■	■					■	■		■			■				
Baker-Henningham, H., et al. (2012)	■																
Bhandari, N., et al. (2012)	■	■			■		■		■								■

Additional file 7 Risk of bias graph for individual studies



Additional file 8 Subgroup analyses

a. Registered and published intervention fidelity assessment, by risk of bias category (N=65)

Low risk of bias category¹ (n=36)

	IF Assessment*		Study ID
	(n)	(%)	
Protocol only	1	2.8	61
Trial report only	16	44.4	78,80,83-85,87,88,90,93-95,99,102,104-106
Both protocol and trial report	19	52.8	43,44,46-50,55,58-60,62,63,67,72,73,75,77,82

Uncertain or high risk of bias category² (n=29)

	IF Assessment*		Study ID
	(n)	(%)	
Protocol only	0	0	
Trial report only	13	44.8	79,81,86,89,91,92,96-98,100,101,103,107
Both protocol and trial report	16	55.2	45,51-54,56,57,64-66,68-71,74,76

*IF assessment was judged to be present if an article proposed or used methods to assess 1 or more key fidelity components

¹. Low Risk = 7 to 10 out of 10 criteria scored as low risk of bias n=36

². Uncertain or High Risk = 0 -6 out of 10 criteria scored as low risk of bias n= 28

b. Comprehensiveness of registered and published intervention fidelity assessment*, by risk of bias category

Low risk of bias category¹ (n=36)

Fidelity components	IF planned (Protocol)		Study ID	IF performed (Trial Report)		Study ID
	(n)	(%)		(n)	(%)	
4 dimensions	2	5.6	50,73	2	5.6	43,50
3 dimensions	3	8.3	43,44, 62	13	36.1	44,47,48,58,59,62,63,73,84,85, 90,94,105
2 dimensions	5	13.9	49,58,59,63,75	11	30.6	46,49,72,75,77,80,82,93,99,104,106
1 dimension	10	27.8	46,47,48,55,60,61,67, 72,77,82,	9	25.0	55,60,67,78,83,87,88,95,102
Not done	16	44.4	78,80,83-85,87,88,90, 93-95,99, 102, 104-106	1	2.8	61

Uncertain or high risk of bias category² (n=29)

Fidelity components	IF planned (Protocol)	Study ID	IF performed (Trial Report)	Study ID
---------------------	-----------------------	----------	-----------------------------	----------

	(n)	(%)		(n)	(%)	
4 dimensions	0	0	-	3	10.3	52,68,101
3 dimensions	2	6.9	57,70	5	17.2	45,53,57,74,79
2 dimensions	6	20.7	45,52,53,64,69,74	10	34.5	64- 66,69,71,81,97,98,100,103
1 dimension	8	27.6	51,54,56,65,66,68, 71, 89	11	37.9	51,54,56,70,76, 86, 89, 91, 92, 96,107
Not done	13	44.8	76,79,81,86,91,92, 96, 98,100,101,103,10 7,	0	0	-

*IF assessment was judged to be present if an article proposed or used methods to assess 1 or more key fidelity components

¹. Low Risk = 7 to 10 out of 10 criteria scored as low risk of bias n=36

². Uncertain or High Risk = 0 -6 out of 10 criteria scored as low risk of bias n= 28

REFERENCES

- Abramsky, T., Devries, K., Kiss, L., Nakuti, J., Kyegombe, N., Starmann, E., . . . Watts, C. (2014). Findings from the SASA! Study: a cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda. *BMC Med*, *12*, 122. doi:10.1186/s12916-014-0122-5
- Alshurafa, M., Briel, M., Akl, E. A., Haines, T., Moayyedi, P., Gentles, S. J., . . . Guyatt, G. H. (2012). Inconsistent definitions for intention-to-treat in relation to missing outcome data: systematic review of the methods literature. *PLoS One*, *7*(11), e49163. doi:10.1371/journal.pone.0049163
- Althabe, F., Belizan, J. M., McClure, E. M., Hemingway-Foday, J., Berrueta, M., Mazzoni, A., . . . Buekens, P. M. (2015). A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. *Lancet*, *385*(9968), 629-639. doi:10.1016/s0140-6736(14)61651-2
- Andersson, N., Nava-Aguilera, E., Arostegui, J., Morales-Perez, A., Suazo-Laguna, H., Legorreta-Soberanis, J., . . . Harris, E. (2015). Evidence based community mobilization for dengue prevention in Nicaragua and Mexico (Camino Verde, the Green Way): cluster randomized controlled trial. *BMJ*, *351*, h3267. doi:10.1136/bmj.h3267
- Andrade, S., Lachat, C., Ochoa-Aviles, A., Verstraeten, R., Huybregts, L., Roberfroid, D., . . . Kolsteren, P. (2014). A school-based intervention improves physical fitness in Ecuadorian adolescents: a cluster-randomized controlled trial. *Int J Behav Nutr Phys Act*, *11*, 153. doi:10.1186/s12966-014-0153-5
- Ansah, E. K., Narh-Bana, S., Affran-Bonful, H., Bart-Plange, C., Cundill, B., Gyapong, M., & Whitty, C. J. (2015). The impact of providing rapid diagnostic malaria tests on fever management in the private retail sector in Ghana: a cluster randomized trial. *BMJ*, *350*, h1019. doi:10.1136/bmj.h1019
- Attanasio, O. P., Fernandez, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., & Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial.[Erratum appears in *BMJ*. 2014;349:g6126]. *BMJ*, *349*, g5785. doi:http://dx.doi.org/10.1136/bmj.g5785
- Aung, T., Montagu, D., Su Su Khin, H., Win, Z., San, A. K., & McFarland, W. (2014). Impact of a social franchising program on uptake of oral rehydration solution plus zinc for childhood diarrhea in myanmar: a community-level randomized controlled trial. *J Trop Pediatr*, *60*(3), 189-197. doi:10.1093/tropej/fmt108
- Awasthi, S., Peto, R., Read, S., Richards, S. M., Pande, V., & Bundy, D. (2013). Population deworming every 6 months with albendazole in 1 million pre-school children in North

- India: DEVTA, a cluster-randomised trial. *Lancet*, 381(9876), 1478-1486. doi:10.1016/s0140-6736(12)62126-6
- Ayles, H., Muyoyeta, M., Du Toit, E., Schaap, A., Floyd, S., Simwinga, M., . . . Godfrey-Faussett, P. (2013). Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet*, 382(9899), 1183-1194. doi:10.1016/s0140-6736(13)61131-9
- Baker-Henningham, H., Scott, S., Jones, K., & Walker, S. (2012). Reducing child conduct problems and promoting social skills in a middle-income country: cluster randomised controlled trial. *Br J Psychiatry*, 201, 101-108. doi:10.1192/bjp.bp.111.096834
- Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*: Public Affairs.
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Treatment Fidelity Workgroup of the, N. I. H. B. C. C. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*, 23(5), 443-451. doi:10.1037/0278-6133.23.5.443
- Bhandari, N., Mazumder, S., Taneja, S., Sommerfelt, H., & Strand, T. A. (2012). Effect of implementation of Integrated Management of Neonatal and Childhood Illness (IMNCI) programme on neonatal and infant mortality: cluster randomised controlled trial. *BMJ*, 344, e1634. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22438367>
- Bird, C., Ame, S., Albonico, M., & Bickle, Q. (2014). Do shoes reduce hookworm infection in school-aged children on Pemba Island, Zanzibar? A pragmatic trial. *Trans R Soc Trop Med Hyg*, 108(5), 297-304. doi:10.1093
- Bland, J. M. (2004). Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol*, 4(1), 21. doi:10.1186/1471-2288-4-21
- Boisson, S., Stevenson, M., Shapiro, L., Kumar, V., Singh, L. P., Ward, D., & Clasen, T. (2013). Effect of household-based drinking water chlorination on diarrhoea among children under five in Orissa, India: a double-blind randomised placebo-controlled trial. *PLoS Med*, 10(8), e1001497. doi:10.1371/journal.pmed.1001497
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V., & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *Br Med J*, 333, 346-349. doi:doi:10.1136/bmj.333.7563.346
- Boone, P., Elbourne, D., Fazio, I., Fernandes, S., Frost, C., Jayanty, C., . . . Walker, P. R. (2016). Effects of community health interventions on under-5 mortality in rural Guinea-Bissau (EPICS): a cluster-randomised controlled trial. *Lancet Glob Health*, 4(5), e328-335. doi:10.1016/S2214-109X(16)30048-1

- Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent, 71(s1)*, S52–63.
- Bousema, T., Stresman, G., Baidjoe, A. Y., Bradley, J., Knight, P., Stone, W., . . . Cox, J. (2016). The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial. *PLoS Med, 13(4)*, e1001993. doi:10.1371/journal.pmed.1001993
- Brierley, G., Brabyn, S., Torgerson, D., & Watson, J. (2012). Bias in recruitment to cluster randomized trials: a review of recent publications. *J Eval Clin Pract., 18(4)*, 878–886.
- Burnhams, N. H., London, L., Laubscher, R., Nel, E., & Parry, C. (2015). Results of a cluster randomised controlled trial to reduce risky use of alcohol, alcohol-related HIV risks and improve help-seeking behaviour among safety and security employees in the Western Cape, South Africa. *Subst Abuse Treat Prev Policy, 10*, 18. doi:10.1186/s13011-015-0014-5
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ, 321(7262)*, 694–696.
- Campbell, M. J., Donner, A., & Klar, N. (2007). Developments in cluster randomized trials and Statistics in Medicine. *Stat Med, 26(1)*, 2-19. doi:10.1002/sim.2731
- Campbell, M. J., Donner, A., & Klar, N. (2007). Developments in cluster randomized trials and statistics in medicine. *Stat Med, 26(1)*, 2–19.
- Campbell, M. K., Elbourne, D. R., Altman, D. G., & group, C. (2004). CONSORT statement: extension to cluster randomised trials. *BMJ, 328(7441)*, 702-708. doi:10.1136/bmj.328.7441.702
- Campbell, M. K., Mollison, J., Steen, N., Grimshaw, J. M., & Eccles, M. (2000). Analysis of cluster randomized trials in primary care: a practical approach. *Fam Pract, 17*, 192–196.
- Campbell, M. K., Piaggio, G., Elbourne, D. R., Altman, D. G., & Group, C. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ, 345*, e5661. doi:10.1136/bmj.e5661
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci, 2*, 40. doi:10.1186/1748-5908-2-40
- Chen-Hussey, V., Carneiro, I., Keomanila, H., Gray, R., Bannavong, S., Phanalasy, S., & Lindsay, S. W. (2013). Can topical insect repellents reduce malaria? A cluster-randomised controlled trial of the insect repellent N,N-diethyl-m-toluamide (DEET) in Lao PDR. *PLoS One, 8(8)*, e70664. doi:10.1371/journal.pone.0070664
- Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., . . . Schmidt, W. P. (2014). Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted

helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *Lancet Glob Health*, 2(11), e645-653. doi:10.1016/S2214-109X(14)70307-9

- Coates, T. J., Kulich, M., Celentano, D. D., Zelaya, C. E., Chariyalertsak, S., Chingono, A., . . . Eshleman, S. H. (2014). Effect of community-based voluntary counselling and testing on HIV incidence and social and behavioural outcomes (NIMH Project Accept; HPTN 043): a cluster-randomised trial. *Lancet Glob Health*, 2(5), e267-277. doi:10.1016/s2214-109x(14)70032-4
- Colbourn, T., Nambiar, B., Bondo, A., Makwenda, C., Tsetekani, E., Makonda-Ridley, A., . . . Costello, A. (2013). Effects of quality improvement in health facilities and community mobilization through women's groups on maternal, neonatal and perinatal mortality in three districts of Malawi: MaiKhanda, a cluster randomized controlled effectiveness trial. *Int Health*, 5(3), 180-195. doi:10.1093/inthealth/iht011
- Corbel, V., Akogbeto, M., Damien, G. B., Djenontin, A., Chandre, F., Rogier, C., . . . Henry, M. C. (2012). Combination of malaria vector control interventions in pyrethroid resistance area in Benin: a cluster randomised controlled trial. *Lancet Infect Dis*, 12(8), 617-626. doi:10.1016/s1473-3099(12)70081-6
- Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., & Thompson, S. (2012). Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health*, 66(12), 1182–1186.
- Dave, P. V., Shah, A. N., Nimavat, P. B., Modi, B. B., Pujara, K. R., Patel, P., . . . Kumar, A. M. (2016). Direct Observation of Treatment Provided by a Family Member as Compared to Non-Family Member among Children with New Tuberculosis: A Pragmatic, Non-Inferiority, Cluster-Randomized Trial in Gujarat, India. *PLoS One*, 11(2), e0148488. doi:10.1371/journal.pone.0148488
- Deressa, W., Yihdego, Y. Y., Kebede, Z., Batisso, E., Tekalegne, A., & Dagne, G. A. (2014). Effect of combining mosquito repellent and insecticide treated net on malaria prevalence in Southern Ethiopia: a cluster-randomised trial. *Parasit Vectors*, 7, 132. doi:10.1186/1756-3305-7-132
- Devries, K. M., Knight, L., Child, J. C., Mirembe, A., Nakuti, J., Jones, R., . . . Naker, D. (2015). The Good School Toolkit for reducing physical violence from school staff to primary school students: a cluster-randomised controlled trial in Uganda. *Lancet Glob Health*, 3(7), e378-386. doi:10.1016/s2214-109x(15)00060-1
- Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4, 3895-3962.

- Durovni, B., Saraceni, V., Moulton, L. H., Pacheco, A. G., Cavalcante, S. C., King, B. S., . . . Golub, J. E. (2013). Effect of improved tuberculosis screening and isoniazid preventive therapy on incidence of tuberculosis and death in patients with HIV in clinics in Rio de Janeiro, Brazil: a stepped wedge, cluster-randomised trial. *Lancet Infect Dis*, *13*(10), 852-858. doi:10.1016/s1473-3099(13)70187-7
- Dusenbury, L., Brannigan, R., & Falco, M., Hansen WB. . (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res*, *18*(2), 237–256.
- Eldridge, S., Ashby, D., Bennett, C., Wakelin, M., & Feder, G. (2008). Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ*, *336*(7649), 876–880.
- Eldridge, S., Campbell, M., Campbell, M., Dahota, A., Giraudeau, B., Higgins, J. P., . . . Siegfried, N. (2016). Revised Cochrane risk of bias tool for randomized trials (RoB 2.0): Additional considerations for cluster-randomized trials. In M. J. In: Chandler J, Boutron I, Welch V (editors). (Ed.), *Cochrane Methods.: Cochrane Database of Systematic Reviews 2016, Issue 10* (Suppl 1).
- Evidence-Based, Medicine Working, & Group. (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*, *268*(17), 2420.
- Fairall, L., Bachmann, M. O., Lombard, C., Timmerman, V., Uebel, K., Zwarenstein, M., . . . Bateman, E. (2012). Task shifting of antiretroviral treatment from doctors to primary-care nurses in South Africa (STRETCH): a pragmatic, parallel, cluster-randomised trial. *Lancet*, *380*(9845), 889-898. doi:10.1016/s0140-6736(12)60730-2
- Feikin, D. R., Bigogo, G., Audi, A., Pals, S. L., Aol, G., Mbakaya, C., . . . Larson, C. P. (2014). Village-randomized clinical trial of home distribution of zinc for treatment of childhood diarrhea in rural Western kenya. *PLoS One*, *9*(5), e94436. doi:10.1371/journal.pone.0094436
- Fottrell, E., Azad, K., Kuddus, A., Younes, L., Shaha, S., Nahar, T., . . . Houweling, T. A. (2013). The effect of increased coverage of participatory women's groups on neonatal mortality in Bangladesh: A cluster randomized trial. *JAMA Pediatr*, *167*(9), 816-825. doi:10.1001/jamapediatrics.2013.2534
- Fylkesnes, K., Sandoy, I. F., Jurgensen, M., Chipimo, P. J., Mwangala, S., & Michelo, C. (2013). Strong effects of home-based voluntary HIV counselling and testing on acceptance and equity: a cluster randomised trial in Zambia. *Soc Sci Med*, *86*, 9-16. doi:10.1016/j.socscimed.2013.02.036
- Gasparian, A. Y., Ayvazyan, L., & Kitas, G. D. (2013). Multidisciplinary bibliographic databases. *J Korean Med Sci*, *28*(9), 1270-1275. doi:10.3346/jkms.2013.28.9.1270

- Go, V. F., Frangakis, C., Minh, N. L., Latkin, C., Ha, T. V., Mo, T. T., . . . Quan, V. M. (2015). Efficacy of a Multi-level Intervention to Reduce Injecting and Sexual Risk Behaviors among HIV-Infected People Who Inject Drugs in Vietnam: A Four-Arm Randomized Controlled Trial. *PLoS One*, *10*(5), e0125909. doi:10.1371/journal.pone.0125909
- Gong, J., Chen, X., & Li, S. (2015). Efficacy of a Community-Based Physical Activity Program KM2H2 for Stroke and Heart Attack Prevention among Senior Hypertensive Patients: A Cluster Randomized Controlled Phase-II Trial. *PLoS One*, *10*(10), e0139442. doi:10.1371/journal.pone.0139442
- Gunawardena, N., Kurotani, K., Indrawansa, S., Nonaka, D., Mizoue, T., & Samarasinghe, D. (2016). School-based intervention to enable school children to act as change agents on weight, physical activity and diet of their mothers: a cluster randomized controlled trial. *Int J Behav Nutr Phys Act*, *13*, 45. doi:10.1186/s12966-016-0369-7
- Halliday, K. E., Okello, G., Turner, E. L., Njagi, K., McHaro, C., Kengo, J., . . . Brooker, S. J. (2014). Impact of intermittent screening and treatment for malaria among school children in Kenya: a cluster randomised trial. *PLoS Med*, *11*(1), e1001594. doi:10.1371/journal.pmed.1001594
- Handlos, L. N., Chakraborty, H., & Sen, P. K. (2009). Evaluation of cluster-randomized trials on maternal and child health research in developing countries. *Trop Med Int Health*, *14*(8), 947–956.
- Hanson, C., Manzi, F., Mkumbo, E., Shirima, K., Penfold, S., Hill, Z., . . . Schellenberg, J. (2015). Effectiveness of a Home-Based Counselling Strategy on Neonatal Care and Survival: A Cluster-Randomised Trial in Six Districts of Rural Southern Tanzania. *PLoS Med*, *12*(9), e1001881. doi:10.1371/journal.pmed.1001881
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci*, *5*, 67. doi:10.1186/1748-5908-5-67
- Hasson, H., Blomberg, S., & Dunér, A. (2012). Fidelity and moderating factors in complex interventions: a case study of a continuum of care program for frail elderly people in health and social care. *Implement Sci*, *7*(23), 1–11.
- Hawe, P. (2015). Lessons from complex interventions to improve health. *Annu Rev Public Health*, *36*, 307–323.
- Hayes, R., & Moulton, L. (2009). *Cluster randomised trials*.
- Hayes, R. J., Alexander, N. D., Bennett, S., & Cousens, S. N. (2000). Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Methods Med Res*, *9*(2), 95–116.

- He, F. J., Wu, Y., Feng, X. X., Ma, J., Ma, Y., Wang, H., . . . MacGregor, G. A. (2015). School based education programme to reduce salt intake in children and their families (School-EduSalt): cluster randomised controlled trial. *BMJ*, *350*, h770. doi:10.1136/bmj.h770
- Higgins, J., Sterne, J., Savović, J., Page, M., Hróbjartsson, A., Boutron, I., . . . Eldridge, S. (2016). Revised Cochrane risk of bias tool for randomized trials (RoB 2.0). In M. J. In: Chandler J, Boutron I, Welch V (editors). (Ed.), *Cochrane Methods*.: Cochrane Database of Systematic Reviews 2016, Issue 10 (Suppl 1).
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, d5928.
- Higgins, J. P., Sterne, J. A., Savovic, J., Page, M. J., Hróbjartsson, A., Boutron, I., . . . Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. *Cochrane database of systematic reviews*, *10*(Suppl 1), 29-31.
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, *348*, g1687. doi:10.1136/bmj.g1687
- Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ*, *319*(7211), 670-674. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10480822>
- Huybregts, L., Hougbe, F., Salpeteur, C., Brown, R., Roberfroid, D., Ait-Aissa, M., & Kolsteren, P. (2012). The effect of adding ready-to-use supplementary food to a general food distribution on child nutritional status and morbidity: a cluster-randomized controlled trial. *PLoS Med*, *9*(9), e1001313. doi:10.1371/journal.pmed.1001313
- Ibrahim, S., & Sidani, S. (2015). Fidelity of intervention implementation: a review of instruments. *Health*, *2015*, 1687.
- Isaakidis, P., & Ioannidis, J. P. (2003). Evaluation of cluster randomized controlled trials in sub-Saharan Africa. *Am J Epidemiol*, *158*(9), 921–926.
- Ivers, N. M., Taljaard, M., Dixon, S., Bennett, C., McRae, A., Taleban, J., . . . Donner, A. (2011). Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. *BMJ*, *343*, d5886. doi:10.1136/bmj.d5886
- Jack, S. J., Ou, K., Chea, M., Chhin, L., Devenish, R., Dunbar, M., . . . Gibson, R. S. (2012). Effect of micronutrient sprinkles on reducing anemia a cluster-randomized effectiveness trial. *Archives of Pediatrics and Adolescent Medicine*, *166*(9), 842-850. doi:http://dx.doi.org/10.1001/archpediatrics.2012.1003

- Keller, L. O., Strohschein, S., Lia-Hoagberg, B., & Schaffer, M. A. (2004). Population-based public health interventions: practice-based and evidence-supported. *Part I. Public Health Nurs*, *21*(5), 453–468.
- Kirkwood, B. R., Manu, A., ten Asbroek, A. H., Soremekun, S., Weobong, B., Gyan, T., . . . Hill, Z. (2013). Effect of the Newhints home-visits intervention on neonatal mortality rate and care practices in Ghana: a cluster randomised controlled trial. *Lancet*, *381*(9884), 2184–2192. doi:10.1016/S0140-6736(13)60095-1
- Lewycka, S., Mwansambo, C., Rosato, M., Kazembe, P., Phiri, T., Mganga, A., . . . Costello, A. (2013). Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial. *Lancet*, *381*(9879), 1721–1735. doi:10.1016/S0140-6736(12)61959-X
- Lutge, E., Lewin, S., & Volmink, J. (2014). Economic support to improve tuberculosis treatment outcomes in South Africa: a qualitative process evaluation of a cluster randomized controlled trial. *Trials*, *15*, 236. doi:10.1186/1745-6215-15-236
- Medical Research, & Council. (2008). *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council.
- Mengistie, B., Berhane, Y., & Worku, A. (2013). Household water chlorination reduces incidence of diarrhea among under-five children in rural Ethiopia: a cluster randomized controlled trial. *PLoS One*, *8*(10), e77887. doi:10.1371/journal.pone.0077887
- Moberg, J., & Kramer, M. (2015). A brief history of the cluster randomised trial design. *J R Soc Med*, *108*(5), 192–198.
- Moher D, Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6* (7), e1000097.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., & D.G., A. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*, *63*(8), e1–e37.
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, *350*, h1258.
- More, N. S., Bapat, U., Das, S., Alcock, G., Patil, S., Porel, M., . . . Osrin, D. (2012). Community mobilization in Mumbai slums to improve perinatal care and outcomes: a cluster randomized controlled trial. *PLoS Med*, *9*(7), e1001257. doi:10.1371/journal.pmed.1001257

- Morello, R. T., Barker, A. L., Ayton, D. R., Landgren, F., Kamar, J., Hill, K. D., . . . Stoelwinder, J. (2017). Implementation fidelity of a nurse-led falls prevention program in acute hospitals during the 6-PACK trial. *BMC Health Serv Res*, *17*(1), 383. doi:10.1186/s12913-017-2315-z
- Morris, S. S., Ranson, M. K., Sinha, T., & Mills, A. J. (2007). Measuring improved targeting of health interventions to the poor in the context of a community-randomised trial in rural India. *Contemp Clin Trials*, *28*(4), 382–390.
- Nikiema, L., Huybregts, L., Kolsteren, P., Lanou, H., Tiendrebeogo, S., Bouckaert, K., . . . Roberfroid, D. (2014). Treating moderate acute malnutrition in first-line health services: an effectiveness cluster-randomized trial in Burkina Faso. *Am J Clin Nutr*, *100*(1), 241-249. doi:10.3945/ajcn.113.072538
- O'Donnell, C. L. (2008). Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K–12 Curriculum Intervention Research. *Review of Educational Research*, *78*(1), 33-84. doi:DOI: 10.3102/0034654307313793
- Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ*, *332*(7583), 413–416.
- Olney, D. K., Pedehombga, A., Ruel, M. T., & Dillon, A. (2015). A 2-year integrated agriculture and nutrition and health behavior change communication program targeted to women in Burkina Faso reduces anemia, wasting, and diarrhea in children 3-12.9 months of age at baseline: a cluster-randomized controlled trial. *J Nutr*, *145*(6), 1317-1324. doi:10.3945/jn.114.203539
- Ononge, S., Campbell, O. M., Kaharuzza, F., Lewis, J. J., Fielding, K., & Mirembe, F. (2015). Effectiveness and safety of misoprostol distributed to antenatal women to prevent postpartum haemorrhage after child-births: a stepped-wedge cluster-randomized trial. *BMC Pregnancy Childbirth*, *15*, 315. doi:10.1186/s12884-015-0750-6
- Onwujekwe, O., Mangham-Jefferies, L., Cundill, B., Alexander, N., Langham, J., Ibe, O., . . . Wiseman, V. (2015). Effectiveness of Provider and Community Interventions to Improve Treatment of Uncomplicated Malaria in Nigeria: A Cluster Randomized Controlled Trial. *PLoS One*, *10*(8), e0133832. doi:10.1371/journal.pone.0133832
- Osrin, D., Azad, K., Fernandez, A., Manandhar, D. S., Mwansambo, C. W., Tripathy, P., & Costello, A. M. (2009). Ethical challenges in cluster randomized controlled trials: experiences from public health interventions in Africa and Asia. *Bull World Health Org*, *87*(10), 772–779.
- Pasha, O., McClure, E. M., Wright, L. L., Saleem, S., Goudar, S. S., Chomba, E., . . . Goldenberg, R. L. (2013). A combined community- and facility-based approach to improve pregnancy outcomes in low-resource settings: a Global Network cluster randomized trial. *BMC Med*, *11*, 215. doi:10.1186/1741-7015-11-215

- Patil, S. R., Arnold, B. F., Salvatore, A. L., Briceno, B., Ganguly, S., Colford, J. M., Jr., & Gertler, P. J. (2014). The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Med*, *11*(8), e1001709. doi:10.1371/journal.pmed.1001709
- Perez, D., Van der Stuyft, P., Zabala, M. C., Castro, M., & Lefevre, P. (2016). A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. *Implement Sci*, *11*(1), 91. doi:10.1186/s13012-016-0457-8
- Perez, M. C., Minoyan, N., Ridde, V., Sylvestre, M. P., & Johri, M. (2016). Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review protocol. *Syst Rev*, *5*(1), 177. doi:10.1186/s13643-016-0351-0
- Pinder, M., Jawara, M., Jarju, L. B., Salami, K., Jeffries, D., Adiamoh, M., . . . Lindsay, S. W. (2015). Efficacy of indoor residual spraying with dichlorodiphenyltrichloroethane against malaria in Gambian communities with high usage of long-lasting insecticidal mosquito nets: a cluster-randomised controlled trial. *Lancet*, *385*(9976), 1436-1446. doi:10.1016/s0140-6736(14)61007-2
- Puffer, S., Torgerson, D., & Watson, J. (2003). Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ*, *327*(7418), 785-789.
- Raj, A., Ghule, M., Ritter, J., Battala, M., Gajanan, V., Nair, S., . . . Saggurti, N. (2016). Cluster Randomized Controlled Trial Evaluation of a Gender Equity and Family Planning Intervention for Married Men and Couples in Rural India. *PLoS One*, *11*(5), e0153190. doi:10.1371/journal.pone.0153190
- Ranson, M. K., Sinha, T., Morris, S. S., & Mills, A. J. (2006). CRTs—cluster randomized trials or “courting real troubles”: challenges of running a CRT in rural Gujarat, India. *Can J Public Health*, *97*(1), 72.
- Richards, D. A., & Hallberg, I. R. (2015). Complex interventions in health: an overview of research methods. *London & New York: Routledge*.
- Rotheram-Borus, M. J., Richter, L. M., van Heerden, A., van Rooyen, H., Tomlinson, M., Harwood, J. M., . . . Stein, A. (2014). A cluster randomized controlled trial evaluating the efficacy of peer mentors to support South African women living with HIV and their infants. *PLoS One*, *9*(1), e84867. doi:10.1371/journal.pone.0084867
- Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*, *56*, 119-127. doi:doi:10.1136/jech.56.2.119
- Rychetnik, L., Hawe, P., Waters, E., Barratt, A., & Frommer, M. (2004). A glossary for evidence based public health. *J Epidemiol Community Health*, *58*(7), 538-545.

- Sangoro, O., Turner, E., Simfukwe, E., Miller, J. E., & Moore, S. J. (2014). A cluster-randomized controlled trial to assess the effectiveness of using 15% DEET topical repellent with long-lasting insecticidal nets (LLINs) compared to a placebo lotion on malaria transmission. *Malar J*, *13*, 324. doi:10.1186/1475-2875-13-324
- Sanson-Fisher, R. W., D'Este, C. A., Carey, M. L., Noble, N., & Paul, C. L. (2014). Evaluation of systems-oriented public health interventions: alternative research designs. *Annu Rev Public Health*, *35*, 9–27.
- Schinckus, L., Van den Broucke, S., Housiaux, M., & Diabetes Literacy, C. (2014). Assessment of implementation fidelity in diabetes self-management education programs: a systematic review. *Patient Educ Couns*, *96*(1), 13-21. doi:10.1016/j.pec.2014.04.002
- Schneeweiss, S. (2006). Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*, *15*(5), 291–303.
- Singla, D. R., Kumbakumba, E., & Aboud, F. E. (2015). Effects of a parenting intervention to address maternal psychological wellbeing and child development and growth in rural Uganda: a community-based, cluster randomised trial. *Lancet Glob Health*, *3*(8), e458-469. doi:10.1016/s2214-109x(15)00099-6
- Soofi, S., Ahmed, S., Fox, M. P., MacLeod, W. B., Thea, D. M., Qazi, S. A., & Bhutta, Z. A. (2012). Effectiveness of community case management of severe pneumonia with oral amoxicillin in children aged 2-59 months in Matiari district, rural Pakistan: a cluster-randomised controlled trial. *Lancet*, *379*(9817), 729-737. doi:10.1016/s0140-6736(11)61714-5
- Sorensen, G., Pednekar, M., Cordeira, L. S., Pawar, P., Nagler, E. M., Stoddard, A. M., . . . Gupta, P. C. (2016). Effects of a worksite tobacco control intervention in India: the Mumbai worksite tobacco control study, a cluster-randomised trial. *Tob Control*. doi:10.1136/tobaccocontrol-2015-052671
- Suchdev, P. S., Ruth, L. J., Woodruff, B. A., Mbakaya, C., Mandava, U., Flores-Ayala, R., . . . Quick, R. (2012). Selling Sprinkles micronutrient powder reduces anemia, iron deficiency, and vitamin A deficiency in young children in Western Kenya: a cluster-randomized controlled trial. *Am J Clin Nutr*, *95*(5), 1223-1230. doi:10.3945/ajcn.111.030072
- Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., & Ye, C. (2013). Debono VB. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol*, *13*(1), 92.
- The World, & Bank. Country and lending groups. Data & statistics: country classification (2016). In. Washington, D.C: . Available at: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519> Accessed 3 July,2016.

- Tian, M., Ajay, V. S., Dunzhu, D., Hameed, S. S., Li, X., Liu, Z., . . . Yan, L. L. (2015). A Cluster-Randomized, Controlled Trial of a Simplified Multifaceted Management Program for Individuals at High Cardiovascular Risk (SimCard Trial) in Rural Tibet, China, and Haryana, India. *Circulation*, *132*(9), 815-824. doi:10.1161/circulationaha.115.015373
- Tripathy, P., Nair, N., Sinha, R., Rath, S., Gope, R. K., Rath, S., . . . Prost, A. (2016). Effect of participatory women's groups facilitated by Accredited Social Health Activists on birth outcomes in rural eastern India: a cluster-randomised controlled trial. *Lancet Glob Health*, *4*(2), e119-128. doi:10.1016/S2214-109X(15)00287-9
- Upshur, R. E. (2002). Principles for the justification of public health intervention. *Can J Public Health*, *93*(2), 101–103.
- Waiswa, P., Pariyo, G., Kallander, K., Akuze, J., Namazzi, G., Ekirapa-Kiracho, E., . . . Peterson, S. (2015). Effect of the Uganda Newborn Study on care-seeking and care practices: a cluster-randomised controlled trial. *Glob Health Action*, *8*, 24584. doi:10.3402/gha.v8.24584
- Wang, P. C., Mwangi, A., Moberley, S., Brockman, B. J., Connor, A. L., Kalesha-Masumbu, P., . . . McCarthy, E. (2015). A Cluster Randomised Trial on the Impact of Integrating Early Infant HIV Diagnosis with the Expanded Programme on Immunization on Immunization and HIV Testing Rates in Rural Health Facilities in Southern Zambia. *PLoS One*, *10*(10), e0141455. doi:10.1371/journal.pone.0141455
- Weiss, S. M., Zulu, R., Jones, D. L., Redding, C. A., Cook, R., & Chitalu, N. (2015). The Spear and Shield intervention to increase the availability and acceptability of voluntary medical male circumcision in Zambia: a cluster randomised controlled trial. *Lancet HIV*, *2*(5), e181-189. doi:10.1016/s2352-3018(15)00042-9
- West, P. A., Protopopoff, N., Wright, A., Kivaju, Z., Tigererwa, R., Mosha, F. W., . . . Kleinschmidt, I. (2014). Indoor residual spraying in combination with insecticide-treated nets compared to insecticide-treated nets alone for protection against malaria: a cluster randomised trial in Tanzania. *PLoS Med*, *11*(4), e1001630. doi:10.1371/journal.pmed.1001630
- World Health, & Organization. (2016). *WHO international clinical trials registry platform (ICRP). New standards for registration of human medical research. Trial registration data set (version 1.2.1)*. Retrieved from
- Yotebieng, M., Labbok, M., Soeters, H. M., Chalachala, J. L., Lapika, B., Vitta, B. S., & Behets, F. (2015). Ten Steps to Successful Breastfeeding programme to promote early initiation and exclusive breastfeeding in DR Congo: a cluster-randomised controlled trial. *Lancet Glob Health*, *3*(9), e546-555. doi:10.1016/s2214-109x(15)00012-1
- Young, S. D., Cumberland, W. G., Nianogo, R., Menacho, L. A., Galea, J. T., & Coates, T. (2015). The HOPE social media intervention for global HIV prevention in Peru: a cluster randomised controlled trial. *Lancet HIV*, *2*(1), e27-32. doi:10.1016/s2352-3018(14)00006-

Yousafzai, A. K., Rasheed, M. A., Rizvi, A., Armstrong, R., & Bhutta, Z. A. (2014). Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial. *Lancet*, *384*(9950), 1282-1293. doi:10.1016/s0140-6736(14)60455-4

3.2 Article 2 Development of an mHealth Behavior Change Communication Strategy: A case-study from rural Uttar Pradesh in India

Pérez, M. C¹., Singh, R²., Chandra, D³., Ridde, V⁴., Seth, A⁵. & Johri, M¹

STATUT : ARTICLE PUBLIÉ

Pérez, M. C., Singh, R., Chandra, D., Ridde, V., Seth, A. & Johri, M. (2020). Development of an mHealth Behavior Change Communication Strategy: A case-study from rural Uttar Pradesh in India. Paper presented at the Proceedings of the ACM COMPASS conference (COMPASS'20). ACM, Guayaquil, Guayas, Equateur, 4 pages. <https://doi.org/10.1145/3378393.3402505>

AUTHOR AFFILIATIONS

1 Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal, Québec, Canada.

2 Gram Vaani Community Media Pvt. Ltd., New Delhi, India

3 Independent Consultant, Tika Vaani, New Delhi, India

4 French Institute for Research on Sustainable Development (IRD), Université de Paris, Paris, France

5 Department of Computer Science, Indian Institute of Technology, Gram Vaani, New Delhi, India

Contribution des auteurs :

Sous la supervision de Dre Mira Johri, Myriam Cielo a réalisé la synthèse documentaire, a rédigé et a intégré les différentes sections de l'article. Rohit Singh, Dinesh Pant et Dr Aaditeshwar Seth ont contribué avec des informations pertinentes pour compléter l'élaboration de l'article, et, avec Dr Valéry Ridde, ont apporté un regard critique à l'ensemble de l'article.

Dre Mira Johri a conçu l'étude, a supervisé l'ensemble des étapes et a apporté un regard critique et constructif à l'ensemble de l'article.

Tous les auteurs ont lu et approuvé le manuscrit final pour soumission.

ABSTRACT

Mobile health interventions are an innovative way to improve health outcomes and may play a powerful role in mitigating health disparities. However, their use poses special challenges and few articles have reported specifically on digital technology interventions for vulnerable populations. This article shares our experience from the Tika Vaani (“vaccine voice”) Intervention which uses a combined face-to-face and mHealth strategy to educate and empower beneficiaries to improve immunization uptake and child health for a poor, low-literate population in rural Uttar Pradesh, India. Based on the mERA checklist, a guide to improve the completeness of reporting mHealth interventions, we provide information about the process of development, implementation, and lessons for scaling up the Tika Vaani intervention. This study contributes to the literature to improve reporting on mHealth interventions and provide researchers with key points and actions to take during intervention development to serve hard-to-reach communities and improve health outcomes.

Author Keywords

mHealth intervention; Child Vaccination; Health Promotion; Behavior Change; Vulnerable Populations; Developing countries; Interactive Voice Response (IVR) systems and Pilot Trial.

CSS Concepts

Human-centered computing~Interaction design, Human-centered computing~Human computer interaction (HCI)

ACM Reference format:

Myriam Cielo Pérez, Rohit Singh, Dinesh Chandra, Valéry Ridde, Aaditeshwar Seth, Mira Johri. 2020. Development of an mHealth Behavior Change Communication Strategy: A case-study from rural Uttar Pradesh in India. Paper presented at the Proceedings of the ACM COMPASS conference (COMPASS’20). ACM, Guayaquil, Guayas, Equateur, 4 pages. <https://doi.org/10.1145/3378393.3402505>

INTRODUCTION

Immunization is one the most successful and cost-effective public health interventions. At present immunization saves two and three million lives each year by protecting children against serious illness, disability and death from vaccine-preventable diseases like measles, polio, diphtheria-pertussis-tetanus and polio (Andre et al., 2008; McGovern & Canning, 2015). Vaccine-preventable disease reduce health care costs, decreasing absenteeism from work or school, and contributing to social and economic development (Bloom, Canning, & Weston, 2017). However, as reported by the World Health Organization (WHO) in 2018, about 19.4 million children in the world are not yet fully vaccinated. Around 60% (11.7million) of these children live in low- and middle-income countries and 29% (2.6million) in India (Organization, 2018). Achieving high and equitable coverage requires targeted actions to ensure access to vaccination, especially for vulnerable populations. Innovative new approaches are required to reach those populations.

Digital technology is an innovative way to improve health outcomes and may play a powerful role in mitigating health disparities (National Academies of Sciences & Medicine, 2016). Mobile health (mHealth) is a subset of digital technologies for health (eHealth) and is defined as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” (Organization, 2011). mHealth interventions have been used in a variety of ways to promote behavior change, improve access to health services, clinical diagnosis, treatment adherence, promote health education, facilitate training and health worker decision support, data collection, registration procedures and maintenance of patient records (Hall, Fottrell, Wilkinson, & Byass, 2014; Marcolino et al., 2018; Wallis, Blessing, Dalwai, & Shin, 2017). A recent overview of systematic reviews concluded that, although the evidence base is small, there is positive and promising evidence that eHealth strategies can improve vaccination uptake, immunization program efficiency, and advance health literacy concerning immunisations (Dumit, Novillo-Ortiz, Contreras, Velandia, & Danovaro-Holliday, 2018). The article concluded that there was a need for more research, particularly in developing country settings, detailing benefits as well as challenges and barriers (Dumit et al., 2018).

Although mHealth interventions offer an important opportunity to reduce health disparities (Peek, 2017), their use poses special challenges and few articles have reported specifically on digital technology interventions for vulnerable populations. This article shares our experience from the Tika Vaani Intervention which uses a combined face-to-face and mHealth strategy to improve child vaccination and health in a low-literacy, resource-poor, rural population in India. The intervention delivers entertaining educational content (edutainment) and vaccination reminder messages through small face-to-face meetings and via mobile phones using strategies adapted to disadvantaged populations.

We recently conducted an early-stage assessment of this model through a cluster-randomised pilot trial in rural Uttar Pradesh, India from January to September 2018. Villages were randomly assigned to either the intervention or control group. The main outcomes were to test feasibility for a larger study, while secondary endpoints tested changes in proxy outcomes, intervention fidelity and user perspectives. Evaluation results have been reported elsewhere (Johri et al., 2020; Pérez et al., 2020). Briefly, all predetermined standards for feasibility were met. Intervention uptake was very high. Overall participation in one or more new interventions was 94% (173/184), 82% (144/184) for the face-to-face strategy and 67% (124/184) for the mHealth strategy. Results showed significantly higher basic health knowledge among the intervention group, thereby supporting the intervention theory of change. The evaluation concluded that the intervention model could be a powerful strategy to deliver information and can be widely scaled to strengthen primary health care and vaccine uptake in poor, low-literate populations.

As recently recognized by the WHO mHealth Technical Evidence Review Group, the evaluation and reporting of mHealth and Information and Communication Technologies (ICT) interventions is currently weak (Agarwal et al., 2016). Presentation of mHealth evidence requires a unique lens, encompassing the need to define what the mHealth intervention is (content), where it is being implemented (context), and how it was implemented (technical features) (Agarwal et al., 2016). Offered as a complement to the pilot trial evaluation (Johri et al., 2020), the purpose of this article is to provide this critical information required to support understanding and replication of the mHealth strategy used in the Tika Vaani intervention.

METHODS

mERA checklist

To improve the completeness of reporting mHealth interventions, the World Health Organization mHealth Technical Evidence Review Group developed the mHealth evidence reporting and assessment (mERA) checklist (Agarwal et al., 2016). The mERA checklist consist of 16 core mHealth items designed and piloted by experts to improve the quality and consistency of mHealth intervention reporting. The guidelines comprise 16 items to inform the context of the study, the content of the mHealth intervention, the process of developing and adapting the intervention, performance of the intervention in practice, and additional considerations related to coherence and fit within the health system, scale up and sustainability. The authors argue that uptake of these guidelines (used in conjunction with an appropriate study-design checklist) will improve evidence on the effectiveness of such approaches to support replication of the interventions (Agarwal et al., 2016). This paper provides the information required by the mERA checklist describing all mERA core items following these five questions: What is the context? What is the Tika Vaani intervention? How was the intervention developed? How did it perform in practice? What are the lessons for scale-up and replication?

WHAT IS THE CONTEXT?

Setting

The study took place in Hardoi a rural district of the state of Uttar Pradesh, India. Uttar Pradesh is the most populous state in India and had the largest numbers of under-5 deaths, contributing 27.1% (325,500) of the national total (1,200,998) largely due to causes that could be potentially addressed by vaccines (Liu et al., 2019). Despite several programs and efforts by the government to reduce infant and child mortality, Uttar Pradesh is one of the districts with the largest immunization coverage gap by disadvantaged population subgroups (the poorest, less educated parents, the lower social caste and by religion beliefs). A recent study, based on the most recent round of the National Family Health Survey (NFHS, 2015, 16), found that in India, among scheduled castes and scheduled tribes (representing the most downtrodden and vulnerable ethnic groups in Indian society), only 55.8% and 63.2% of their children, respectively, were fully

immunized, while in Uttar Pradesh, those numbers were 34.1% and 52.1%, respectively (Vishwakarma, Shekhar, Dutta, & Yadav, 2019). Another study investigated factors affecting childhood vaccination in the Indian states, and found that Uttar Pradesh had the lowest percentage of fully vaccinated children aged 12 to 36 months among Muslim children (30%), compared to children from other religious groups (70% among Christian children) and in other states (in Tamil Nadu and Arunachal Pradesh, over 80% have full vaccination coverage) (Shrivastwa, Wagner, & Boulton, 2019). At the same time, Hardoi, with a population of 4 million (more than 70% of whom are in rural districts), only 34% of children aged 12–23 months are fully immunized in the rural areas compared to 39% in all districts (Sciences, 2007).

A cross-sectional study in Hardoi (rural site) and Kirti Nagar (urban site) assessed the association between maternal health literacy (measured by the degree to which people have the ability to identify, process, understand and respond to basic health information) and by DTP3 vaccination (standard measure of routine immunization coverage measured by receipt of three-dose of diphtheria-tetanus-pertussis vaccine among children 12–23 months of age). The results revealed a positive association between maternal health literacy and DTP3: in Hardoi, DTP3 vaccination was 49.1% (192/490) among high health literacy populations compared to 33.5% (156/490) in low health literacy populations. The lack of awareness or parents' perceptions of non-health benefits was reported in 35.9% (306/852) as the main reason given by mothers for their child's incomplete immunization (Johri et al., 2015).

Developing interventions tailored to local beliefs and context are needed to reduce vaccine-preventable deaths and to reduce geographical disparities for achieving the Sustainable Development Goal (SDG) child survival targets in India by 2030. India sees promise in technological innovations to improve rural healthcare due to their broad geographical coverage, the low cost, availability of mobile handset and widely deployed infrastructure (Bassi et al., 2018; Ganapathy & Ravindra, 2008; Madanian, Parry, Airehrour, & Cherrington, 2019; Rajasekera, Mishal, & Mori, 2020). According to a report from Telecom Regulatory authority of India, of the 1203.77 million number of telephone subscribers in January 2019, 45% (530.86 million) were rural telephone subscribers (Shri S. K. Mishra, 2019), opening the landscape to develop mHealth

interventions that could reduce health disparities, improve healthcare outcomes, with leaving no one behind and achieving universal health coverage.

Overview of the Tika Vaani pilot study

The Tika Vaani intervention model is a new approach to educate and empower beneficiaries to improve immunisation uptake and child health developed in Uttar Pradesh, India. Although vaccination is the primary focus, the intervention also addresses health care entitlements, prevention, recognition and management of common infectious diseases (diarrhoea, pneumonia, dengue, chikungunya), nutrition, and water, sanitation and hygiene (WASH). The Tika Vaani intervention was tested in a pilot study using a cluster-randomised design. Clusters were rural villages in the Hardoi district, Uttar Pradesh, India. Villages were eligible for inclusion if they had less than 4000 inhabitants and were located in Bawan Block, Hardoi district. Villages were assigned to either intervention or control using simple randomization with a 1:1 allocation following a computer-generated randomization schedule. A total of 387 (184 intervention; 203 control) households with children ages 0 to 12 months in 26 (13 intervention, 13 control) villages were included in the study. All inhabitants (n= 8516) residing in the study villages, as well as the 33 health workers of the intervention group (12ASHA, 11AWWs and 10 AWW Helpers) were also invited to participate.

WHAT IS THE TIKA VAANI INTERVENTION?

Tika Vaani operates through two mechanisms:

I. mHealth strategy:

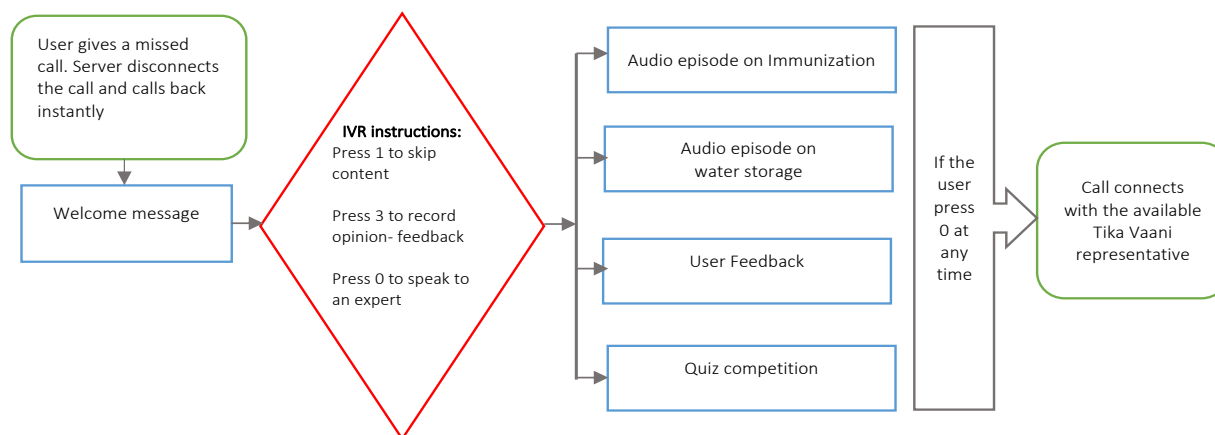
This consists of three contact pathways:

- a) Audio messages are “pushed” via an IVR (Interactive Voice Response) system to mobile phones of families with young children in the age range of 0-12 months. At the beginning of each week, families receive a call with a specific health information capsule. Five days later, they receive a call with the corresponding reminder capsule. Information capsules range in length from 3 to 9 minutes, and reminder messages from 1 - 3 minutes.

- b) The reminder voice messages for immunization are based on the child’s birthdate and the GoI (Government of India) immunization schedule. When a child’s vaccination is due, families receive three calls (four days prior, one day prior, and on the day of the VHND (vaccination day)) informing them that the child’s vaccinations are due and to go to the local vaccination center.
- c) “On-demand” access to content is also provided where people can give a missed call to the IVR phone number, and receive an automated call back over which the same pre-recorded informative stories/capsules are made available. Any person residing in an intervention village can contact the IVR system at any time to repeat the information or to obtain other information of interest.

In all the above contact pathways, people can navigate to two sections (Figure 1). The first section contains infotainment messages on vaccination and WaSH issues. In case a caller has a specific query which is not answered by the pre-recorded messages, they can press 0 at any time to reach the second section which takes them to a helpline where a Tika Vaani team member speaks with them to understand their query, and makes the relevant content available to them.

Figure 1: IVR (Interactive Voice Response) structure for Tika Vaani.



Access to this mHealth strategy required physical infrastructure such as electricity availability, mobile network coverage in the villages, and ownership of a mobile phone or access to mobile

phones to receive the information. For the present study, only one village in intervention group did not have electricity supply and one village had less than 6 hours electricity supply per day. 387 households, 94% (173/184) households owned a mobile phone in the intervention group, and 92% (188/203) household owned a mobile phone in the control group. Mobile phone network quality was reported poor only in 11% (21/184) households in the intervention group while the control group did not report any problems on the network.

To run the IVR, we chose the Mobile Vaani (MV) platform built by a social technology company located in New Delhi, called Gram Vaani (Moitra et al., 2016). The MV platform was developed using open-source components. The hardware (servers) are housed in a data center located in New Delhi, with all the state-of-the-art security features deployed on the server. The Gram Vaani team has accumulated over a decade of experience in designing IVR (Interactive Voice Response) platforms for less-literate rural populations. In addition to the physical and technological requirements already mentioned, the mHealth strategy for the pilot study required to have the telephone number of the target people of the intervention group in order for them to receive messages with content about child health and timely reminder voice messages for immunization. For this, a database was created for storing phone numbers of the participants. The telephone numbers of the participants were obtained from the household surveys that were carried out before the intervention implementation phase. Although the intervention did not require integration with existing health information systems for the pilot study, some upcoming Health Management Information Systems (HMIS) of the government of India also maintain the phone numbers of the patients/ beneficiaries. Therefore, The Tika Vaani platform could leverage these systems to push out relevant messages or information to different cohorts of patients/ beneficiaries, based on their needs and stages of life.

A cornerstone when using mobile devices in health care is to ensure the privacy and confidentiality of the participant's identity and health information. To ensure data security of the activities related to the Interactive Voice Response system (IVR), while at the same time protecting the privacy and safety of research participants, the Gram Vaani team takes the following measures: (i) The call logs, survey data, and audio recordings, obtained through the Mobile Vaani IVR and app systems, are stored on Gram Vaani owned servers which are hosted in a data center in India.

The servers are configured with standard Linux security methods such as password-less logins, disabled ports, and are placed behind a series of firewalls to prevent unauthorized access. Access to the servers is restricted to a small team of technology operation managers; (ii) Processed data of the call logs, surveys, and audio recordings, can be obtained through an access-controlled login restricted only to the project team members, via a login/ password-based web dashboard. Only the dashboard is accessed over HTTPS to ensure encrypted and authorized communication. Suitably anonymized data can be released to the rest of the collaboration team for analysis. Access is restricted only to the Gram Vaani research team members.

II. Community mobilization strategy:

Community mobilization is useful for informing and keeping in touch with community concerns, clearing information, answering questions and has significant potential to create change and improve health (Lassi, Das, Salam, & Bhutta, 2014; Organization, 2017). We used this approach to facilitate the mHealth strategy (Sondaal et al., 2016). It consists of one large introductory meeting and three small meetings in the 13 intervention villages. In the large introductory meeting, the field team introduces the project to participating communities and explains the operation of the Tika Vaani IVR system. A mobile phone and speaker are used to disseminate information. Field team members request permission to paint village walls to promote the intervention by displaying the intervention name, logo, and phone number. The meeting is announced throughout the village by going door-to-door, with a focus on target households with children in the age range of 0 to 12 months (previously identified in the baseline survey). Pamphlets (sheets) and stickers containing the Tika Vaani phone number are distributed. Overall, in the 13 intervention villages, a total of 1692 participants attended these large introductory meetings.

Additionally, a series of 3 small group meetings (held in the day, a time and place that is convenient for the participants) are replicated several times per village to enable widespread participation. Small group meetings are held separately for men and women. The purpose of these meetings is to receive feedback on a series of educational audio capsules on different topics related

to child health delivered via the mobile phone. The meetings help get feedback about whether the programs were heard, about the usability of the technology, and to provide any further usage guidance to the participants. A structured meeting guide is used to learn about the experiences and the perception of the community regarding the Tika Vaani intervention. A total of 96 meetings (n=915 participants) were held during the intervention period. In total, 12 capsules of specific health content and their corresponding reminder capsules were presented to the people.

The intervention is implemented by paid field staff. Equal numbers of males and females are recruited since our experience indicated that some activities are more successful if delivered by someone of the same gender. Selection criteria for field staff includes an interest in promoting healthy lifestyles, residence in the intervention area (Hardoi district), completion of at least 12 years of schooling, experience in working with the community, and ability to speak Hindi and the local dialect. To develop the skills necessary and to standardize the knowledge of the field staff, a manual containing key concepts for working with the community, and a schedule of intervention activities, is used. Competence in relevant skills is assessed prior to the intervention.

HOW WAS THE INTERVENTION DEVELOPED?

The intervention was developed in two stages: (1) definition of technical public health content; (2) adaptation to local populations. Our approach was informed by social marketing theory, which encourages an understanding of key audiences, identification of intervention channels, and development of appropriate messages (Evans, 2006; Gittelsohn et al., 2006):

1) Key orientations for the behavior change communication messages were established through a process evaluation of Village Health and Nutrition Days (VHNDs), the Government of India's primary strategy to deliver health care to rural residents. Results showed that a subset of high-priority services for antenatal care, institutional delivery, and vaccination were being delivered quite well, while other basic health promotion and prevention services were neglected in the VHNDs (Johri et al., 2019). We consulted the public health literature, giving central place to resources such as Facts for Life (UNICEF; WHO; UNESCO; UNFPA; UNDP; UNAIDS; WFP and the World Bank, 2010) and Alive & Thrive (Thrive, 2017) to develop technical guidance and

key messages related to these topics of importance for child health and nutrition. The behavior change messages were reviewed by national and international scientific experts and members of the Indian Ministry of Health.

2) The content was then adapted to local populations through formative research conducted from January 1st, 2017 to January 10th, 2018 in the Hardoi district. A dynamic iterative and participatory approach was favored to validate the need for the interventions and to develop a linguistically and culturally appropriate strategy.

Additionally, a formative research was conducted, with three objectives:

- a) To assess the perceived need for a new service offering basic health promotion and protection information to families with young children, and to identify specific topics of interest

To assess need, appropriateness, and acceptability of the planned interventions, we conducted an interviewer-administered survey of households with a child less than 24 months. Villages were purposively selected from a single administrative block of Hardoi district, Uttar Pradesh based on available census data (Office of the Registrar General and Census Commissioner, 2014). Households were identified through a door-to-door census in randomly selected villages. All eligible households in the village were selected and additional villages were included to achieve a minimum study size target of 100 households. The household survey assessed possession of a mobile phone, quality of the local cell phone network, ability of participants to use a mobile phone, interest in receiving health information, specific topics of interest, a self-evaluation of the level of knowledge about child health, and appropriate channels for receiving health information. A sample health information capsule was also played via mobile phones to evaluate interest and understanding of key messages.

A structured survey of front-line health workers was also conducted. The front-line health worker survey assessed similar themes to those of the household survey, and asked in addition about the front-line workers experience, past-training, and interest in getting additional training. It also requested the front-line workers to perform a self-evaluation of their level of knowledge about child health. These surveys helped to identify which issues and communications methods were

likely to be most important for different audience segments. Data collection took place from June 6th, 2017 to August 8th, 2017. The final sample included respondents from 10 villages: 18 frontline workers and 119 households with at least one child less than 24 months of age. Within these households, 221 independent structured interviews were conducted involving 106 mothers, 56 fathers, and 59 mothers-in-law.

- b) To contribute to the design and refinement of the interventions and increase probable intervention uptake, acceptability, satisfaction
- c) To design and pre-test the evaluation instruments

To inform the design and refinement of the study interventions and improve evaluation tools and methods, we conducted individual meetings, outreach activities and community meetings in 38 villages. Activities took place from January 1st, 2017 through January 10th, 2018. Villages were purposively selected from a single administrative block of Hardoi district, Uttar Pradesh based on available census data (Office of the Registrar General and Census Commissioner, 2014). Village selection was designed to ensure representation of a range of population characteristics potentially relevant to the success of the interventions. The target populations for meetings were mothers, fathers, and families of young children, but frontline workers and members of the general public were also invited. There were 4460 participants over the 76 meetings, including 424 family members/ caregivers of a child less than 23 months of age, 55 frontline workers (ANM, ASHA, and AWW), 2761 adult members of the general public, and 1220 children (less than 15 years from the general public).

We used qualitative and quantitative methods to accomplish these aims. Where possible, multiple methods were used to gain information from different sources and data convergence was assessed by triangulation. All participants strongly agreed on the need for the intervention. The proposed communications channels (via face-to-face small group meetings and mobile phone) were felt to be promising, although limits to achievable coverage through each method were also raised and documented. For each study objective, Table 1 summarises the formative research methods, findings, and implications for intervention design.

Table 1 Key findings of the formative research

Objective	Methods used	Key findings	Implications for intervention design
1. To assess the perceived need for a new service offering health promotion and protection information to families with young children, and to identify specific topics of interest	Structured (quantitative) surveys with some open questions	<ul style="list-style-type: none"> - Interest was very high; no other viable information sources - Self-assessed health knowledge was medium to low - Child health topics were of most interest 	<ul style="list-style-type: none"> - Intervention addresses an important need, and interest is very high - Content & topics adjusted to the needs
2. To contribute to the design and refinement of the interventions, and increase intervention uptake, acceptability, satisfaction	Structured (quantitative) surveys with some open questions Individual meetings ^a	<ul style="list-style-type: none"> - Basic knowledge to protect and promote child health, is weak - In these poor rural communities, there are few health information sources, particularly for women due to low levels of literacy - Access to mobile phones is widespread - Women have less access to a mobile phone than men - Digital literacy is low, particularly for women 	<ul style="list-style-type: none"> - Need for the intervention is high - Pushed messages adopted as the primary mobile phone strategy, judged to be easier than user-initiated dialling - Capsules and key messages simplified to an elementary school level but with entertaining content oriented towards adults

		- Low levels of female empowerment and decision making	
		- Families find it suitable to listen to content individually or together	
- Advantages and disadvantages of different intervention modalities (mobile phone versus face-to-face) for each user segment (women, men, health workers)	Structured (quantitative) surveys with some open questions Community meetings ^b (direct observations, structured discussions)	- Mobile phone & community meetings are both appropriate information channels - Gender issues relate to empowerment and literacy, and affect access to and ability to use mobile phones - Gender issues affect mobility, particularly for some groups	- Both face-to-face and mobile phone channels are required – each channel meets the needs of a distinct user segment
- Inform recruitment and retention strategies for participants	Outreach activities ^c Internal evidence review & discussion among team members	- Outreach and recruitment activities were appropriate and effective	- Use of a low-cost loudspeaker for community mobilisation was added to the intervention
- Intervention content: efforts to increase interest, cultural appropriateness, comprehension and retention of key messages	Structured (quantitative) surveys with some open questions Community meetings ^b	- Entertaining “edutainment” formats preferred - Barriers related to language and comprehension were important	- An “edutainment” approach to content development was adopted - Content was extensively refined in community meetings to ensure that it

	(direct observations, structured discussions)	retained interest and was understandable. - Development of engaging content was extremely challenging. It was most successful when done by a team including those closest to communities. - A variety of promising formats were retained for testing in the pilot study.
3. To design and pre-test the evaluation instruments	Structured (quantitative) surveys with some open questions	- Baseline and end-line questionnaires were largely appropriate - Questions to measure social attitudes and women's empowerment (from standard questionnaire models) were not well understood. - Questions to measure changes in knowledge, attitudes, and perceptions were proposed.

^aIndividual meetings (n=19) with village leaders. ^bCommunity meetings (n=76) for different user segments (males, females, and older versus younger participants). ^cOutreach activities: wall painting, door-to-door mobilization, and distribution of pamphlets and stickers in all villages

mHealth approaches used in combination with other approaches have been shown to be more effective in improving maternal, newborn and child health, as compared to strategies using mHealth approaches alone (Sondaal et al., 2016). Our intervention design therefore retained both channels. Development of technically accurate, comprehensible, and engaging content for this rural, low-literate population was especially challenging. Content was initially not well received by communities for a variety of reasons; the iterative formative research process continued until research and implementation teams were satisfied that the content could be of genuine benefit.

The iterative formative research process was challenging, substantially delayed project timelines, and did not generate high-quality research outputs as traditionally defined. Notwithstanding, formative research was essential to improving intervention design and ensuring that the Tika Vaani program was aligned with the needs and realities of the target population.

HOW DID IT PERFORM IN PRACTICE?

Based on the evaluation of the pilot study, implementation fidelity and acceptability assessment results showed intervention feasibility for all criteria related to the study process (Johri et al., 2020). The intervention was designed according to the needs of the context and community's needs, favoring high uptake and high adherence among the target population. In effect, 94% (173/184) of the target population participated in at least one of the components of the intervention. An mHealth item was considered received if the caller remained connected for 80% or more of the item duration to receive information. Particularly, uptake of mHealth audio vaccination reminders and mHealth edutainment capsules were accessed by 63% (115/184) of target households and 60% (11/184) of target households respectively. The small group meetings were attended by 78% (144/184) of the target households. Owning a phone or having easy access to the phone by mothers were factors that shaped the intensity of uptake for mHealth strategies, while living far from the meeting site was an important factor for attendance in small group meetings. Combining

both strategies to reach the target population was a key strategy given the factors that influenced the participation of the target population.

Designing health interventions to better meet the needs of communities not only improves uptake and adherence but also improves participants' satisfaction and behaviors. The intervention was well-received by the participants and showed evidence of improvement in basic health knowledge, supporting the potential for impact at scale. For example, to the question related to immunization knowledge: "Do you know the immunization schedule for children from birth to 5 years of age?", the probability of correct responses among those receiving the intervention at end line was 66.67% (n=102/153) compared to 26.63% (n=49/184) at the baseline. For the control group the probability of correct responses was 44.58% (n=74/166) at the end-line compared with 34.48% (n=70/203) at baseline. Furthermore, the participants found the content friendly, entertaining and easy to understand. In addition, a strength of the study is that it was not limited to verifying with the participants the usefulness of the content but also asked them about the use of the telephone as a way to receive the information and improve health outcomes. Participants feel comfortable receiving the information at home and liked the fact that information could be accessed on-demand. In addition, frontline workers were convinced that the different components of the intervention did modify the behavior of the community towards childhood vaccination:

FLW 17 "They [mothers] don't pay attention to us, but they can understand information through TikaVaani".

FLW 12 "Those [mothers] who did not come for vaccination, now they come".

Finally, fidelity assessment, an evaluation to determine if the intervention as planned (Carroll et al., 2007) can be implemented on a large scale, revealed a high percentage (86.7%) for implementation fidelity (Pérez et al 2020). The commitment of the field team to implement the intervention, training supervision and monitoring process, economic, human and logistic resources were factors that contributed to the success of the

design and development of the intervention, as planned (Aranda-Jan, Mohutsiwa-Dibe, & Loukanova, 2014). The duration of the pilot study was too short to assess changes in the vaccination coverage, but the objective for which the pilot was tested was indeed met and it was shown that the Tika Vaani Intervention can be implemented at a large scale.

WHAT ARE THE LESSONS FOR SCALE-UP AND REPLICATION?

Our results support the potential contribution of mHealth interventions to increase access to health information in low-resource settings. Our findings are consistent with other studies that have documented that Interactive Voice Response system can provide information access with low technical and functional literacy (Chakraborty, Gupta, & Seth, 2019; Mudliar, Donner, & Thies, 2012; Patel et al., 2009; Patel, Chittamuru, Jain, Dave, & Parikh, 2010; Raza et al., 2016; Vashistha, Cutrell, Borriello, & Thies, 2015). Additionally, the IVR system can be used on all types of mobile phones, and thus can enable reach to many users. Studies carried out in rural India showed that mobile handsets equipped with a voice response technology greatly outperform text-based interfaces and are more useful to meet the needs of less-literate users (Medhi et al., 2011). We opted by audio messaging due to the low literacy and technological comfort levels of the participants.

Among the strengths, the intervention was implemented as planned and results showed significantly higher basic health knowledge among the intervention group supporting the intervention theory of change. In addition, the field implementers understood and believed in the program objectives which facilitated the implementation. The motivation of the target group and commitment of the implementers and other stakeholders was also high, which are likely to be important drivers for scaling.

We found that an mHealth strategy offers the potential to improve knowledge, but the content must be adapted to meet the needs of less-literate users. In our study, stories and pure information formats were both effective, but story formats were especially appreciated. Dramas that touched the heart and included a crisis, denouement, and humour,

were most successful. Having the underdog win and happy endings were popular. The ideal length was 1 to 5 minutes. Another key lesson was that the content should be designed to be inclusive of all social groups, particularly the marginalised. By thoughtful choice of names and positive portrayal of characters, those designing content can increase relatability and foster a sense of inclusion and respect. Aspirational stories placing female, minority and disabled characters in positions of authority (e.g. doctors, government officers) were universally well received.

The intervention is also low cost, both for the participants as well as for institutional expenses. Access by participants through mobile phones or community meetings is kept free for them. Although we did not do a formal economic evaluation for the pilot study, the basic cost regarding the mobile strategy, and a one-year intervention that on-boards and engages 50,000 beneficiaries, costs USD \$ 92,000. Per beneficiary cost for 1 year will be USD \$1.92. Content testing and community meetings (n=3), during the intervention period, further costed USD \$ 3,949.3: Table 2.

Table 2. Basic cost regarding the mobile strategy

	Activities	Cost
The one-time cost	IVR Setup and Deployment	USD \$ 6,000
	Content Development for 30 pieces of content – plot development, scripting, testing, and production (recording and mixing).	USD \$ 40,000
The recurring costs	Call Cost per minute	USD \$ 0.02
	Infra Rental per year	USD \$10,000
	Monitoring & Evaluation: USD per year	USD \$ 6,000
Content testing	- Total man days on content: 352 (one-man day cost 429 INR (Indian Rupee))	USD \$2,270.66 ^a
	- Total cost of content testing in the field= 429 * 352 = 150,999 INR	
Small group meeting	- Total man days on community meetings (3 meeting per village) = 232 (one-man day cost 481 INR (Indian Rupee))	USD \$1,678.64 ^a
	- Total cost of organizing small group meeting = 481 * 232 = 111,630 INR	

^a 1 USD = INR (Indian Rupee) 66.5 (exchange rate for May 2018)

However, several issues require further attention if Tika Vaani is to thrive at greater scale:

- i. The mHealth was very useful to facilitate standardization of health promotion content at low cost, but uptake was uneven. mHealth capsules heard were favoured by mother's possession of a mobile phone, mother's easy access to phone and maternal education.
- ii. For community mobilization, distance from place of residence to the meeting site influenced mothers' attendance at small groups meetings. Women's empowerment was a transversal factor influencing uptake for both strategies. Empowerment women strengthens their role in the family and community, increase awareness of their role in their children's health and can improve the effectiveness of interventions (Janssens, 2011); Men participated actively and were own a mobile phone. Active involvement in the Intervention for men and families to share phone becomes necessary to attenuate gender barriers.
- iii. To achieve equity and reach, both strategies in a complementary way are necessary. However, the logistical and economic challenge to offer the community mobilization component a large scale is important. Given the interest and acceptability of the intervention by frontline health workers (Pérez et al., 2020), the intervention could be incorporated with the Indian health system. Frontline health workers could be trained to become the pivots for community mobilization strategies given their experience, credibility in the communities and their role to implement existing large-scale programs (Bhutta Z, 2010; Pallas et al., 2013). Adoption of this strategy would require adequate training, supervision and remuneration to facilitate their role in the intervention. It might also be necessary to work with community-based NGOs to offer targeted support for community mobilization activities.
- iv. Additionally, other technical aspects to consider for Tika Vaani's replication will be (a) to ensure robust data protection and storage security at the intervention site

(phone numbers and names stored in the database must be confidential data); (b) The content of the stories is transmitted in the local language (Hindi). In other places, where Hindi is not spoken, the content will require translation; (c) The platform is compatible with any type of cell phone, even with landlines, and hence does not require advanced mobile technology to receive the content. However, network quality and electricity must be guaranteed to avoid technical problems in communicating with the users. Finally, (c) Audio messaging were chosen in the study due to the low literacy and technological comfort levels of the participants in the study. Text messages in the Roman script or even the local language can be sent via Mobile Vaani. However, studies carried out in rural India showed that mobile handsets equipped with a voice response technology greatly outperform text-based interfaces and are more useful to meet the needs of less-literate users (Medhi et al., 2011). Therefore, the choice of designing innovations should be based on the sociocultural context and user profile, to ensure better learning and easier access for users.

CONCLUSION

A mHealth intervention strategy adapted to the local context and combined with face-to-face communication is a promising method to educate and empower communities in resource-poor rural areas and an excellent way to reduce health inequities. While the mHealth strategy increased standardization, scalability and reach at low cost, face-to-face communication remains necessary to reach the most vulnerable. This manuscript also corroborates the importance of conducting formative research to meet local conditions. In fact, formative research was a key element to identify individuals' beliefs, interests and perceptions to develop an intervention adapted to the context and develop interesting, entertainment and diversified content for the target population. The methodological rigor with which the intervention was designed, implemented and evaluated, gives us the confidence to promote a large-scale intervention. The challenges to consider for large-scale expansion have also been presented. This study contributes to the literature to improve

reporting on mHealth interventions and provides researchers with key action points during the intervention development to serve hard-to-reach communities and improve health outcomes.

ACKNOWLEDGMENTS

We thank the communities and stakeholders who participated in the formative research phase facilitating the development of the intervention.

References

- Agarwal, S., LeFevre, A. E., Lee, J., L'Engle, K., Mehl, G., Sinha, C., & Labrique, A. (2016). Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ*, 352, i1174.
- Andre, F. E., Booy, R., Bock, H. L., Clemens, J., Datta, S. K., John, T. J., . . . Ruff, T. (2008). Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bulletin of the World Health Organization*, 86, 140-146.
- Aranda-Jan, C. B., Mohutsiwa-Dibe, N., & Loukanova, S. (2014). Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. *BMC public health*, 14(1), 188.
- Bassi, A., John, O., Praveen, D., Maulik, P. K., Panda, R., & Jha, V. (2018). Current status and future directions of mHealth interventions for health system strengthening in India: systematic review. *JMIR Mhealth Uhealth*, 6(10), e11440.
- Bloom, D. E., Canning, D., & Weston, M. (2017). The value of vaccination. In *Fighting the Diseases of Poverty* (pp. 214-238): Routledge.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci*, 2, 40. doi:10.1186/1748-5908-2-40
- Chakraborty, D., Gupta, A., & Seth, A. (2019). Experiences from a mobile-based behaviour change campaign on maternal and child nutrition in rural India. Paper presented at the Proceedings of the Tenth International Conference on Information and Communication Technologies and Development.
- Cross, A., Gupta, N., Liu, B., Nair, V., Kumar, A., Kuttan, R., . . . Rodrigues, R. (2019). 99DOTS: a low-cost approach to monitoring and improving medication adherence. Paper presented at the Proceedings of the Tenth International Conference on Information and Communication Technologies and Development.
- Dumit, E. M., Novillo-Ortiz, D., Contreras, M., Velandia, M., & Danovaro-Holliday, M. C. (2018). The use of eHealth with immunizations: An overview of systematic reviews. *Vaccine*, 36(52), 7923-7928. doi:10.1016/j.vaccine.2018.06.076
- Evans, W. D. (2006). How social marketing works in health care. *BMJ*, 332(7551), 1207-1210.
- Ganapathy, K., & Ravindra, A. (2008). mHealth: A potential tool for health care delivery in India. *Proceedings of the Making the ehealth Connection: Global Partnerships, Global Solutions*.

- Gittelsohn, J., Steckler, A., Johnson, C. C., Pratt, C., Grieser, M., Pickrel, J., . . . Staten, L. K. (2006). Formative research in school and community-based health programs and studies: "State of the art" and the TAAG approach. *Health Education & Behavior*, 33(1), 25-39.
- Hall, C. S., Fottrell, E., Wilkinson, S., & Byass, P. (2014). Assessing the impact of mHealth interventions in low-and middle-income countries—what has been shown to work? *Global health action*, 7(1), 25606.
- Janssens, W. (2011). Externalities in program evaluation: the impact of a women's empowerment program on immunization. *Journal of the European Economic Association*, 9(6), 1082-1113.
- Johri, M., Subramanian, S., Sylvestre, M.-P., Dudeja, S., Chandra, D., Koné, G. K., . . . Pahwa, S. (2015). Association between maternal health literacy and child vaccination in India: a cross-sectional study. *J Epidemiol Community Health*, 69(9), 849-857.
- Johri, M., Chandra, D., Koné, G. K., Sylvestre, M.-P., Mathur, A. K., Harper, S., & Nandi, A. (2020). Social and behaviour change communication interventions delivered face-to-face and by mobile phone to strengthen vaccination uptake and improve child health in rural India: A randomised pilot study for a cluster randomised controlled trial. Under review.
- Johri, M., Rodgers, L., Chandra, D., Abou-Rizk, C., Nash, E., & Mathur, A. K. (2019). Implementation fidelity of village health and nutrition days in Hardoi District, Uttar Pradesh, India: a cross-sectional survey. *BMC Health Services Research*, 19(1), 756.
- Lassi, Z. S., Das, J. K., Salam, R. A., & Bhutta, Z. A. (2014). Evidence from community level inputs to improve quality of care for maternal and newborn health: interventions and findings. *Reproductive health*, 11(2), S2.
- Liu, L., Chu, Y., Oza, S., Hogan, D., Perin, J., Bassani, D. G., . . . Dhingra, N. (2019). National, regional, and state-level all-cause and cause-specific under-5 mortality in India in 2000–15: a systematic analysis with implications for the Sustainable Development Goals. *The Lancet Global Health*, 7(6), e721-e734.
- Madanian, S., Parry, D., Airehrour, D., & Cherrington, M. (2019). mHealth and big-data integration: promises for healthcare system in India. *BMJ Health & Informatics*, 26(1).
- Marcolino, M. S., Oliveira, J. A. Q., D'Agostino, M., Ribeiro, A. L., Alkmim, M. B. M., & Novillo-Ortiz, D. (2018). The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. *JMIR Mhealth Uhealth*, 6(1), e23. doi:10.2196/mhealth.8873

- McGovern, M. E., & Canning, D. (2015). Vaccination and all-cause child mortality from 1985 to 2011: global evidence from the Demographic and Health Surveys. *American journal of epidemiology*, 182(9), 791-798.
- Mechael, P., Batavia, H., Kaonga, N., Searle, S., Kwan, A., Goldberger, A., Ossman, J. (2010). Barriers and gaps affecting mHealth in low- and middle-income countries: Policy white paper: Columbia university. Earth institute. Center for global health and economic
- Medhi, I., Patnaik, S., Brunskill, E., Gautama, S. N., Thies, W., & Toyama, K. (2011). Designing mobile interfaces for novice and low-literacy users. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 18(1), 1-28.
- Moitra, A., Das, V., Vaani, G., Kumar, A., & Seth, A. (2016). Design lessons from creating a mobile-based community media platform in Rural India. Paper presented at the Proceedings of the Eighth International Conference on Information and Communication Technologies and Development
- Mudliar, P., Donner, J., & Thies, W. (2012). Emergent practices around CGNet Swara, voice forum for citizen journalism in rural India. Paper presented at the Proceedings of the Fifth International Conference on Information and Communication Technologies and Development.
- National Academies of Sciences, E., & Medicine. (2016). The promises and perils of digital strategies in achieving health equity: workshop summary: National Academies Press.
- Noordam, A. C., George, A., Sharkey, A. B., Jafarli, A., Bakshi, S. S., & Kim, J. C. (2015). Assessing scale-up of mHealth innovations based on intervention complexity: two case studies of child health programs in Malawi and Zambia. *Journal of health communication*, 20(3), 343-353.
- Office of the Registrar General and Census Commissioner. (2014). Hardoi PCA CDB 0924-F: C.D. Block Wise Primary Census Abstract Data (PCA) - UTTAR PRADESH: India: Census 2011. Retrieved from New Delhi: http://censusindia.gov.in/pca/cdb_pca_census/Houselisting-housing-UP.html
- Organization, W. H. (2011). mHealth: new horizons for health through mobile technologies. mHealth: new horizons for health through mobile technologies.
- Organization, W. H. (2017). WHO Strategic Communications Framework for effective communications. Geneva: WHO.
- Organization, W. H. (2018). Immunization coverage- Key challenges. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/immunization-coverage>
- Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P., & Parikh, T. S. (2009). A comparative study of speech and dialed input voice interfaces in rural India. Paper

presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Patel, N., Chittamuru, D., Jain, A., Dave, P., & Parikh, T. S. (2010). Avaaj otalo: a field study of an interactive voice forum for small farmers in rural india. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Peek, M. E. (2017). Can mHealth interventions reduce health disparities among vulnerable populations. *Divers Equal Health Care*, 14(2), 44-45.

Pérez, M. C., Chandra, D., Koné, G. K., Singh, R., Ridde, V., Sylvestre, M.-P., . . . Johri, M. (2020). Implementation fidelity of the Tika Vaani intervention to improve vaccination uptake and other health outcomes among rural Indian children: A Mixed Methods Evaluation of a pilot cluster randomized controlled trial. Under review.

Rajasekera, J., Mishal, A. V., & Mori, Y. (2020). Innovative mHealth Solution for Reliable Patient Data Empowering Rural Healthcare in Developing Countries. In *Big Data Analytics in Healthcare* (pp. 83-104): Springer.

Raza, A. A., Kulshreshtha, R., Gella, S., Blagsvedt, S., Chandrasekaran, M., Raj, B., & Rosenfeld, R. (2016). Viral spread via entertainment and voice-messaging among telephone users in india. Paper presented at the Proceedings of the Eighth International Conference on Information and Communication Technologies and Development.

Sciences, I. I. f. P. (2007). National Family Health Survey (NFHS-4), 20015–16: India. 1. Retrieved from <http://rchiips.org/nfhs/NFHS5tenders.shtml>

Shri S. K. Mishra, P. A. F. E. (2019). Telecom Regulatory Authority of India. Technology. Retrieved from https://main.trai.gov.in/sites/default/files/PR_No.22of2019_0.pdf

Shrivastwa, N., Wagner, A. L., & Boulton, M. L. (2019). Analysis of State-Specific Differences in Childhood Vaccination Coverage in Rural India. *Vaccines*, 7(1), 24.

Sondaal, S. F. V., Browne, J. L., Amoakoh-Coleman, M., Borgstein, A., Miltenburg, A. S., Verwijs, M., & Klipstein-Grobusch, K. (2016). Assessing the Effect of mHealth Interventions in Improving Maternal and Neonatal Care in Low- and Middle-Income Countries: A Systematic Review. *PLoS One*, 11(5), e0154664. doi:10.1371/journal.pone.0154664

Thrive, A. (2017). *Alive & Thrive: Resources Toolbox*.

UNICEF; WHO; UNESCO; UNFPA; UNDP; UNAIDS; WFP and the World Bank. (2010). *Facts for Life: Fourth Edition*. Retrieved from New York

- Vashistha, A., Cutrell, E., Borriello, G., & Thies, W. (2015). Sangeet swara: A community-moderated voice forum in rural india. Paper presented at the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.
- Vishwakarma, M., Shekhar, C., Dutta, M., & Yadav, A. (2019). Gaps in infant and child mortality among social groups and its linkages with institutional delivery and child immunization using census and National Family Health Survey (2015-16). *Journal of Public Health*, 1-11.
- Wallis, L., Blessing, P., Dalwai, M., & Shin, S. D. (2017). Integrating mHealth at point of care in low-and middle-income settings: the system perspective. *Global health action*, 10(sup3), 1327686.

3.3 Article 3 Implementation fidelity and acceptability of an intervention to improve vaccination uptake and child health in rural India: A mixed methods evaluation of a pilot cluster randomized controlled trial

Pérez, M. C^{1,2}., Chandra, D³., Koné, G. K⁴., Singh, R⁵., Ridde, V^{6,7}., Sylvestre, M.-P^{1,2}., Seth, A^{5,8}. & Johri, M^{1,9}

STATUT : ARTICLE PUBLIÉ

Pérez, M. C., Chandra, D., Koné, G. K., Singh, R., Ridde, V., Sylvestre, M.-P., Seth, A., & Johri, M. (2020). Implementation fidelity and acceptability of an intervention to improve vaccination uptake and child health in rural India: A mixed methods evaluation of a pilot cluster randomized controlled trial. *Implementation Science Communications*. 2020

AUTHOR AFFILIATIONS

¹ Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montréal, Québec, Canada.

² Département de Médecine Sociale et Préventive, École de Santé Publique (ESPUM), Université de Montréal, Montréal, Québec, Canada

³ Independent Consultant, New Delhi, India

⁴ Principal Technical Advisor for Health Financing at Management Sciences for Health (MSH)/USAID, Port-au-Prince, Haiti

⁵ Gram Vaani Community Media Pvt. Ltd., New Delhi, India

⁶ Centre de recherche en santé publique, Université de Montréal, 7101 avenue du Parc, Montréal, Québec, Canada

⁷ IRD (French Institute for Research on Sustainable Development), CEPED (IRD-Université Paris Descartes), Université de Paris, ERL INSERM SAGESUD, 45 rue des Saints-Pères, 75006, Paris, France

⁸ Department of Computer Science, Indian Institute of Technology Delhi, New Delhi, India

⁹ Département de gestion, d'évaluation, et de politique de santé, École de Santé Publique (ESPUM), Université de Montréal, Montréal, Québec, Canada

Contribution des auteurs :

Dans le cadre du projet de recherche de Dre Mira Johri, sous la supervision de cette dernière, et conformément au mandat confié à l'étudiante-chercheuse, Myriam Cielo Pérez a développé la conceptualisation de l'évaluation, a conçu la méthode, a développé les outils de collecte de données (liste de contrôle, entretiens et focus groups). Les données provenant de la communauté ont été colligées par le groupe du terrain et les données du groupe de terrain ont été colligées par deux personnes externes à la mise en œuvre de l'intervention à l'aide de différents outils développés par l'étudiante-chercheuse. Par la suite, Myriam Cielo Pérez a réalisé l'analyse des données, l'interprétation des résultats et a rédigé la version finale de l'article. Dre Mira Johri a supervisé l'ensemble de ces étapes, et a apporté un regard critique et constructif à tout le processus de la démarche, à la révision de l'article et le soutien financier pour la publication de l'article.

Dinesh Chandra a participé à la collecte des données du terrain, à l'interprétation des données documentaires et à la révision de l'article.

Georges Koné a analysé les données quantitatives de l'intervention et a participé à la révision de l'article.

Singh Rohit a participé à la collecte des données de la plateforme mobile et à la révision de l'article.

Dr Valery Ridde a porté un regard critique et constructif au protocole de recherche pour ce volet de la thèse et à la révision de l'article.

Dre Marie-Pierre Sylvestre a participé à la révision de l'article.

Dr Adi Seth a participé à la révision de l'article.

Tous les auteurs ont lu et approuvé le manuscrit final pour soumission.

ABSTRACT:

Background

The Tika Vaani intervention, an initiative to improve basic health knowledge and empower beneficiaries to improve vaccination uptake and child health for underserved rural populations in India, was assessed in a pilot cluster randomized trial. The intervention was delivered through two strategies: mHealth (using mobile phones to send vaccination reminders and audio-based messages) and community mobilization (face-to-face meetings) in rural Indian villages from January to September 2018. We assessed acceptability and implementation fidelity to determine whether the intervention delivered in the pilot trial can be implemented at larger scale.

Methods:

We adapted the Conceptual Framework for implementation fidelity to assess acceptability and fidelity of the pilot interventions using a mixed methods design. Quantitative data sources include a structured checklist, household surveys, and mobile phone call patterns. Qualitative data came from field observations, intervention records, semi-structured interviews and focus groups with project recipients and implementers. Quantitative analyses assessed whether activities were implemented as planned, using descriptive statistics to describe participant characteristics and the percentage distribution of activities. Qualitative data were analyzed using content analysis and in the light of the implementation fidelity model to explore moderating factors and to determine how well the intervention was received.

Results:

Findings demonstrated high (86.7%) implementation fidelity. 94% of the target population benefited from the intervention by participating in a face-to-face group meeting or via mobile phone. The participants felt that the strategies were useful means for obtaining information. The clarity of the intervention theory, the motivation and commitment of the implementers as well as the periodic meetings of the supervisors largely explain the high

level of fidelity obtained. Geographic distance, access to a mobile phone, level of education and gender norms are contextual factors that contributed to heterogeneity in participation.

Conclusions:

Although the intervention was evaluated in the context of a randomized trial that could explain the high level of fidelity obtained, this evaluation provides confirmatory evidence that the results of the study reflect the underlying theory. The mobile platform coupled with community mobilization was well-received by the participants and could be a useful way to improve health knowledge and change behavior.

Trial registration: ISRCTN 44840759 (22 April 2018)

Keywords: Pilot study, Child Health, mHealth program, Implementation Science, Implementation fidelity, Adherence, Process evaluation, mixed methods evaluation, Developing Countries and Global Health.

Contributions to the literature

- Fidelity assessment is a useful method for pilot trials to strengthen community interventions, improve their sustainability, and replicate them in other contexts.
- The implementation of community interventions may be affected by diverse factors, including cultural and socioeconomic dimensions, gender, and education. These factors should be systematically studied because they can affect program adherence.
- The training, understanding, and credibility of implementers and the availability of resources and feedback from supervisors are key understudied factors contributing to successful implementation and high implementation fidelity.

BACKGROUND

Public health interventions should be based on the best available evidence, and randomized field experiments are considered the strongest evaluation design by many. However, these interventions may vary during implementation due to diverse factors related to the complexity of the intervention, context, participants, and implementers. Their evaluation is, therefore, usually complex and challenging (Brownson, Fielding, & Maylahn, 2009). These factors need to be considered at the time of evaluation and before replication in other contexts (Campbell et al., 2000; Perez, Minoyan, Ridde, Sylvestre, & Johri, 2018; Richards & Hallberg, 2015; Rychetnik, Frommer, Hawe, & Shiell, 2002). A pilot study is a smaller-sized study that aims to investigate the feasibility of the crucial components of the main study and guide the planning of a large-scale intervention (Thabane et al., 2010). Implementation science is a field of methods to promote the systematic uptake of research findings into routine practice (Eccles & Mittman, 2006). Pilot studies play an essential role in the development of Cluster Randomized Trials (CRT). They can also contribute to the implementation of science goals by helping to identify the factors that may affect intervention effectiveness before conducting a large-scale study. Applying an implementation science lens to pilot studies is crucial to identify the factors that restrict or facilitate their replication (Glasgow & Emmons, 2007; Layde et al., 2012; Medical Research & Council., 2008) and is particularly necessary in low-income countries, which need to optimize resources to benefit as many people as possible and improve public health.

Implementation fidelity, sometimes called adherence or integrity, refers to the degree to which an intervention is delivered as planned (Carroll et al., 2007). Implementation fidelity is crucial for the successful application of evidence-based interventions (Breitenstein et al., 2010; Cohen et al., 2008; Resnick et al., 2005). Fidelity assessment permits the following: (i) the assessment of cause and effect relationships in studies of complex interventions; (ii) the explanation of variations in study results; (iii) the identification of components that require improvement; (iv) the identification of potential barriers to the successful delivery of interventions at scale; and (v) the collection of information needed to identify problems, propose solutions, and help increase the chances

of success of these interventions (Breitenstein et al., 2010; Bruckenthal & Broderick, 2007; Carroll et al., 2007; Cohen et al., 2008; Craig et al., 2008; Hasson, 2010; Resnick et al., 2005).

Context of the Tika Vaani Intervention

It is estimated that among the 19.5 million children worldwide who did not receive all basic vaccines in the first year of life in 2016, 16% were from India (Shrivastwa, Wagner, & Boulton, 2019). India has achieved considerable improvements in child vaccination coverage; however, significant challenges to closing the immunization gap remain, especially those related to beneficiary demand for immunization (Gurnani et al., 2018). Sociodemographic factors, levels of knowledge, and public beliefs may affect vaccination coverage in children (Gurnani et al., 2018; Shrivastwa et al., 2019). Educational approaches targeting the general population and adapted to local beliefs are necessary to address these barriers (Gurnani et al., 2018). Vaccination coverage remains especially low in Uttar Pradesh (Shrivastwa et al., 2019).

The Tika Vaani (“voice of vaccination” in Hindi) intervention was designed to address this issue by using an interactive mobile platform combined with community mobilization to reduce the gap in childhood vaccination. This intervention primarily aimed to increase immunization coverage among children 0 to 2 years old living in rural areas of Uttar Pradesh, India. In addition, to strengthen primary health care, the intervention seeks to improve health literacy among community members, health workers, and families with young children.

Mobile health (mHealth) is defined as medical and public health practice supported by mobile devices (Organization, 2011). Despite the inherent challenges, especially when combined with non-mHealth interventions, mHealth is an effective tool to improve maternal and child health in low- and middle-income countries (LMICs) (Sondaal et al., 2016). In particular, Interactive Voice Response (IVR) technology has become an increasingly popular approach for sending educational messages about behavior change and has been widely used in rural contexts in India to help illiterate communities improve

maternal and child health (Chakraborty, Gupta, & Seth, 2019; Moitra, Das, Vaani, Kumar, & Seth, 2016). A recent study pointed out that although literacy in India has improved over the past three decades, Uttar Pradesh is one of the states with the highest rates of illiteracy (Chandra, 2019), and women are especially disadvantaged. The 23% literacy gender gap in the state is more than four times the 2016 global average. According to the most recent round of the National Family Health Survey (NFHS, 2015), in Uttar Pradesh, 61% of women are literate in comparison to 82.4% of men (IIPS, 2016). In Uttar Pradesh, 55% of rural households and 78% of urban households own a mobile phone (Paul & Murti, 2016). The rapid growth in mobile phone coverage in India is an excellent opportunity to launch mobile-learning programs that could reduce health disparities and improve healthcare outcomes. In addition, social mobilization is recognized as a key health-promotion strategy used by LMICs to promote vaccination (Ofosu-Amaah, 1989). The Tika Vaani intervention uses both the strategies to overcome illiteracy barriers, reach users who do not have a mobile phone, and motivate telephone users to participate in achieving their goals.

Fidelity of the Tika Vaani intervention

The Tika Vaani intervention was evaluated through a pilot CRT study conducted in rural villages of Hardoi district, Uttar Pradesh, between January and September 2018. A quantitative assessment of study feasibility, intervention uptake, coverage, and early impacts is provided in the main pilot trial report (Johri et al., 2020). Overall, quantitative results showed that all predetermined standards for feasibility were met. Intervention uptake was extremely high, and basic health knowledge was significantly higher among the intervention group. This mixed methods study complements the main analysis by addressing implementation research questions that help to shed light on whether the program requires changes before further assessment and large-scale replication. The specific objectives were to (i) measure the level of implementation fidelity of the interventions (quantitative approach), (ii) assess factors that could potentially influence fidelity (qualitative approach), and (iii) determine whether the proposed interventions are acceptable to participants (qualitative and quantitative approach, see Table 1).

Table 1. Outcome variables and data sources for the Tika Vaani pilot study¹

OUTCOME	DEFINITION	APPROACH	ANALYSIS POPULATION	DATA SOURCES
PRIMARY				
FEASIBILITY OF THE FUTURE MAIN STUDY	<i>Ex-ante criteria related to key study processes and measures</i>	<i>Quantitative</i>	<i>Intervention & control groups</i>	<i>& reported in main study</i>
UPTAKE OF INTERVENTIONS (ADOPTION)	<i>Evidence of participation in the new interventions</i>	<i>Quantitative</i>	<i>Intervention group</i>	<i>reported in main study</i>
SECONDARY				
<i>ACCEPTABILITY*</i>	Perception among stakeholders that an intervention is agreeable, suitable, relevant, useful, and credible.	Mixed methods	Intervention group	QUANT IVR platform (mHealth); Household surveys QUAL Semi-structured interviews Discussion groups
<i>FIDELITY*</i>	Ability to deliver the interventions as planned	Mixed methods	Intervention group (some information from controls)	QUANT Project records Structured observation with checklist;

				IVR platform (mHealth); Household surveys QUAL Semi- structured interviews Discussion groups Field observation Document review
COVERAGE	<i>The degree to which a (sub) population that is eligible to benefit from an intervention actually receives it</i>	<i>Quantitative</i>	<i>Intervention group</i>	<i>reported in main study</i>
PROGRAMME THEORY	<i>Anticipated changes in knowledge, attitudes, and practices of end users</i>	<i>Quantitative</i>	<i>Intervention & control groups</i>	<i>reported in main study</i>

IVR - interactive voice response mHealth - mobile health

¹. Adapted from Peters et al., 2013.

* Outcomes analyzed for the present study

METHODS

The Tika Vaani intervention

The study interventions took place over a three-month period and offered social and behavior change communication (SBCC) for members of the general public in rural Indian villages addressing topics related to child health. The intervention is described in detail in the main article (Johri et al., 2020). The program rationale and design were iteratively developed through a comprehensive review of the literature, formative evaluations, (Claudia Bojanowski* et al., 2018) and training of implementers to develop activities that responded to local needs and were sustainable in the long term (Johri, Chandra, et al., 2015). SBCC materials were delivered through two channels: (1) Community (face-to-face) mobilization strategy, including one large introductory meeting offered to each village and three small group meetings offered to each participant; and (2) Individual audio messages via mobile phone (mHealth) strategy delivered through the IVR system. For the mHealth component, automated dial-outs featuring entertaining educational audio capsules and voice immunization reminders were favored, and "On-demand" access to all content was offered. Table 2 summarizes the activities, objectives, and beneficiaries of each key component of the intervention. To facilitate implementation, monitoring, and evaluation, a well-defined role was established for each of the actors involved in the intervention (Table 3) (Saré et al., 2018). More details about the interventions and delivery processes are reported in the associated TiDIER Checklist (Johri et al., 2020). All intervention components were offered free of charge to end users. The control group received all standard health services provided by the Government of India, including monthly Village Health and Nutrition Days (VHND).

Table 2 Intervention components and beneficiaries for the Tika Vaani Intervention

Key Components	Activities per village	Purpose	General public	Primary caregivers and families of children 0 to 12 months of age	Frontline workers*
Community Mobilization (face-to-face strategy)	-Large introductory meeting (n=1)	To inform the community about the intervention and invite participation	x	x	x
	-Small group meetings (n=3)	To educate and reinforce basic health knowledge	x	x	x
Messages via mobile phone (mHealth strategy)	“Pushed” edutainment & summary capsules via mobile phone (n=13)	To educate and reinforce basic health knowledge.	x	x	x
	Vaccination “reminders” via mobile phone	To inform when the child's vaccination is due		x	
	“On-demand” (callback) access via mobile phone to content through the IVR portal	To provide convenient access	x	x	x

* Considered for the present study, Accredited Social Health Activist [ASHA], *Anganwadi* Worker [AWWs] and AWW Helpers (Sahaika).

* All interventions were offered free-of-cost to end users

Table 3 Role of the implementers and the research group in the Tika Vaani Intervention

Actors	Role
Implementers	<p>Field staff</p> <p>Conduct a series of 4 community meetings:</p> <ul style="list-style-type: none"> - A large group introductory meeting to inform the purpose the intervention and encouraging participation; to promote the Tika Vaani intervention by painting the logo, and telephone number on village walls and distribute pamphlets (sheets) and stickers containing the Tika Vaani phone number. - Three small group meetings held at monthly intervals to discuss the themes assigned for each meeting in order to assess the interest, acceptability, and understandability of the information capsules and comprehension and retention of key messages.
Research group	<p>Field staff</p> <ul style="list-style-type: none"> - Monitor the progress of activities carried out in the field - Provide technical support. <p>Intervention coordinator</p> <hr/> <p>-Interactive voice-response (IVR) system coordinator</p> <ul style="list-style-type: none"> - Monitor the progress of information capsules delivered - Provide technical support.

Study Framework to assess Implementation fidelity

We used the conceptual framework proposed by Carroll and colleagues (Carroll et al., 2007) as modified by Hasson (Hasson, 2010) to evaluate the fidelity of the Tika Vaani intervention. This framework, useful for assessing the fidelity of complex interventions (Perez et al., 2018), allows us to assess whether (i) the activities were implemented as

planned (content), (ii) the number of planned activities and the designated area were respected (coverage), and (iii) the activities were delivered with the regularity planned by its designers and occurred over a specified period (frequency and duration). This framework also allows us to explore whether there are specific moderating factors that can explain the degree of fidelity obtained. The framework used does not consider evaluating the control group, but its evaluation is crucial to determine the degree of treatment differentiation, which is the systematic variance that is expected to account for any differences in outcomes (Durlak & DuPre, 2008; Hasson, Blomberg, & Dunér, 2012). We investigated whether members of the control group participated to evaluate possible contamination between study groups. This fidelity assessment used a Mixed-Methods Concurrent design (Creswell & Clark, 2017) because the research question focused on a single phenomenon—the fidelity of the Tika Vaani intervention—and measured fidelity and factors that could explain the achieved level of fidelity, which could not be adequately assessed by either qualitative or quantitative methods alone. Based on the GRAMMS checklist, a guide to improve Good Reporting of A Mixed Methods Study (O'cathain, Murphy, & Nicholl, 2008), we provide information about the fidelity assessment of the Tika Vaani intervention through a Mixed-Methods Concurrent Triangulation design (Additional file #1).

Data collection

Qualitative and quantitative data were obtained from multiple sources and methods (Table 4). Data were collected between April and August 2018. The qualitative tools were translated to Hindi (the local language) and validated before use by field team members who were native Hindi speakers residing in the study area. To preserve meaning in the translation process, two bilingual authors (DC, MJ) fluent in Hindi and English with a detailed understanding of the study context, including the cultural characteristics of the participants, led this process. Qualitative data were collected by a research assistant not involved in the intervention. The triangulation of these methods was used to strengthen the validity of the constructs, ensure consistency, confirm data reliability, and increase the validity of the study (Creswell & Clark, 2017; Yin, 2017).

Table 4 Overview the conceptual framework to assess acceptability, fidelity implementation and type of data collection for each dimension

FIDELITY COMPONENTS		
Definition	Purpose of the information	Type of data
Content Defined as an attempt to establish the “active ingredients” of the intervention, for example, in a theory of change or logic model, and assess whether they have been delivered as planned	Number of components implemented as planned Community mobilization strategy # community meetings mHealth strategy # “pushed” edutainment capsules # “pushed” vaccination reminder messages	Quantitative ^a
Coverage Refers to the degree to which all persons who met study inclusion criteria received the intervention	Community mobilization strategy # and characteristics of individuals (general population & target group) attending the different scheduled sessions mHealth strategy: # and characteristics of individuals using the IVR platform by content type (edutainment (general population & target group), vaccination reminders (target	

group only))

<p>Frequency Refers to whether the intervention was delivered with the regularity or frequency planned by its designers.</p>	<p>Number of activities delivered defined in time and frequency according to the scheduled calendar: Community mobilization strategy # community meetings mHealth strategy # “pushed” edutainment capsules # “pushed” vaccination reminder messages “On-demand” to content through the IVR portal</p>
<p>Duration Establishes whether the intervention was delivered with the duration planned by its designers</p>	<p># “pushed” edutainment capsules # “pushed” vaccination reminder messages “On-demand” to content through the IVR portal</p>

MODERATING FACTORS

Definition	Purpose of the information	Type of data
<p>Comprehensiveness of intervention description</p>	<p>To evaluate the implementers' understanding of: -the theory of intervention</p>	<p>Qualitative ^b</p>

<p>Factors such as the degree of intervention complexity, and whether the intervention description is complete or incomplete, vague or clear, may influence the degree of implementation fidelity</p>	<p>-the activities and resources allocated to the different components of the intervention</p> <p>-the role in the intervention</p> <p><i>This dimension is evaluated in three moments: (i) before beginning the data collection from intervention records, (ii) at the end of the implementation based on discussion groups and interviews with the intervention implementers and (iii) through feedback meetings with program designers</i></p>
<p>Strategies to facilitate implementation</p> <p>Several support strategies may be used to optimize and to standardize implementation fidelity</p>	<p>According to the perspective of the implementers:</p> <p>What were the strategies that facilitated the implementation?</p> <p>What were the facilitating elements and the challenges encountered during the implementation phase?</p>
<p>Quality of delivery</p> <p>Concerns whether an intervention is delivered in a way that increases the likelihood of achieving the desired health outcomes</p>	<p>According to the perspective of the implementers:</p> <p>To assess the quality of the material used, the delivery of the content of the intervention and the participation of the participant</p>

<p>Participant responsiveness</p> <p>Intervention uptake depends on its acceptance by and acceptability to those receiving it. Low participant involvement or responsiveness may negatively impact intervention fidelity</p>	<p>To know the acceptability and usefulness of the activities and the different key messages delivered and according to the perspective the</p> <ul style="list-style-type: none"> -Frontline workers -General public and primary caregivers and families of children 0 to 12 months of age <p>To understand the reasons for non-participation of member of the target group in the proposed activities:</p> <ul style="list-style-type: none"> -Community mobilization (face-to-face) activities -Messages via mobile phone (mHealth) activities
<p>Recruitment*</p> <p>Refers to procedures that were used to attract potential program participants.</p>	<p>According to the perspective of the implementers:</p> <p>To assess recruitment process, recruitment strategies and challenges to attract participants in each group (to compare the level of fidelity achieved in each village).</p>
<p>Context*</p> <p>Refers to surrounding social systems, such as structures and cultures of organizations and groups, and historical and concurrent activities and events.</p>	<ul style="list-style-type: none"> -Reasons for any deviation from the planned activities according to the point of view of the implementers -To assess which contextual factors influence the fidelity obtained in the different components of the intervention -Information on context factors regarding the delivery and receipt of the intervention in the different villages

-The actions of the Government of India to provide primary health care in rural areas

Control group	-Components of the intervention that took place in the control group
Monitoring of events in the control group	during the intervention period.
	-Strategies adopted to prevent contamination in the control group.

Adapted from Carrol et al (Carrol et al., 2007) and these components* added by Hasson (Hasson 2010)

^a **Research methods:** Structured observation with checklist, survey records & data from the interactive voice-response (IVR) system and household surveys. **Data source:** Implementers, Administrative records of the intervention Tika Vaani and the mobile platform IVR

^b **Research methods:** Semi Structured Interviews, discussion groups, documentary review and field observation. **Data source** Implementers, Frontline workers, general public and primary caregivers and families of children 0 to 12 months of age, records of the intervention Tika Vaani and the mobile platform IVR an journal of the main author

Quantitative data:

Structured observation with a checklist was used by each implementer to verify if planned activities were implemented as specified in terms of content, coverage, duration, and frequency. For each planned activity, we evaluated whether it was Implemented as planned (I), Modified (M), or Cancelled (C). Further, implementers were asked to note if any activity was added during implementation (Additional file #2). Data from the IVR system and household surveys conducted at the study baseline and end line in both the study groups were also consulted to understand the degree of fidelity obtained.

Qualitative data:

Qualitative data were obtained from semi-structured interviews, focus group discussions, semi-structured observations (Additional file #3), and analysis of records to provide complementary information on different factors that affected the fidelity of implementation and adjustments to the program.

Participants:

All implementers of the intervention (n=8) provided data from semi-structured interviews and discussion groups to identify possible modifications, difficulties encountered during the implementation phase, and strategies to address these difficulties and understand the program rationale. The semi-structured interviews (n=8) and two discussion groups with implementers lasted 30–60 minutes and were recorded and conducted in a confidential setting.

Community:

Small discussion groups: We analyzed information collected from 915 participants during 96 small discussion groups (Additional file #4) on the program rationale, mobile platform, records, and feedback that the lead author obtained during the development and

implementation phases. Seven meetings were conducted with the group coordinator to clarify and validate information.

Community exit meetings: Community discussions (n=25) were held after the end line survey in all intervention villages to provide an opportunity for partner communities to express their views. A purposive sample based on the characteristics of the target group for the intervention (parents of children younger than 2 years), regardless of participation, and who were available to attend a discussion group, was used to recruit participants for the qualitative phase (Palinkas et al., 2015; Sandelowski & health, 1995). However, all members of the community were invited to give their feedback on the intervention. Several meetings were held in each village to ensure convenience and facilitate discussion. We used a structured interview guide to conduct meetings and a pre-defined data collection tool including closed-form responses and some space for open comments to record feedback. The discussion groups aimed to understand the participants' perceptions, understanding, and acceptance of the activities. Discussions were held in a place easily accessible to participants and lasted approximately 60 to 90 minutes. In total, in the 13 intervention villages, 25 exit meetings (three meetings were attended by men and 22 by women) were held with an average of 20 participants (n=292 participants) per meeting.

Frontline workers: A purposive sample comprising personnel from the target area who were available to attend interviews was recruited (Palinkas et al., 2015; Sandelowski & health, 1995). Semi-structured interviews with frontline workers (n=17) lasting 30–40 minutes were performed at the workplace to examine these professionals' perceptions of the program.

Data Analysis

Quantitative data analysis:

A matrix was created to assess whether activities were modified, canceled, or implemented as planned. For each component, the number of implemented activities relative to that of planned activities was calculated and multiplied by 100 to obtain the percentage representing the degree of fidelity. The percentages were summed up to obtain the overall degree of fidelity of the intervention. Although no specific guideline was used to define the optimal degree, values between 80% and 100% were typically considered high (Borrelli, 2011; Noell, Gresham, & Gansle, 2002). The following scoring categories were used in this study: 80–100%, high; 79–51%, moderate; and $\leq 50\%$, low.

Data from questionnaires, intervention records, and mobile platform records were analyzed by a researcher not involved in the project using STATA software version 15. Subsequently, data considered relevant to measuring fidelity was validated and analyzed by the first author (not involved in the implementation phase). Descriptive statistics were used to describe the characteristics of the participants and the distribution of activities.

Qualitative data analysis:

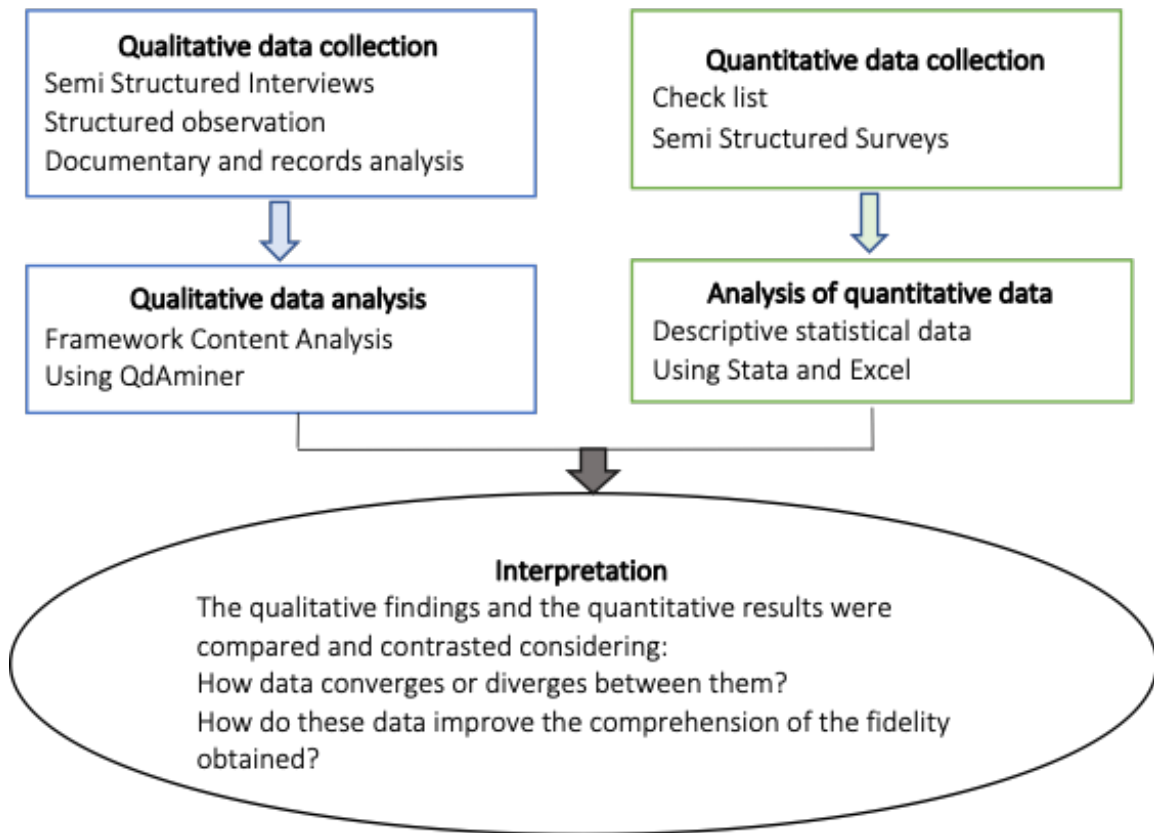
Qualitative data from semi-structured interviews and discussion groups were transcribed and translated from Hindi to English by a researcher not involved in the interventions. Interviews were chosen at random to assess the veracity of the translation. The Framework Method was used for management and analysis of qualitative data (Gale, Heath, Cameron, Rashid, & Redwood, 2013; Ritchie & Spencer, 2002). This phase included four stages: (i) Transcription and translation of the edited data in word format, and reading of the transcribed material to familiarize the researchers with the data; (ii) data coding following two approaches: deductive, to identify the fidelity parameters of the conceptual framework, and inductive, to identify emerging information (Miles & Huberman, 2003); (iii) we kept the quotation that we considered the most relevant to assess

implementation fidelity and assigned them codes according to our conceptual framework, using QDA Miner software to store and organize the data making it accessible for the analysis process; (iv) we use a matrix to summarize the data. The analysis included reading and evaluating the connections between codes by the first author to obtain a complete picture of the fidelity assessment until data saturation was reached. Subsequently, the information was verified and validated by a second author to resolve divergent interpretations and obtain an agreement on the identified findings.

Data integration

Although most data sources were qualitative, both data types had the same priority in this study. Quantitative and qualitative data were collected, analyzed, and integrated in parallel (Mixed-Methods Concurrent Triangulation design) to shed light on the degree of fidelity obtained and increase the validity of our results and conclusions (Plano Clark et al., 2015; Pluye, Bengoechea, & Granikov, 2018) (Fig. 1).

Fig. 1 Mixed methods study flow diagram of collection, analysis, and integration of study data



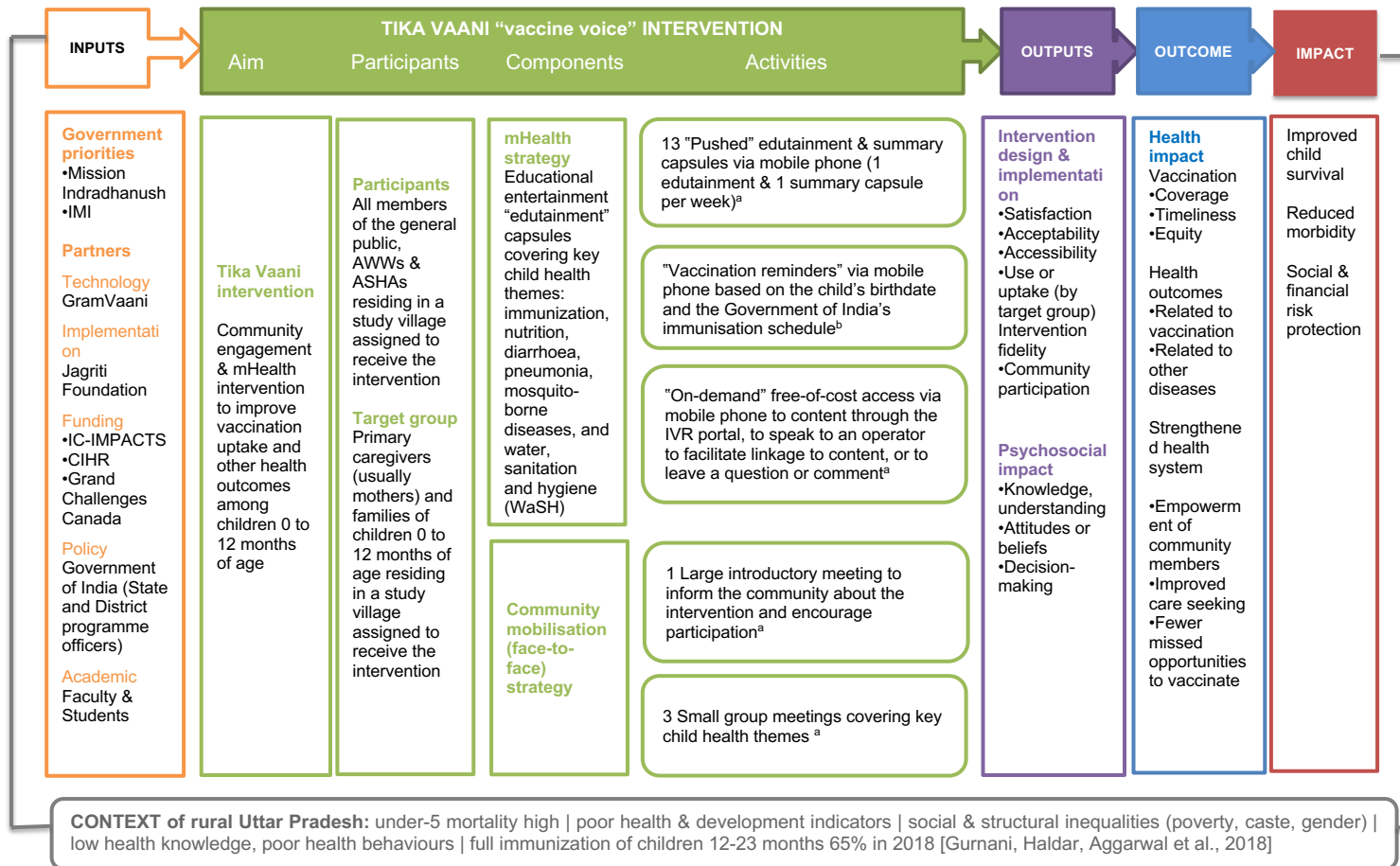
RESULTS

The results are presented according to our conceptual framework. However, comprehensiveness of the policy description (a moderating factor) is presented first to give greater clarity about the intervention evaluated. Additional information relevant to understanding the level of fidelity achieved is provided (Additional file #5).

Comprehensiveness of the policy description

The policy description followed the Template for Intervention Description and Replication (TIDieR) guidelines (Hoffmann et al., 2014). During the study's formative phase, implementers received information and training about the theory of the program, its goals, and the strategies for achieving these goals. In addition, they participated in the development of the project. Discussion groups (n=2) and interviews (n=8) with the implementers allow us to corroborate their understanding and their belief in the program objectives and intervention components, lending validity to the intervention logic model (Fig. 2).

Fig. 2 Logic model for the "Tika Vaani" to improve vaccination uptake and other health outcomes



^a. Activity directed towards target group households but open to all village residents, ASHAs & AWWs

^b. Service offered exclusively to target group households

ADHERENCE

Content

According to different activities of the intervention, the results demonstrated that the fidelity of the intervention was high, with an overall rate of 86.7% (Table 5). For the face to face strategy, two of the 13 villages did not have the large introductory meeting. In one village, this omission was in accordance with the pre-established intervention protocol due to the small population of the village. In another village, the attempt to organize a meeting was unsuccessful.

R-FDBK: “one (1) village there were only 4 households and, in another (1) village, the team tried to organize the meeting, but people didn’t come to the meeting.”

A total of 96 small group meetings were held throughout the study period. The number of meetings per village depended on that of target families in the area. To collect additional data for research purposes, target families that lived far and did not attend the meeting were visited at their homes. For the mHealth strategy, messages via mobile phone were provided to participants with mobile phones who agreed to receive and send messages through the mobile platform.

Table 5 Content dimension

Key Components	Activities for selected villages	Tasks of each planned activity	Planned activity	Implemented activity	% fidelity achieved
Community Mobilization (face-to-face strategy) 90,3%	Large introductory meeting	<ul style="list-style-type: none"> a) First contact with the community leader to get permission. b) To visit each target household to invite them to the meeting c) Collected mobile numbers from the target households d) Collected mobile numbers of the people who wish to receive more health information through the platform. e) Invite ASHA and AWW to participate in the big meeting f) Paint a wall with the logo and the number TV g) Pasted TV poster in the villages h) TV team gave their introduction to community in the introductory meetings i) Demonstration about how to access the TV platform j) Distribution the stickers (sheets) with information on the TV phone number. k) Street play 	13	11	84.6%
	Small group meetings	<ul style="list-style-type: none"> a). To visit each target household to invite them to the small group meetings b) To collect mobile numbers from the target households who wish to receive health information through the platform (This activity was done only during 1st and 2nd small group meetings) 	GD #1 13	GD #1 13	GD #1 100%

	c) To visit ASHA and AWW workers to invite them for small group meetings	GD #2	GD #2	GD #2	
	d) Wall painting the logo and the number TV in the villages	13	11	84,6%	
	e) To Paste TV poster in the villages				
	f) Our team also collected information about newborn children during 1st and 2nd small group meeting.	GD #3	GD #3	GD #3	
	Small group activities:				
	g) Introduction activity with the meeting participants	13	12	92%	
	h) Use a guide sheet to know the experiences and the perception of the community regarding the TV Intervention and participants attending.				
	i) To play reminder capsule or straight content and then TV team discussed these capsules with the participants.				
	j) TV number demonstration				
	k) To distribute TV number slips				
	“Pushed” edutainment & summary capsules	- Transmit 13 educational capsules and 13 reminder messages during implementation period	26	26	100%
Messages via mobile phone (mHealth strategy)	Vaccination “reminders”	-Reminder messages for each target family to remember vaccination period	184	122	66,3%
83,15%	“On-demand” free-of-cost access via mobile phone to content (IVR portal)	-The entire population was invited to participate spontaneously to dial the number to obtain information on the different capsules	-	-	-

Coverage

This study was conducted in 13 villages and included 184 families, which composed the target group. All residents (n=8516) from these villages and 33 frontline workers from the intervention group (12ASHA, 11AWWs, 10 AWW Helpers) were invited to participate. A total of 94% (173/184) of the target population benefited from the program either by participating in the face-to-face strategy (community mobilization activities) or by mHealth strategy (mobile phone messages or listening to at least 80% of the information) through different types of calls: “Pushed” edutainment & summary capsules, vaccination reminder, or on-demand (callback). Table 6 presents the coverage by the different study participants.

Table 6 Coverage dimension

Key Components	Activities	General public	Primary caregivers and families of children 0 to 12 months of age	Frontline workers
	Large introductory meeting	19% (1692/8516)	100% (184/184)	30% (10/33)
Community Mobilization (face-to-face strategy)	Small group meetings	GD #1: 184 GD #2: 269 GD #3: 225	GD #1: 45,6% (84/184) GD #2: 46,7% (86/184) GD #3: 40,2% (74/184)	GD #1: 15% (5/33) GD #2: 36% (12/33) GD #3: 21% (7/33)
	“Pushed” edutainment & summary capsules	-	44.6% (82/184) of households listened to at least one OBD	85% (28/33)
Messages via mobile phone (mHealth strategy)	Vaccination “reminders”	-	66,3% (122/184) received at least 1 VHD reminder message	-
	Use of the IVR system to access health information (on demand access [callback])	-	29.3% (54/184) of households received at least 1 callback	88% (29/33) listened to at least one OBD

Community mobilization (face to face strategy)

- Of the 1694 participants in the large introductory meeting, 75% (n=1264) were women and children, and 25% (n=430) were men. Although men were highly interested in the interventions, the rate of participation was low because most actions were conducted during the day, when they were working.

-Small group meetings: Throughout the study period, 428 people attended at least one small group meeting. Of them, 66% (284/428) came from the community, and 34% (144/428) were caregivers (representing 78.2% of the target group households (144/184)). The target population who lived far from the intervention site constituted 22.8% (42/184), and their participation was low (Fig. 3). The main reason for the non-participation of mothers was being away from home at the time of meetings (Fig. 4). Frontline workers did not attend all meetings because they were not present at the time of the visit to the community. The non-remuneration of this group to attend meetings might also explain the low rate of participation.

Fig. 3 Participation of mothers to the small group meetings

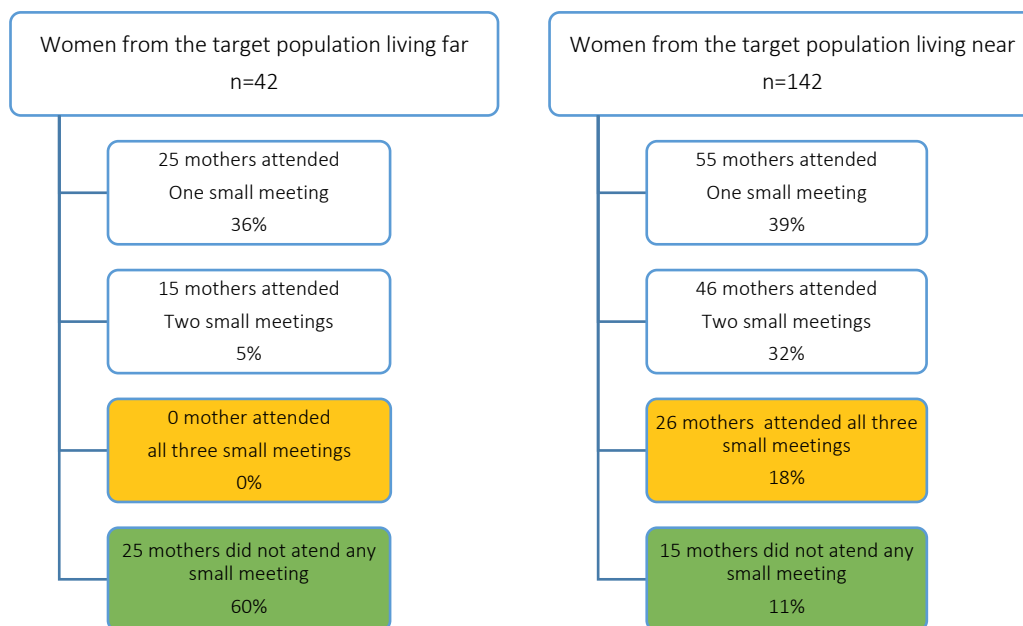
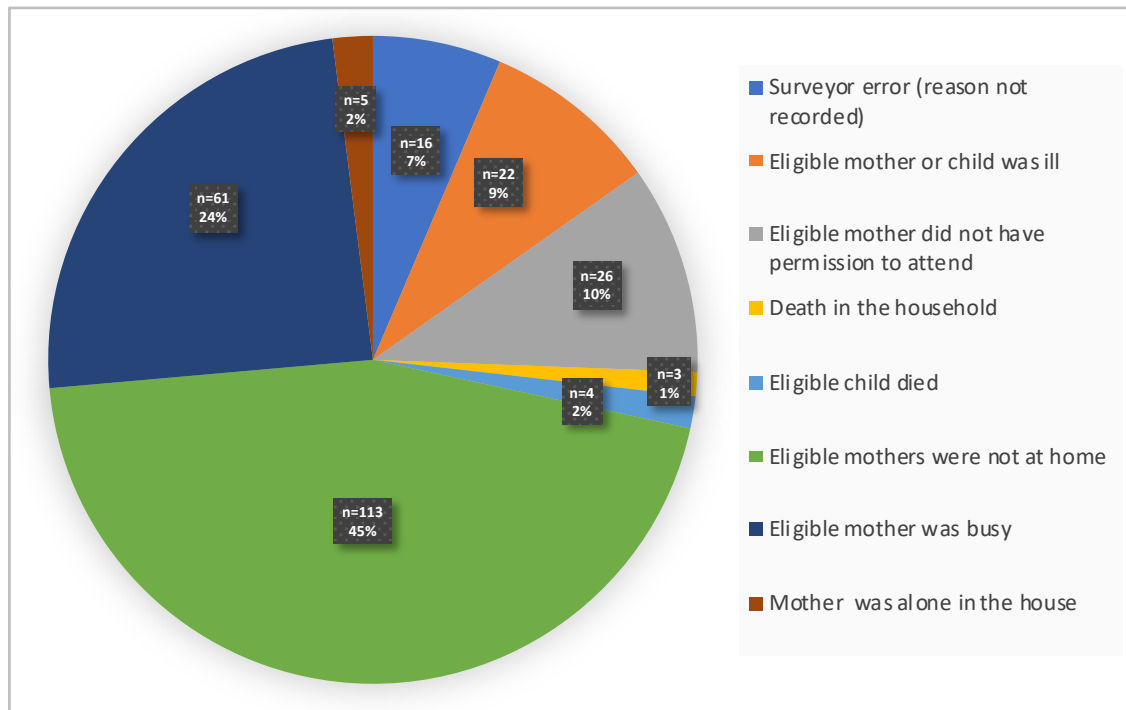


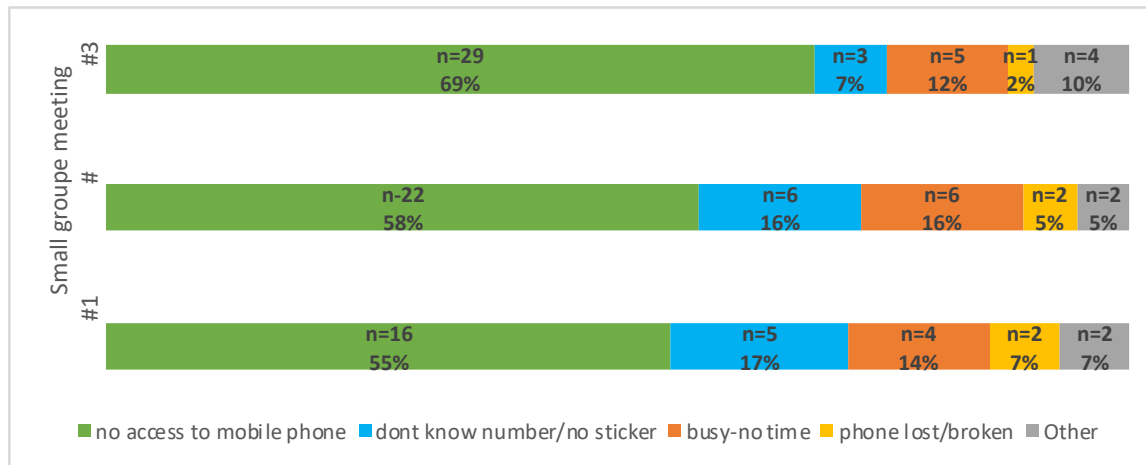
Fig. 4 Reasons for non-participation of mothers in small group meetings.



Messages via mobile phone (*mHealth strategy*)

-The population attending small meetings benefited from “Pushed” edutainment & summary capsules. Among the target population, 70.7% (130/184) received at least one type of “Pushed” edutainment & summary capsules, vaccination reminder, or on-demand (callback). For “On-demand” (callback), the lack of access to a telephone was the main reason for not receiving calls (Fig. 5). Of the 29 frontline workers with a mobile phone, 100% (29/29) heard at least one capsule, and 34.5% (10/29) received at least one call through the platform.

Fig. 5 Reasons given by meeting participants for not listening to the educational capsules via mobile phone



Frequency and duration

The program began a few weeks later than scheduled because of local factors: failure to deliver the information capsules on time, technical problems related to the mobile platform, and official religious holidays. However, there were no significant changes to the interventions during the study period.

MODERATING FACTORS

Strategies to facilitate implementation

During the study’s formative phase, a manual containing key concepts for working with the community—together with a schedule of intervention activities—was used to develop the necessary skills and standardize the knowledge of the field staff. The implementers of the intervention felt that training and training manuals facilitated implementation:

C22: “How we should fill formats and with whom to fill them. We also received training on how to demonstrate Tika Vaani to reach the people of the village ... all this has simplified our work.”

C22: “What and how to do after reaching a village. The manual and guidelines provided were very useful.”

Field experience, motivation, periodic meetings, field notes, and the presence of frontline workers were factors that facilitated community contacts and interactions with participants.

Quality of delivery

The information capsules were evaluated and tested to ensure the quality of the sound and content. During the implementation phase, community events were monitored through meetings with supervisors, and messages via mobile phone were evaluated by the technical team in charge of the mHealth component of the project (RS, AS).

Participant responsiveness

The participants felt that the strategies used were useful means for obtaining information because it was easy to access from home, available on demand, and both easy to understand and entertaining. In addition, they recognize that the information obtained is important for their health.

Community#12 “Yes, method is good [getting information through phones and community meetings]. You people come, do meetings and provide information...We get information while sitting at home.”

Community18: “An eligible mother didn't use to take her children for vaccination but after listing stories on TikaVaani platform now she takes her children for vaccination.”

The content of the calls was considered useful, entertaining, and easily understood by the participants. Moreover, they considered reminder messages about childhood vaccination useful.

Community#2 “there is no other method [reminding families on child vaccination through calls], phone is the best method.”

Community#12 “this information [reminding families on child vaccination through calls] is very helpful for all.”

From meetings and mHealth capsules, the participants received information on how to recognize the main symptoms of conditions such as diarrhea, pneumonia, dengue, and chikungunya, and manage these diseases.

The participation of mothers in different activities was influenced by multiple factors (Figs. 3, 4, and 5). The results showed that the benefits of the program were affected by the level of education and socioeconomic status, whereas sociodemographic characteristics and access to mobile phones directly affected access to information through meetings and mobile devices, respectively. Among the women who attended small group meetings, 29% (54/184) were illiterate, 35% (65/184) were from the poorest quintile, and 69% (127/184) lived close to the meeting place (Fig. 6). Among the women who accessed information through mobile devices, 19% (35/184) were illiterate, 24% (45/184) were from the poorest quintile, and 30% (56/184) had direct mobile phone access (Fig. 7). Among the women who lacked access to a mobile phone, 63% (67/106) were able to access the information through a family member's mobile phone either to listen to vaccination

reminders and edutainment capsules or to communicate with the platform at some point in the intervention period.

Fig. 6 Factors influencing participation of women in community mobilization activities n=184

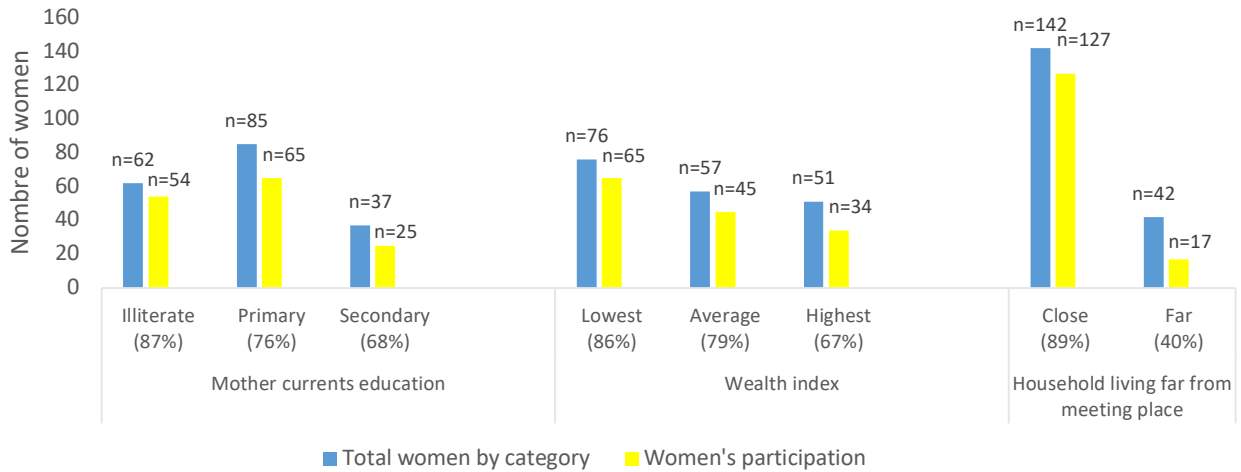
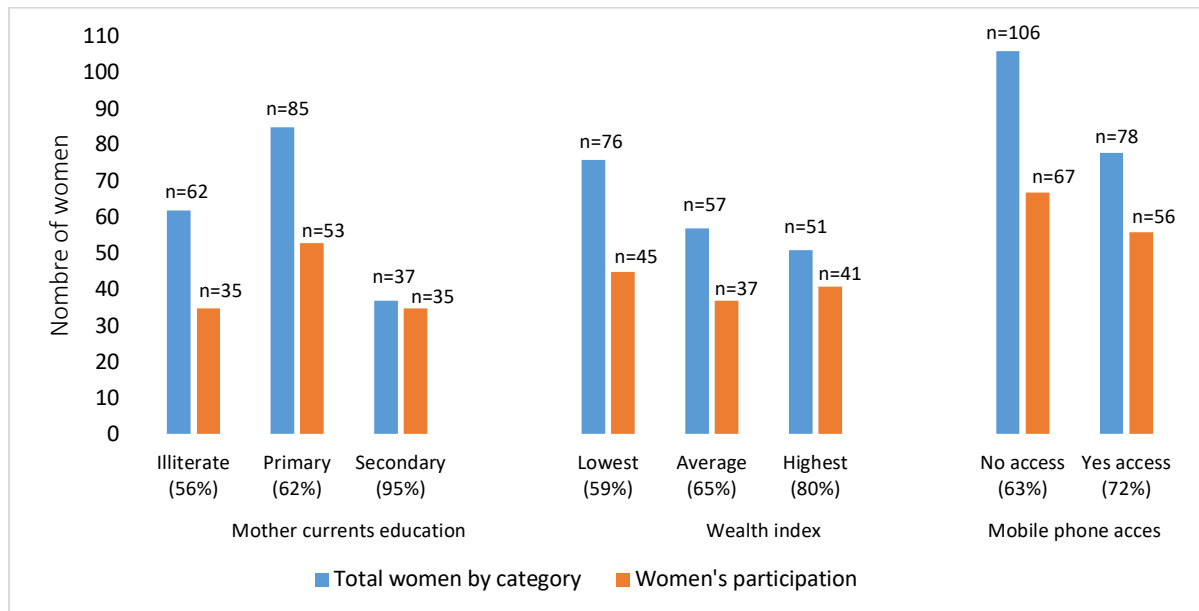


Fig. 7 Factors influencing participation of women in educational capsules via mobile phone n=184

n=184



Frontline workers reported that mothers changed their behavior and were more aware of childhood vaccinations and adherence to the vaccination schedule.

Community15: “Complete vaccination give full protection from diseases”, “Timely vaccination is important to protect from diseases” “there are side effects of vaccines but not serious.”

HW8 “Women say that their children did not get all the vaccine because vaccines are not written in the card. People are becoming aware... After listening content over this number, people started coming for vaccination.”

HW3 “now vaccination is happening successful people take their children immediately if their child is sick... Now everyone is coming for vaccination.”

In view of the benefits to the community, frontline workers reported that the intervention should continue over time and were interested in receiving further training.

HW8 “Your team should keep coming to make difference among people.”

HW2 “Our village people liked the TikaVaani program. Team should be sent to every place to spread information.”

Recruitment

The recruitment process was the same for the 13 villages; nonetheless, certain factors such as the level of education, time constraints, or interest of the participants limited this process.

C66: “Mobilizing men and women for meetings is the most challenging task.”

C44: “People did not have time to attend meetings. It was difficult to mobilize people to attend meetings.”

C66: [It was not easy] “Because some villagers were less educated, it was not easy.”

C33: “some of the villagers are educated people who motivate others to attend meetings and teach them how to use the Tika Vaani number. Due to which our work becomes much easier.”

However, the commitment and motivation of the implementers to carry out the different activities led them to develop strategies to attract participants and keep their attention during the planned activities:

C33: “We tried to organize meetings in a quiet place. If during the meeting some children created a disturbance, then one of the Tika Vaani team members took the children to a different location and did some activities with them.”

C77: “To keep the attention of participants, we used speakers and storytelling methods to disseminate information about Tiki Vaani.”

Despite the obstacles encountered to attract participants, implementers and frontline workers found that the intervention was gaining credibility among the participants.

C33: “There was a substantial change in people’s demeanor. Those people who did not trust Tika Vaani at the beginning of the pilot study developed faith in the program by the end and found the services beneficial.”

C88: “Behavior improved slowly over time.”

HW14 “Women who did not come for immunization, they now also come for vaccination after listening to the Tika Vaani number.”

Context

The contextual factors that influenced the fidelity of implementation were classified into two levels:

(i) Implementers: the presence of a motivated fieldwork team that understood the context, had experience in community work, and had the support of supervisors, was crucial to implementing the program.

(ii) At the community level, access to mobile phones, level of education, distance from home to meeting places, local weather conditions, and the quality of mobile networks were the main factors that explained the level of community participation and affected the level of fidelity.

C33: “We faced difficulties due to the lack of availability of mobile phones for women in the villages.”

C66: “In the villages, we sometimes met people who had taken alcohol. It was very challenging to make them understand.”

C77: “It was difficult to give information about TikaVaani to the less educated women of the villages.”

Community#12“problem is only of mobile network otherwise no problem was faced. We get information free of cost, no money deducted on incoming calls.”

R-FDBK “Today we cannot meet with men here. Because only a few days ago, the rain had occurred due to which all the men were engaged in the work of their fields.”

Control group:

The main study presents an evaluation of this dimension. According to the analysis of the records, only people who belonged to the intervention group attended the community meetings. The people who communicated with or received calls from the IVR platform were from the intervention group, except for one (1/166) control group member.

DISCUSSION

We used an established conceptual framework and a mixed methods design to study implementation fidelity as part of a randomized pilot trial. Fidelity assessment shed light on the trial interventions, demonstrating that the Tika Vaani intervention was implemented with high fidelity, supporting the conclusion that the study results faithfully reflected the underlying program theory and demonstrating acceptability and interest to participants. At a methodological level, our study corroborates the importance of assessing the fidelity of pilot projects before assessing whether activities are viable and optimize the intervention before promoting its large-scale implementation (Bruckenthal & Broderick, 2007).

Factors affecting and facilitating implementation fidelity in Tika Vaani Intervention

The modified conceptual framework for implementation fidelity adopted in this study proved useful for assessing fidelity and identifying factors that affect it. The integration of this framework in the study protocol enabled a structured assessment of the key dimensions of adherence (content, coverage, frequency, and duration) and comparison of targeted versus achieved levels of implementation. Moreover, it facilitated structured data collection and analysis of a rich range of interrelated moderating factors that together influenced fidelity. Among the moderating factors, recruitment, participant responsiveness, and context were especially relevant to understanding the level of coverage. The comprehensiveness of the policy description, strategies to facilitate implementation, and quality of delivery had more influence on the level of fidelity achieved. Consistent with other studies, quantitative and qualitative findings converged to reveal how some factors can affect implementation fidelity at different levels:

- Implementers of the intervention: The characteristics of implementers, especially their levels of knowledge, experience, motivation, and perception of the intervention, facilitated implementation and helped maintain the high level of fidelity (Bellg et al., 2004; Carroll et al., 2007; Durlak & DuPre, 2008; Wang et al., 2015). Additionally, the implementers in the communities gained the people's trust over time, thereby increasing their confidence in doing the job.

Beneficiaries: Availability of resources, geographic location, access to mobile phones, and the women's level of education were barriers to implementing activities and influenced participation in meetings and mHealth components. In this study, gender influenced the effectiveness of both components of the program.

- Program designers: A robust conceptual framework was useful to develop training modules, create manuals, hold periodic follow-up meetings, facilitate implementation, and maintain a high level of fidelity. Therefore, although complex interventions tend to have a lower level of fidelity (Gearing et al., 2011), fidelity can be maintained when programs are

well-founded and their theory is understood by implementers (Crawford, Freeman, Huscroft-D'Angelo, Quebec Fuentes, & Higgins, 2019).

The study of implementation fidelity allowed us to identify these and other programmatic and contextual factors that could affect the results of large-scale interventions (Bellg et al., 2004; Carroll et al., 2007; Hasson et al., 2012; Wang et al., 2015), and to take steps to strengthen the intervention design.

Considerations to improve the future intervention and optimize scale-up

Approaches focused on community participation and mHealth are known to improve childhood vaccination coverage (Johri, Pérez, et al., 2015; Kaufman et al., 2018; Oliver-Williams, Brown, Devereux, Fairhead, & Holeman, 2017; Oyo-Ita et al., 2016). However, more evidence is needed regarding mHealth in LMICs facing immunization barriers (Lee et al., 2016; Marcolino et al., 2018). This intervention showed that mHealth projects in conditions of limited resources might be useful in improving health knowledge and childhood vaccination (Johri et al., 2020). However, in the context of this study, it was not possible to reach the entire community using mobile phones alone. Combining community mobilization with access via mobile phone was essential to increase participation and reach the entire target population. Considering the context where the intervention is implemented is important to discover and understand the variations in the effects of the interventions (Lechat et al., 2019).

Developing context-based approaches combined with complementary actions that increase the probability of reaching different community groups and reduce health disparities is fundamental to achieve the Sustainable Development Goals (Frohlich & Potvin, 2008; Gorin, Badr, Krebs, & Das, 2012). Fidelity assessment allows anticipating changes and identifying adjustments that can be made before implementing large-scale interventions (Allen et al., 2012). Adjustments based on context should be made to improve program effectiveness and maintain a high level of fidelity (Blakely et al., 1987; Durlak & DuPre, 2008; Hasson et al., 2012). The following recommendations can improve the program and ensure intervention fidelity during scale-up:

Community mobilization components

- Local leaders are necessary to promote health education and community mobilization (Castle, Wendel, Pryor, & Ingram, 2017; Yajima, Takano, Nakamura, & Watanabe, 2001). Implementers found that the recruitment process was challenging. Strategies focused on identifying leaders that adopt recruitment strategies that favor community participation in the long term should be stressed to attract participants to join the different activities.

- Fathers play a significant role in improving the health of their children and stimulating childhood vaccination (Brugha, Kevany, & Swan, 1996). Parents who participated in our study acknowledged the usefulness of receiving health messages through a mobile phone and liked the small meetings. Scheduling meetings that favor paternal participation without interfering with work activities improve community involvement and increase awareness about the role of women in decision-making about children's health (Kumar et al., 2015). In turn, this approach could increase phone access to women through which they can receive information and increase coverage.

- Geographical location influenced the participation of women in meetings. Continuing to organize meetings in different geographical areas to reach more women will be fundamental to maximize the participation of mothers and empowering them in decision-making about childhood vaccination and increasing vaccination coverage (Thorpe et al., 2016).

- The role of health workers in promoting childhood vaccination in large-scale programs to improve the health of less favored populations is widely recognized (Lewin et al., 2010; Pallas et al., 2013). The involvement of these professionals is critical to developing contextualized work agendas that favor their recruitment and engagement over

time. Other strategies, such as certified training and non-monetary incentives, including free access to mobile phones or refreshments throughout the intervention period, can be adopted to stimulate health workers' involvement (Strachan et al., 2012).

Messages via mobile phone component

- Acceptance of technology to improve behavior is necessary to promote technological use (Dabbs et al., 2011; Huang, 2015). The use of mobile phones is crucial in large-scale interventions, and this is an opportunity to improve its functionality. For instance, given the active participation of frontline workers and their acceptance of mobile devices, this technology could be useful to provide training, facilitate data collection to plan vaccination schedules, and improve community health worker performance (Abejirinde, Ilozumba, Marchal, Zweekhorst, & Dieleman, 2018; Braun, Catalani, Wimbush, & Israelski, 2013; DeRenzi et al., 2012). This approach involves training, improving access to telephones, and developing policies that favor the exchange of information through mobile platforms. The involvement and commitment of different stakeholders can enhance the success of this strategy and optimize the activities of health workers as mediators between the health system and community (Braun et al., 2013; Strachan et al., 2012).

Strengths and Limitations

Given the complexity of public health interventions, this study reinforces the value of using mixed methods to assess the fidelity of implementation (Bellg et al., 2004; Palinkas et al., 2011; Proctor et al., 2011). The triangulation of multiple sources, data validation, and the rigorous methodology used for analyzing data led to a richer understanding of the findings and the generation of reliable recommendations for a large-scale study. This strategy involves organizing time and mobilizing resources, but the importance of implementing evidence-based actions justifies its planning to reduce costs of ineffective and unreliable programs over time (Bellg et al.,

2004). One main limitation using triangulation design is the discrepancies between different types of data. However, the framework used allows us to develop instruments for data collection in a complementary manner; therefore, findings were not contradictory, and minimal differences could be logically reconciled. Data were collected—by a researcher not involved in the project to avoid information bias—immediately after completing the program to avoid recall bias (Althubaiti, 2016). The analyses were conducted by an expert in the program, who did not participate in the implementation to ensure independence in the analysis and presentation of results.

The implementers who conducted the community interviews have delivered the mobilization component, possibly favoring a relationship with the participants around the intervention and social desirability bias (Althubaiti, 2016). However, the evaluation was formative, and the presence of implementers allowed examining the experiences of participants and the need for adjustments before starting a large-scale intervention. Fidelity was evaluated in a research context to determine short-term feasibility, but results may not be fully comparable with real-life situations to detect the factors affecting fidelity. However, the research context is an essential phase to assess program fidelity before conducting a large-scale study.

CONCLUSION

This study offers an example of how to conduct and report a study of intervention fidelity in the context of a pilot trial. The results showed that the fidelity of the implementation was high and highlighted the role of specific strategies to improve intervention coverage and reach at scale. Our findings reinforce the importance of opening the “black box” to understand the interactions that occur during the implementation phase in a better manner. Further, they demonstrate the importance of conducting fidelity assessment during pilot studies as a key element of the evaluation process to assess the viability of an intervention, refine it, and transfer it with the best possible evidence to other contexts.

Abbreviations

ASHA: Accredited Social Health Activist,

AWWs: *Anganwadi* Workers

AWW: Helpers (Sahaika)

CRT: Cluster Randomized Trial

IVR: Interactive Voice response mHealth - mobile health

SBCC: Social and Behaviour Change Communication

VHND: Village Health Nutrition Day

Declarations

Acknowledgements

We would like to thank the entire Tika Vaani Team for the warm welcome, and their energy, dedication and commitment during the process of this research. We also thank the community participants and frontline workers for their patience and time. We acknowledge the help of Pawan Kumar for conducting the interviews in the local language (Hindi) with field staff, to Kalpa Sharma for the translation of the data to English and to Frédérique Sauvé for her support with data management.

Authors' contribution

MCP led the drafting of all sections of this manuscript in conjunction with MJ. MCP performed the data abstraction and the data analysis. DC and GK assisted with the data management and cleaning.

MCP, DC, GK, RS, VR, MPS, AS and MJ all contributed to the critical revision and final approval of the manuscript to be published.

Funding

IC-IMPACTS (the India-Canada Centre for Innovative Multidisciplinary Partnerships to

Accelerate Community Transformation and Sustainability) provided funding for this study in the form of doctoral scholarships for MCP. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request

Ethics approval and consent to participate

The evaluation of this study is part of the Tika Vaani pilot study project that has received ethical approval from the by the Research Ethics Committee of the University of Montreal Hospital (CRCHUM), and the Institutional Committee for Ethics and Review for Research, IHMR, Jaipur, India. Informed consent of the participants was obtained before conducting the interviews and discussion groups.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Additional file 1,2,3 and 4 in annex to this thesis

Dossier supplémentaire 1, 2, 3 et 4 en annexe à cette thèse

Additional file 5 Additional information assessing fidelity of the intervention Tika Vaani

This document provides an overview of other findings illustrating fidelity assessment for each dimension according to the conceptual framework developed by Carroll et al. (2007) and modified by Hasson et al. (2010).

Fidelity Components	Summary of results by each component
Adherence	
Content	<p>Community mobilization strategy</p> <p>Two villages did not receive a <u>large introductory meeting</u> as planned, but the target families were visited directly to inform them about the intervention.</p> <p style="text-align: center;"><i>R-FDBK: “one (1) village there were only 4 households and, in another (1) village, the team tried to organize the meeting, but people didn’t come to the meeting.”</i></p> <p>Not all villages received the three series of small group meeting (fig a). To compare the experiences among those target families that lived near those who lived far and did not attend the meeting, they were visited at their homes:</p> <p style="text-align: center;"><i>FDBK: *For small group meeting#2: In two villages it was not conducted because in 1 village there was only 1 household and in the second village there were total 5 household in the village but 3 households had a long distance from each other and our team did individual visits to these households. Our team organized the 1st small group meeting with remaining 2 households because they were very near to each other but in the second small group meeting one target mother (the mother attended first small group meeting) was not present at her home during 2nd small group meeting and our team only did individual visit.</i></p> <p style="text-align: center;"><i>*For small group meeting#3: in one village had only 1 household and our team did individual visit.”</i></p>

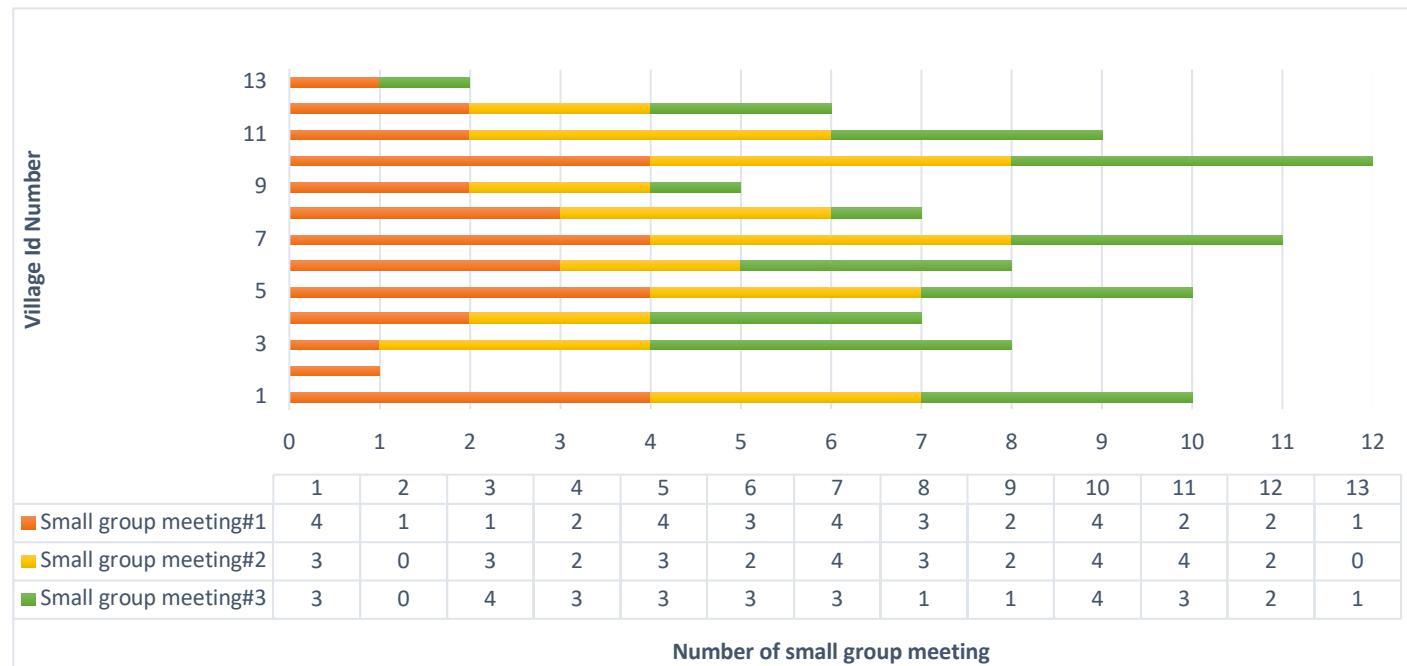
mHealth strategy

“Pushed” edutainment & summary capsules: they were delivered to people who indicated their interest and had a mobile phone to receive the message or to communicate with the platform.

Vaccination “reminders”: they were delivered to the target families, according to the age of each child and if the family had the mobile phone to receive the message.

“On-demand”: this activity comes from the initiative of each participant, so the number of calls was not planned. The platform, despite some technical / network difficulties at the beginning of the intervention, worked without problem during the intervention period.

Fig. a. Number of small group meeting held per village (n=13 villages)



<p>Coverage</p>	<p>Community mobilization strategy</p> <p>Community: The implementers did not record the target population during the large introductory meeting, so it is difficult to know precisely how many target family members have directly benefited from the introductory meeting.</p> <p>-The participation of the community in a small group meeting was represented by 250 women (123 mothers, 88 grandmothers and 39 aunts), 18 men (10 fathers, 6 grandparents and 2 uncles) and 16 children. The participation of the target population in the discussion groups varied during the intervention phase. Overall, 38% (70/184) of families participated in a single meeting, 15.2% (28/184) of families in two meetings while 21.7% (40/184) target families were not present at any meeting. 14% (26/184) of families benefit the all educational capsules as planned through small group meeting.</p> <p>Health workers: Their participation was low during the entire implementation phase:</p> <p>-small group meeting # 1: 33 meetings were held, and 30 health workers were invited (ASHAs = 10 AWW = 11 AWH = 9), of which only 5 workers attended (17%). For this first phase of the 33 meetings carried out, 5 meetings were attended by at least 1 health worker.</p> <p>- Small group meeting # 2: 32 small meetings were held, and 43 health workers were invited (ASHAs = 13 AWW = 16, AWH = 14) of which 12 (28%) attended. For this second phase, of the 32 meetings carried out, 9 meetings were attended by at least one health worker.</p> <p>-Small group meeting # 3: 31 small meetings were held and 36 health workers were invited (ASHAs = 10 AWW = 11, AWH = 15) of which 7 (27%) attended. For this third and final phase, of the 31 meetings held, 5 meetings were attended by at least one health worker</p> <p>mHealth strategy</p> <p>For Vaccination "reminders" 22% of target families (40/184) did not have a telephone to receive the message. Of the people who had a mobile phone, 80% (122/144) received at least one reminder message and found it useful:</p>
------------------------	---

	<p><i>Community#2 “there is no other method [reminding families on child vaccination through calls], phone is the best method. ASHA should come and take us”</i></p> <p><i>Community#12 “this information [reminding families on child vaccination through calls] is very helpful for all.”</i></p>
<p>Frequency and duration</p>	<p>Community mobilization strategy:</p> <p>The frequency of each discussion group took place every four weeks, once the period of four informative capsules were delivered to the community. The average duration of the meetings was between 40 to 60 minutes depending on the interactions and the different topics addressed</p> <p>mHealth strategy:</p> <p>There was no variation according to the schedule. Of the target families that had a portable cell phone, 56.9% (82/144) heard at least one OBD call (outbound dialing) with an average listening time per household of 151 minutes (range 0 min to 2180 min).</p> <p>37.5% (54/144) received at least one callback with a listening average of 39 minutes (range between 0 to 1333.7 minutes) per call. The capsule that was most heard by families were the pneumonia capsules.</p> <p>96.6% (28/29) of health workers have heard at least one capsule with a listening time of 80% or longer.</p>
<p>Moderating factors</p>	
<p>Comprehensiveness of intervention description</p>	<p><i>-The implementers believe in intervention theory and in the different activities proposed:</i></p>

	<p><i>C55: “Every child’s mother should have information about vaccination and on diseases such as dengue, diarrhoea, etcetera If all women understand everything and have full knowledge of vaccination, then our target will be achieved.”</i></p> <p><i>C33: “When we make people aware about the vaccination programme through the IVR [interactive voice response] and meetings, then (specific) people will definitely get their children vaccinated on time leading to development of a healthy society.” During small meetings, we learned that how much health information the participants got from Tika Vaani... “Creating awareness about health in the community through the medium of the IVR [interactive voice response] seems quite effective to me. We can do this on a big scale”</i></p> <p><i>C11: “By dialling the Tika Vaani number, people are provided with information on diseases. The Tika Vaani number is free. This is very good, people will receive information about children’s health. If children receive full vaccination, then they will not fall ill due to deadly diseases and will remain healthy.”</i></p> <p><i>C22: “Because through these approaches [doing meetings, providing information through home visits and mobile phone], we went to every corner and hamlet and informed people about the Tika Vaani programme so everyone received full information about this program.”</i></p> <p><i>C44: “If we keep on informing people in the village about vaccination in this way, then people will become aware.”</i></p>
<p>Strategies to facilitate implementation</p>	<p>The training received and the function’s manual were key elements that facilitated the implementation:</p> <p><i>C33 : “How to use Tika Vaani; How to send an item published on Tika Vaani to a friend or anyone in the community via mobile phone; I received training on the Tika Vaani callback system; How to talk to the Tika Vaani moderator... It was only by doing all these meetings and activities that our intervention was successful: Through meetings, we were introduced to the people; Through the IVR, people got information about</i></p>

vaccination; Through wall painting, everyone was made aware of the vaccination (Tika Vaani) number; ASHA and Anganwadi, so people gained trust in the Tika Vaani team ”

C77: “We developed a plan, made a village map and indicated selected households on it, and with the help of the map were able to reach these households”

C88: “[I] worked according to the instructions in the manual ... Because of these, it was easy to explain to people”

Regular meetings were held throughout the implementation phase to verify the planning, evaluate what was implemented and share the difficulties and ways to remedy them. In addition, they used personal records to plan the workday

C77: “We kept a diary and wrote any plan that we made in it”

The motivation of the implementers also influenced the implementation of the different activities:

C22: “I very much liked working in the field (village). In the village, I like discussing with the Pradhan (Mayor) and senior residents.”

C33: “Creating awareness about health in the community through the medium of the IVR [interactive voice response] seems quite effective to me. We can do this on a big scale”

C66: “I very much enjoy meeting and talking with the people in the villages”

C88: “I like to inform women about the health of children ”

An element that facilitated community participation was to find a quiet place and use strategies to capture the attention of the participants:

	<p><i>C33: “We tried to organise meetings in a quiet place. If during the meeting some children created a disturbance, then one of the Tika Vaani team members took the children to a different location and did some activities with them”</i></p> <p><i>C77: “To keep the attention of participants we used speakers to dissemination information about Tika Vaani and story telling methods”</i></p> <p><i>C88: “First of all, we talked to the women about their daily lives. Afterwards, we would do meetings with them. By this, they would listen carefully to us”</i></p> <p><i>C11: “When we asked questions, we used a conversational approach, so that people would not feel that we were asking questions”</i></p> <p><i>C66: “During meetings, when we saw that any woman was distracted, we immediately approached her and gave her information about Tika Vaani”</i></p>
<p>Quality of delivery</p>	<p>Before beginning the intervention, the Implementers were trained to balance their knowledge of the different issues addressed to the community.</p> <p>Knowing the user’s point of view is a way of assessing the quality of the intervention itself. The people who interacted with the platform were dialing the number, receiving call from the platform and who attended the group meetings stated that the information on the content of the different capsule and the way to access the platform was clear, understandable and useful. Likewise, the time in which calls were made to give information and give a reminder message to vaccinate the children was adequate and convenient and they considered it as a very useful strategy to attend the vaccination day.</p>

	<p>The health workers found that the information offered on the platform was useful, practical for their work and easy to understand and found no difficulty in entering the platform and following the instructions in each capsule.</p>
<p>Participant responsiveness</p>	<p>The response of the community and of the health workers was evaluated according to the perception, comprehension and acceptability of the intervention activities, as well as the level of participation in the different activities proposed.</p> <p>In order to improve women's access to the mobile phone we have asked participants what options could be considered to improve access to information through the platform. Involving men more is an important factor to consider</p> <p style="padding-left: 40px;">Community #13 <i>“women should explain their husband then only husbands allow women to listen the information”</i></p> <p style="padding-left: 40px;">Community #11 <i>“explain to men through mobile phones or explain them face to face”</i></p> <p style="padding-left: 40px;">Community #9 <i>“when older people listen storied then only, they will inform women”</i></p> <p style="padding-left: 40px;">Community #12 <i>“explain our husbands that whenever they get calls, give call to mother of the child”</i></p> <p style="padding-left: 40px;">Community” #15 <i>“put the phone near to the household members, eligible mothers will listen to it”</i></p> <p>The lack of time / interest was evoked by some to not participate in community meetings</p> <p style="padding-left: 40px;">Community#8 <i>“did not feel the need and did not get the time”</i></p> <p style="padding-left: 40px;">Community#15 <i>“we do not have time, we do not need this, we do not have money. We do embroidery work and save money, so do not get time, also not have mobile”</i></p>

Although there is a better understanding of the importance of childhood vaccination, this is an issue that still presents certain prejudices in some people in the community.

Community 8: “An eligible mother said that if they go to vaccinated their child her husband scold to her mother-in-law”

Community 12 [If a child has been given all the vaccines but still, the child is weak and frequent fall ill, then what is the benefit of vaccination?] “No benefits of vaccination go to doctor”

The Implementers of the intervention agree that increasing the number of visits to communities and including other health issues can improve not only participation but also the intervention itself

C44: “There should be capsules to provide information on different diseases. Each village should be visited at least once per month”

C66: “More visits to the community, information capsules about different diseases”

Finally, health workers consider the mobile platform a good strategy to reach the community and other health workers and also suggest that other issues can be attached to be presented to the community

HW8 “Yes people will get information on vaccination through this number and they will come for vaccination”

HW9 “Women in the village ask when their child will get vaccines. Now started giving attention on vaccination and cleanliness”

HW3 “now vaccination is happening successful people take their children immediately if their child is sick... Now

	<p><i>everyone is coming for vaccination”</i></p> <p><i>HW1 “all Anganwadi, ANMs will get support”</i></p> <p><i>HW7 “ASHA have to go house to house they will get good information”</i></p> <p>In the same way, the workers express interest to receive training on topics such as child vaccination and family planning and propose as strategies for learning the use of mobile phones, classrooms and video projection. Particularly 76% (13/17) of the health workers interviewed are motivated and would like to receive training through the mobile platform, who do not agree to receive it through this form expressed as the main reason for not having a telephone (n = 2) or because they don't have time (n = 2).</p>
Recruitment	<p>The field workers visited each target family to invite them to the different meetings, provide them with educational information and to motivate them to communicate with the platform. Both field workers and health workers believe that the participation of the latter can facilitate the recruitment and participation of the community</p> <p><i>HW10: “Whenever a meeting is organized, the village head, ASHA [Accredited Social Health Activist], Anganwadi worker should be present in the meeting so that women could make aware”</i></p>
Context	<p>The lack of telephone was an important factor that influenced the access of information through this medium:</p> <p><i>Community#7 “our son has mobile phone, so we were not able to listen”</i></p> <p><i>Community#8 “my husband keeps mobile phones, so we do not get information”</i></p> <p><i>C77: “Women do not have mobile phones. We faced difficulties when we asked them to dial a number”</i></p>

	<p>Another contextual factor that influenced the reception of educational messages is the fact of extensive working hours or rainy seasons during the intervention period that did not allow, especially parents, to participate</p> <p><i>R-FDBK:</i></p> <p><i>“Men's meeting could not be held because the market was started in this village today...”</i></p> <p><i>“Meeting with men was not organized because all the men were busy working on their farm.”</i></p> <p>Apparently, the fact that the platform used the same information at the beginning of each capsule could have been a factor that explains why some parents received the call, but did not listen to the story completely</p> <p><i>Community 3: “Because of same introduction in every capsule an eligible mother used to disconnect TikaVaani call because she thought that she was getting same story.”</i></p> <p>Finally, people change the number very quickly when they receive a text message offering them a new, cheaper card. Then the numbers that were initially registered to receive the information over time are invalid numbers to receive a call, a factor that may explain why some parents who have a telephone did not receive the information as planned.</p>
Grupo control	The control group received the mobile phone components (OBD, Calback, Reminders) of the intervention one month after the implementation period between September and December 2018 ended.

REFERENCES

- Abejirinde, I.-O. O., Ilozumba, O., Marchal, B., Zweekhorst, M., & Dieleman, M. (2018). Mobile health and the performance of maternal health care workers in low- and middle-income countries: A realist review. *21*(3), 73-86. doi:10.1177/2053434518779491
- Allen, J. D., Linnan, L. A., Emmons, K. M., Brownson, R., Colditz, G., Proctor, E. J. D., & practice, i. r. i. h. T. s. t. (2012). Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination. 281-304.
- Althubaiti, A. J. J. o. m. h. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *9*, 211.
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Treatment Fidelity Workgroup of the, N. I. H. B. C. C. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*, *23*(5), 443-451. doi:10.1037/0278-6133.23.5.443
- Blakely, C. H., Mayer, J. P., Gottschalk, R. G., Schmitt, N., Davidson, W. S., Roitman, D. B., & Emshoff, J. G. J. A. j. o. c. p. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *15*(3), 253-268.
- Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent*, *71*(s1), S52-63.
- Braun, R., Catalani, C., Wimbush, J., & Israelski, D. J. P. o. (2013). Community health workers and mobile technology: a systematic review of the literature. *8*(6), e65772.
- Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Res Nurs Health*, *33*(2), 164-173. doi:10.1002/nur.20373
- Brownson, R. C., Fielding, J. E., & Maylahn, C. M. (2009). Evidence-Based Public Health: A Fundamental Concept for Public Health Practice. *30*(1), 175-201. doi:10.1146/annurev.publhealth.031308.100134
- Bruckenthal, P., & Broderick, J. E. (2007). Assessing treatment fidelity in pilot studies assist in designing clinical trials: an illustration from a nurse practitioner community-based intervention for pain. *ANS Adv Nurs Sci*, *30*(1), E72-84. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/17299277>
- Brugha, R., Kevany, J., & Swan, A. J. I. J. o. E. (1996). An investigation of the role of fathers in immunization uptake. *25*(4), 840-845.
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ*, *321*(7262), 694-696.

- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci*, 2, 40. doi:10.1186/1748-5908-2-40
- Castle, B., Wendel, M., Pryor, B. N. K., Ingram, M. J. J. o. P. H. M., & Practice. (2017). Assessing Community Leadership: Understanding Community Capacity for Health Improvement. 23, S47-S52.
- Chakraborty, D., Gupta, A., & Seth, A. (2019). *Experiences from a mobile-based behaviour change campaign on maternal and child nutrition in rural India*. Paper presented at the Proceedings of the Tenth International Conference on Information and Communication Technologies and Development.
- Chandra, T. (2019). Literacy in India: The Gender and Age Dimension. *Literacy*.
- Claudia Bojanowski*, Dinesh Chandra, Aaditeshwar Seth, Alok K Mathur, Mira Johri, & Collaboration, o. b. o. t. T. V. (2018). Learning from context: Formative research to enhance intervention design for Tika Vaani, a m-health platform to improve child health in rural India. Oral presentation. *the 2018 International Social and Behavior Change Communication (SBCC) Summit Featuring Entertainment Education, 2nd edition, April 16th to 20th, 2018; Nusa Dua, Bali, Indonesia*.
- Cohen, D. J., Crabtree, B. F., Etz, R. S., Balasubramanian, B. A., Donahue, K. E., Leviton, L. C., . . . Green, L. W. J. A. j. o. p. m. (2008). Fidelity versus flexibility: translating evidence-based research into practice. 35(5), S381-S389.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. 337, a1655. doi:10.1136/bmj.a1655 BMJ
- Crawford, L., Freeman, B., Huscroft-D'Angelo, J., Quebec Fuentes, S., & Higgins, K. N. J. L. D. Q. (2019). Implementation fidelity and the design of a fractions intervention. 0731948719840774.
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*: Sage publications.
- Dabbs, A. D., Song, M.-K., Hawkins, R., Aubrecht, J., Kovach, K., Terhorst, L., . . . Callan, J. J. N. r. (2011). An intervention fidelity framework for technology-based behavioral interventions. 60(5), 340.
- DeRenzi, B., Findlater, L., Payne, J., Birnbaum, B., Mangilima, J., Parikh, T., . . . Lesh, N. (2012). *Improving community health worker performance through automated SMS*. Paper presented at the Proceedings of the Fifth International Conference on Information and Communication Technologies and Development.

- Durlak, J. A., & DuPre, E. P. J. A. j. o. c. p. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *41*(3-4), 327-350.
- Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. In: Springer.
- Frohlich, K. L., & Potvin, L. (2008). Transcending the known in public health practice: the inequality paradox: the population approach and vulnerable populations. *Am J Public Health, 98*(2), 216-221. doi:10.2105/AJPH.2007.114777
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. J. B. m. r. m. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *13*(1), 117.
- Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. J. C. p. r. (2011). Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *31*(1), 79-88.
- Glasgow, R. E., & Emmons, K. M. J. A. R. P. H. (2007). How can we increase translation of research into practice? Types of evidence needed. *28*, 413-433.
- Gorin, S. S., Badr, H., Krebs, P., & Das, I. P. (2012). Multilevel Interventions and Racial/Ethnic Health Disparities. *JNCI Monographs, 2012*(44), 100-111. doi:10.1093/jncimonographs/lgs015 %J JNCI Monographs
- Gurnani, V., Haldar, P., Aggarwal, M. K., Das, M. K., Chauhan, A., Murray, J., . . . Sudan, P. J. B. (2018). Improving vaccination coverage in India: lessons from Intensified Mission Indradhanush, a cross-sectoral systems strengthening strategy. *363*, k4782.
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci, 5*, 67. doi:10.1186/1748-5908-5-67
- Hasson, H., Blomberg, S., & Dunér, A. (2012). Fidelity and moderating factors in complex interventions: a case study of a continuum of care program for frail elderly people in health and social care. *Implement Sci, 7*(23), 1–11.
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ, 348*, g1687. doi:10.1136/bmj.g1687
- Huang, Y.-M. J. A. J. o. E. T. (2015). Exploring the factors that affect the intention to use collaborative technologies: The differing perspectives of sequential/global learners. *31*(3).
- IIPS. (2016). National Family Health Survey (NFHS-4), 2015–16: India: . Retrieved from <http://rchiips.org/nfhs/pdf/NFHS4/India.pdf>
- Johri, M., Chandra, D., Koné, G. K., Dudeja, S., Sylvestre, M.-P., Sharma, J. K., & Pahwa, S. (2015). Interventions to increase immunisation coverage among children 12–23 months of

age in India through participatory learning and community engagement: pilot study for a cluster randomised trial. *5*(9), e007972. doi:10.1136/bmjopen-2015-007972 J BMJ Open

- Johri, M., Chandra, D., Koné, G. K., Sylvestre, M.-P., Mathur, A. K., Harper, S., & Nandi, A. (2020). Social and behaviour change communication interventions delivered face-to-face and by mobile phone to strengthen vaccination uptake and improve child health in rural India: A randomised pilot study for a cluster randomised controlled trial. *Under review*.
- Johri, M., Pérez, M. C., Arsenault, C., Sharma, J. K., Pai, N. P., Pahwa, S., & Sylvestre, M.-P. J. B. o. t. W. H. O. (2015). Strategies to increase the demand for childhood vaccination in low-and middle-income countries: a systematic review and meta-analysis. *93*, 339-346.
- Kaufman, J., Ryan, R., Walsh, L., Horey, D., Leask, J., Robinson, P., & Hill, S. J. C. D. o. S. R. (2018). Face-to-face interventions for informing or educating parents about early childhood vaccination. (5).
- Kumar, N., Perrier, T., Desmond, M., Israel-Ballard, K., Kumar, V., Mahapatra, S., . . . Lal, P. (2015). *Projecting health: community-led video education for maternal health*. Paper presented at the Proceedings of the Seventh International Conference on Information and Communication Technologies and Development.
- Layde, P. M., Christiansen, A. L., Peterson, D. J., Guse, C. E., Maurana, C. A., & Brandenburg, T. J. A. j. o. p. h. (2012). A model to translate evidence-based interventions into community practice. *102*(4), 617-624.
- Lechat, L., Bonnet, E., Queuille, L., Traore, Z., Some, P.-A., & Ridde, V. (2019). Relevance of a Toll-free call service using an interactive voice server to strengthen health system governance and responsiveness in Burkina Faso. *International Journal of Health Policy and Management*, *8*(6), 353-364.
- Lee, S. H., Nurmatov, U. B., Nwaru, B. I., Mukherjee, M., Grant, L., & Pagliari, C. (2016). Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Systematic review and meta-analysis. *J Glob Health*, *6*(1), 010401. doi:10.7189/jogh.06.010401
- Lewin, S., Munabi-Babigumira, S., Glenton, C., Daniels, K., Bosch-Capblanch, X., van Wyk, B. E., . . . Zwarenstein, M. J. C. d. o. s. r. (2010). Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases. (3).
- Marcolino, M. S., Oliveira, J. A. Q., D'Agostino, M., Ribeiro, A. L., Alkmim, M. B. M., & Novillo-Ortiz, D. (2018). The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. *JMIR Mhealth Uhealth*, *6*(1), e23. doi:10.2196/mhealth.8873
- Medical Research, & Council. (2008). *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council.
- Miles, M. B., & Huberman, A. M. (2003). *Analyse des données qualitatives*: De Boeck Supérieur.

- Moitra, A., Das, V., Vaani, G., Kumar, A., & Seth, A. (2016). *Design lessons from creating a mobile-based community media platform in Rural India*. Paper presented at the Proceedings of the Eighth International Conference on Information and Communication Technologies and Development.
- Noell, G. H., Gresham, F. M., & Gansle, K. A. J. J. o. B. E. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *11(1)*, 51-67.
- O'cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research & Policy*, *13(2)*, 92-98.
- Ofori-Amaah, S. (1989). Social mobilization for immunization and primary health care. *Asia Pacific Journal of Public Health*, *3(3)*, 200-204.
- Oliver-Williams, C., Brown, E., Devereux, S., Fairhead, C., & Holeman, I. (2017). Using Mobile Phones to Improve Vaccination Uptake in 21 Low- and Middle-Income Countries: Systematic Review. *JMIR Mhealth Uhealth*, *5(10)*, e148. doi:10.2196/mhealth.7792
- Organization, W. H. (2011). mHealth: new horizons for health through mobile technologies. *mHealth: new horizons for health through mobile technologies*.
- Oyo-Ita, A., Wiysonge, C. S., Oringanje, C., Nwachukwu, C. E., Oduwole, O., & Meremikwu, M. M. J. C. D. o. S. R. (2016). Interventions for improving coverage of childhood immunisation in low-and middle-income countries. (7).
- Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlain, P., Hurlburt, M., Landsverk, J. J. A., . . . Research, M. H. S. (2011). Mixed method designs in implementation research. *38(1)*, 44-53.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., Hoagwood, K. J. A., . . . Research, M. H. S. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *42(5)*, 533-544.
- Pallas, S. W., Minhas, D., Pérez-Escamilla, R., Taylor, L., Curry, L., & Bradley, E. H. J. A. j. o. p. h. (2013). Community health workers in low-and middle-income countries: what do we know about scaling up and sustainability? , *103(7)*, e74-e82.
- Paul, G. B., & Murti, A. B. (2016). Socio-economy of Mobile Phone Ownership in India. In *Technology* (pp. 157-168): Springer.
- Perez, M. C., Minoyan, N., Ridde, V., Sylvestre, M. P., & Johri, M. (2018). Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review. *Trials*, *19(1)*, 410. doi:10.1186/s13063-018-2796-z

- Plano Clark, V. L., Anderson, N., Wertz, J. A., Zhou, Y., Schumacher, K., & Miaskowski, C. J. J. o. M. M. R. (2015). Conceptualizing longitudinal mixed methods designs: A methodological review of health sciences research. *9*(4), 297-319.
- Pluye, P., Bengoechea, E. G., & Granikov, V. J. I. J. M. R. A. (2018). A world of possibilities in mixed methods: review of the combinations of strategies used to integrate qualitative and quantitative phases, results and data. *10*, 41-56.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., . . . Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*, *38*(2), 65-76. doi:10.1007/s10488-010-0319-7
- Resnick, B., Bellg, A. J., Borrelli, B., De Francesco, C., Breger, R., Hecht, J., . . . Ernst, D. J. A. o. B. M. (2005). Examples of implementation and evaluation of treatment fidelity in the BCC studies: where we are and where we need to go. *29*(2), 46.
- Richards, D. A., & Hallberg, I. R. (2015). Complex interventions in health: an overview of research methods. *London & New York: Routledge*.
- Ritchie, J., & Spencer, L. J. T. q. r. s. c. (2002). Qualitative data analysis for applied policy research. *573*(2002), 305-329.
- Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*, *56*, 119–127. doi:doi:10.1136/jech.56.2.119
- Sandelowski, M. J. R. i. n., & health. (1995). Sample size in qualitative research. *18*(2), 179-183.
- Saré, D., Pérez, D., Somé, P.-A., Kafando, Y., Barro, A., & Ridde, V. (2018). Community-based dengue control intervention in Ouagadougou: intervention theory and implementation fidelity. *Global health research and policy*, *3*(1), 21.
- Shrivastwa, N., Wagner, A. L., & Boulton, M. L. (2019). Analysis of State-Specific Differences in Childhood Vaccination Coverage in Rural India. *Vaccines*, *7*(1), 24.
- Sondaal, S. F. V., Browne, J. L., Amoakoh-Coleman, M., Borgstein, A., Miltenburg, A. S., Verwijs, M., & Klipstein-Grobusch, K. (2016). Assessing the effect of mHealth interventions in improving maternal and neonatal care in low-and middle-income countries: a systematic review. *PLoS One*, *11*(5).
- Strachan, D. L., Källander, K., ten Asbroek, A. H., Kirkwood, B., Meek, S. R., Benton, L., . . . hygiene. (2012). Interventions to improve motivation and retention of community health workers delivering integrated community case management (iCCM): stakeholder perceptions and priorities. *87*(5_Suppl), 111-119.

- Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., . . . Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol*, *10*, 1. doi:10.1186/1471-2288-10-1
- Thorpe, S., VanderEnde, K., Peters, C., Bardin, L., Yount, K. M. J. M., & Journal, C. H. (2016). The Influence of Women's Empowerment on Child Immunization Coverage in Low, Lower-Middle, and Upper-Middle Income Countries: A Systematic Review of the Literature. *20*(1), 172-186. doi:10.1007/s10995-015-1817-8
- Wang, B., Stanton, B., Deveaux, L., Poitier, M., Lunn, S., Koci, V., . . . Li, X. J. I. s. (2015). Factors influencing implementation dose and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program. *10*(1), 44.
- Yajima, S., Takano, T., Nakamura, K., & Watanabe, M. (2001). Effectiveness of a community leaders' programme to promote healthy lifestyles in Tokyo, Japan. *Health Promot Int*, *16*(3), 235-243. doi:10.1093/heapro/16.3.235
- Yin, R. K. (2017). *Case study research and applications: Design and methods*: Sage publications.

Chapitre 4 – Discussion

La présente thèse avait pour objectif d'examiner la fidélité des interventions en santé publique implémentées dans des pays du Sud et évaluées sous le modèle des essais randomisés en grappes, afin de déterminer si les interventions mises en place sous un contexte contrôlé doivent tenir compte de ce type d'évaluation.

Donc, l'objectif de cette thèse visait deux volets de recherche : le premier volet consistait à examiner systématiquement la littérature scientifique pour étudier les pratiques actuelles concernant l'évaluation de la fidélité de l'intervention dans les essais randomisés en grappes des interventions de santé publique mises en place dans les pays du Sud. Pour ce volet, nous avons mené une revue systématique pour laquelle nous avons élaboré et publié le protocole (Article 1A) et les résultats de cette révision (Article 1B). Ensuite, afin de répondre à un besoin des concepteurs de l'intervention Tika Vaani pour évaluer la fidélité de la mise en œuvre de l'intervention, le deuxième volet de recherche évaluative cherchait à fournir des résultats utiles à l'amélioration de l'intervention avant de l'appliquer à grande échelle. Pour ce volet, nous avons mené l'évaluation de la fidélité de la mise en œuvre de l'intervention Tika Vaani à travers une étude de cas unique à trois unités d'analyse, selon une approche mixte concomitante. Les résultats de cette évaluation sont présentés dans deux des trois articles (Article 2 et Article 3) dans le Chapitre des résultats. L'Article 2 permet de mieux comprendre son fondement théorique, le développement de l'intervention ainsi que le contexte dans lequel l'intervention s'inscrit. L'Article 3 présente l'évaluation de la fidélité de la mise en œuvre de l'intervention Tika Vaani, ainsi que des recommandations pour l'améliorer.

Des enseignements, tirés du point de vue méthodologique concernant l'évaluation de la fidélité de la mise en œuvre des interventions en santé publique, et issus de deux volets de recherche, sont présentés dans cette section. Ensuite, la contribution empirique de cette pratique évaluative pour renforcer la recherche dans les sciences de l'implémentation est abordée. Enfin, ce chapitre aborde une réflexion sur les forces et les limites de la thèse, ainsi que sur les pistes d'action pour la recherche.

4.1 Synthèse, contribution des résultats et pistes d'action

Globalement, les résultats de cette thèse permettent de constater que l'évaluation de la fidélité de la mise en œuvre des interventions en santé publique sous la conception des essais randomisés en grappes, mises en place dans des pays du Sud, n'est pas considérée systématiquement au moment de développer une intervention, ni au moment de l'évaluer (M. C. Perez et al., 2018). Dans la révision systématique, 90 essais randomisés en grappes qui répondaient à nos critères d'inclusion ont été identifiés. Parmi ces 90 études, 25 (28%) ne prévoyaient pas l'évaluation de la fidélité de la mise en œuvre de l'intervention, et les 65 études restantes (72%) traitaient au moins d'une dimension de la fidélité de la mise en œuvre de l'intervention. Les résultats font ressortir que, concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention qui n'est pas systématiquement faite dans ce type d'étude, la qualité de l'évaluation et du rapport est très hétérogène, et est souvent sans méthode bien définie pour ce type d'étude. Les dimensions plus fréquemment mesurées visent l'adhésion, en mesurant la couverture et le contenu de l'intervention à partir d'une liste de contrôle, de consultations de manuels ou de notes de terrain. Malgré l'importance d'explorer les facteurs qui modèrent la fidélité de la mise en œuvre de l'intervention, ils sont moins évalués ou rapportés dans les études ciblées par la revue systématique.

Bien que : (i) les interventions de santé publique peuvent présenter des variations pendant leur mise en œuvre, et occasionner des écarts de ce qui a été planifié en compromettant la validité des conclusions ; et (ii) l'évaluation de la fidélité de la mise en œuvre de l'intervention peut aider à évaluer l'ampleur de ces problèmes, et à atténuer les effets négatifs, nous constatons que les lignes directrices actuelles, pour le rapport de ce type d'étude, ne donnent aucune indication sur l'évaluation de la fidélité de la mise en œuvre de l'intervention. En effet, le groupe CONSORT fournit des directives sur la façon d'améliorer le rapport des essais randomisés en grappes en élaborant des guides pour rapporter les résultats de façon claire, complète et transparente (M. K. Campbell, Elbourne, Altman, & group, 2004; M. K. Campbell, Piaggio, Elbourne, Altman, & Group, 2012; Moher et al., 2010). Cependant, ces lignes directrices n'offrent aucun conseil sur les méthodes permettant d'évaluer la fidélité de la mise en œuvre de l'intervention, et ne l'exigent pas dans la publication des résultats. Donc, un outil qui facilite la planification, l'évaluation et le

rapport de cette pratique dans ce type d'étude s'avère utile voire indispensable.

Impliquer les acteurs concernés, avec une méthode structurée de consultation, dans l'élaboration des outils permet de dégager des informations plus précises. Ces informations peuvent être utilisées pour développer des guides pratiques pouvant aider les chercheurs à mieux orienter leur pratique. Des outils comme le TIDieR, outil pour la description et la réplique des interventions (Hoffmann et al., 2014), outil actualisé par l'outil TIDieR-PHP pour étendre la portée aux interventions en matière de santé et des politiques de la population (M. Campbell et al., 2018), l'ASTAIRE, outil d'analyse de la transférabilité et de l'accompagnement à l'adaptation des interventions en promotion de la santé (Cambon, Minary, Ridde, & Alla, 2013), et le SAGE, outil pour évaluer l'engagement des décideurs à utiliser la recherche dans l'élaboration des politiques de santé (R. Makkar et al., 2016), sont des exemples largement utilisés dans la littérature et dans le domaine de la santé publique. Ces outils ont été construits et validés par des acteurs concernés par les pratiques d'intervention développées et des collègues d'experts dans le domaine de l'évaluation. Donc, l'implication des participants concernés facilite la construction d'un outil fiable, pratique et adapté aux interventions en santé publique. De plus, cette implication est un facteur qui favorise leur utilisation, car ces outils répondent aux préoccupations des utilisateurs qui mènent la recherche en milieu de pratique (Dagenais, 2006).

À la suite des résultats de la révision systématique et afin de mieux comprendre les pratiques, les enjeux et les perspectives actuelles des chercheurs concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention, les 90 auteurs principaux de chaque étude identifiée dans notre revue systématique ont été invités à participer pour connaître davantage leur point de vue, concernant cette pratique évaluative à travers un court entretien par téléphone ou par courriel. Donc, pour disséminer les principaux résultats de la révision, une note de synthèse (*policy brief*) a été élaborée et envoyée aux chercheurs (Annexe 9). Cette extension de la révision systématique a été exemptée de l'approbation éthique par le comité d'éthique de la recherche du Centre Hospitalier de l'Université de Montréal (CRHUM). En effet, nous avons contacté la présidente du comité d'éthique de la recherche et nous avons expliqué notre intérêt à contacter les auteurs de notre revue systématique et le but des entretiens. Elle nous a autorisée à procéder sans besoin de déposer une demande d'approbation éthique.

Les questions ainsi que les résultats d'entretien faits auprès des auteurs de notre revue systématique sont présentés dans l'annexe 10. À partir des réponses des participants, nous avons conclu de façon globale que : (i) les chercheurs sont familiarisés avec le concept, et ils trouvent très utile et important de faire l'évaluation de la fidélité de la mise en œuvre de l'intervention dans ce type d'intervention. Cependant, ils ont souligné la difficulté pour bien comprendre la façon de procéder pour le faire, car il n'existe pas de guides précis qui les orientent; en plus, (ii) les chercheurs suggèrent qu'un guide pratique sous forme de liste de contrôle, appuyé par un document de référence fournissant des explications détaillées point par point, serait des plus utile pour adopter cette pratique évaluative dans de futures études (Annexe 10). Donc, le besoin d'élaborer un guide qui facilite la compréhension pour élaborer et pour mener une telle évaluation, afin d'améliorer cette pratique et ainsi bonifier les résultats de l'intervention, est bien réel.

4.2 Contribution méthodologique concernant l'évaluation de la fidélité de la mise en œuvre de l'intervention

L'état des connaissances a mis en relief que les essais randomisés en grappes sont largement utilisés dans l'évaluation des interventions du système de santé, compte tenu de leur importance en tant que source de données probantes pour établir le lien de causalité d'une intervention au sein de la population cible. Ces interventions sont encore plus complexes, car différents facteurs peuvent favoriser des écarts entre ce qui a été planifié et sa mise en œuvre (Grant et al., 2013; Rychetnik et al., 2002). Des écarts peuvent se ressentir encore plus dans les pays du Sud étant donné la pénurie des ressources économiques, logistiques et humaines à laquelle ces pays sont confrontés (Valéry Ridde & Dagenais, 2019). Pour atténuer ces problèmes méthodologiques, qui peuvent affecter la validité interne et externe des résultats de ce type d'intervention, l'évaluation de la fidélité de la mise en œuvre de l'intervention devient un outil clé à favoriser pour toute évaluation d'une intervention.

Les résultats qui émergent de cette thèse offrent une contribution méthodologique à la pratique évaluative concernant la fidélité de la mise en œuvre de l'intervention sur plusieurs aspects.

Premièrement, la fidélité de la mise en œuvre de l'intervention reflète l'adhésion au contenu, à la fréquence, à la durée et à la couverture de l'intervention. Des facteurs modérateurs pouvant influencer le degré de la fidélité de la mise en œuvre de l'intervention sont également importants à explorer (Carroll et al., 2007). Cependant, les résultats qui émergent de cette pratique évaluative ont permis de constater que *les évaluations de la fidélité de la mise en œuvre de l'intervention examinent rarement l'évaluation complète de tous les facteurs qui peuvent l'influencer et se concentrent souvent dans la mesure de l'adhésion* (M. C. Perez et al., 2018). En revanche, les résultats issus de cette recherche mettent en lumière la *difficulté pour mesurer la fidélité de la mise en œuvre de l'intervention au niveau de l'adhésion*, difficulté constatée également par d'autres chercheurs dans des interventions sur la promotion de la santé (Dusenbury et al., 2003). *Quant aux facteurs qui modèrent le niveau de fidélité atteint, à notre connaissance,*

c'est la première étude qui explore tous ces facteurs dans des interventions de santé publique dans le cadre des essais randomisés en grappes. Les facteurs qui modèrent le degré de fidélité obtenu sont importants voire fondamentaux à explorer dans ce type d'intervention, car ils offrent des informations riches pour mieux comprendre les résultats obtenus (Breitenstein et al., 2010; Hasson et al., 2012). Cependant, ces interventions n'évaluent pas systématiquement ces facteurs, bien qu'elles soient étroitement liées et dépendantes des facteurs contextuels favorisant des écarts au protocole et en affectant la fidélité de la mise en œuvre de l'intervention (M. C. Perez et al., 2018). Ainsi, les chercheurs devront considérer chercher des informations supplémentaires qui vont au-delà de l'évaluation de l'efficacité pour décrire ce qui se passe réellement dans un essai randomisé en grappes, et pour explorer des hypothèses sur les raisons expliquant que les interventions sont efficaces ou inefficaces (Flottorp, Håvelsrud, & Oxman, 2003).

Deuxièmement, nos résultats suggèrent qu'une plus grande considération et une plus grande importance devraient être accordées à la compréhension sur la façon d'aborder la fidélité de la mise en œuvre par les concepteurs des interventions, au-delà de se concentrer à évaluer l'efficacité de l'intervention. Bien que les chercheurs qui ont mené des interventions en santé publique sous le cadre des essais randomisés en grappes considèrent l'importance d'évaluer la fidélité de la mise en œuvre au sein des interventions, *les chercheurs soulèvent la difficulté pour savoir comment aborder cette évaluation dans ces études.* En effet, malgré la diversité sur la conceptualisation de la fidélité de la mise en œuvre de l'intervention, le « quoi » et le « pourquoi » de cette évaluation concernant ses avantages et sa force pour renforcer les résultats d'une étude sont bien documentés dans la littérature. Cependant, bien que des cadres des références pour évaluer la fidélité de la mise en œuvre de l'intervention existent, nous constatons que le « comment » est plus hétérogène, car *il n'y a pas de directives spécifiques sur la façon de procéder et les auteurs de ces interventions questionnés reconnaissent qu'il existe peu d'orientations normalisées sur la façon d'opérationnaliser chaque dimension.* Ce manque de compréhension peut mener les chercheurs à éviter de faire ce type d'évaluation, et ainsi de ne pas profiter de tous les avantages que cette évaluation apporte. Une compréhension éclairée est fondamentale quand les ressources sont limitées, car l'évaluation de la fidélité de la mise en œuvre de l'intervention implique du temps, des coûts et de la main-d'œuvre pour s'engager.

Nous constatons la présence des lignes directrices pour mieux rapporter les résultats selon la conception de l'étude, mais il n'existe pas de lignes directrices qui exigent de faire et de déclarer systématiquement l'évaluation de la fidélité de la mise en œuvre de l'intervention pour fournir une image plus complète et précise de l'efficacité du programme. Le guide TIDieR-PHP exige des informations à mentionner dans la description des interventions en matière de santé et des politiques de la population dans les rapports d'études d'évaluation (M. Campbell et al., 2018). Parmi ces informations, il fait référence à l'importance de mentionner si l'adhésion à l'intervention a été évaluée, en décrivant comment, par qui et quelles stratégies ont été utilisées pour préserver ou améliorer la fidélité de la mise en œuvre de l'intervention. Cependant, il ne fournit pas d'instructions aux auteurs sur la façon de procéder.

Pour combler cette lacune dans le domaine de l'implémentation, qui exige de plus en plus des approches fondées sur la base des meilleures preuves scientifiques disponibles à partir, entre autres, d'une évaluation solide de la fidélité de la mise en œuvre d'une intervention (Moir, 2018; Titler, 2008), un outil pratique et adaptable aux interventions se révèle nécessaire. Bien que nous mettions l'accent sur les essais randomisés en grappes, car ils sont considérés par la médecine fondée sur les évidences, comme faisant partie des meilleurs moyens d'évaluer les effets d'une intervention, il est important de reconnaître que des problèmes similaires se posent dans l'évaluation des interventions de santé publique, qui peuvent utiliser une gamme de modèles d'évaluation différents. En effet, les interventions de santé publique peuvent être influencées par un ensemble des facteurs internes et externes à l'intervention qui vont contribuer au succès ou à l'échec de l'intervention. Des variations peuvent survenir suite aux interactions des acteurs locaux et des acteurs nationaux, voire internationaux, auxquels sont confrontés les pays du Sud (Craig et al., 2008; Valéry Ridde & Capelle, 2011; Valéry Ridde & Dagenais, 2019). Donc, une approche plus structurée, mais orientée sur la conception de l'évaluation de la fidélité de la mise en œuvre, va bonifier la validité et la diffusion du résultat de ces études, afin que ces interventions puissent se reproduire en pratique dans d'autres contextes. *Dans le but de combler ce besoin, nous suggérons un outil pour aborder cette évaluation qui sera présenté plus en détail dans la section suivante (4.2.1).*

Troisièmement, les résultats issus de l'étude de cas font ressortir l'importance *d'aborder l'évaluation de la fidélité de la mise en œuvre de l'intervention dès la planification d'une intervention et de l'aborder sous une approche de méthodes mixtes pour mieux comprendre le niveau de fidélité atteint*. Bien que des approches quantitatives soient couramment utilisés dans les essais randomisés en grappes pour évaluer l'efficacité d'une intervention, des approches qualitatives sont fortement recommandées pour évaluer des interventions complexes pour contribuer à mieux comprendre les résultats de l'intervention (Medical Research & Council., 2008). En effet, les méthodes qualitatives à l'intérieur d'un essai contrôlé peuvent explorer des facteurs qui affectent la fidélité de la mise en œuvre de l'intervention (Scantlebury et al., 2020), et peuvent éclairer la prise de décision sur la poursuite ou l'arrêt d'un essai ou de ses composantes (Murtagh et al., 2007). Bien qu'évaluer les interventions sous une approche de méthodes mixtes implique du temps pour la planification et la mobilisation des ressources, cette évaluation se voit essentielle pour mieux comprendre la mise en œuvre de l'intervention et mieux répondre à la question de la recherche concernant entre autres la fidélité de la mise en œuvre de l'intervention (Creswell & Clark, 2017).

4.2.1 Outil pour évaluer la fidélité de la mise en œuvre de l'intervention dans les essais randomisés en grappes

Une évaluation de la mise en œuvre, qui s'intéresse particulièrement à la variable « fidélité de la mise en œuvre », c'est-à-dire au degré auquel l'intervention a été mise en œuvre comme elle a été prévue, mais qui explore tous les facteurs qui l'influencent, s'impose dans les interventions de santé publique évaluées par des essais randomisés en grappes. Cette évaluation est encore plus importante quand les interventions sont mises en place dans les pays du Sud, qui ont une dynamique encore plus complexe pour faire face aux différentes menaces auxquelles ces pays sont confrontés.

Afin de d'amener une réflexion plus systématique aux chercheurs impliqués dans la planification ou l'évaluation de la fidélité de la mise en œuvre des interventions en santé publique, et pour combler un écart dans la littérature concernant la disponibilité d'un outil pratique à utiliser pour les chercheurs qui veulent évaluer la fidélité de la mise en œuvre de l'intervention dans ce type d'intervention, *cette thèse contribue à l'avancement des connaissances du domaine de*

l'évaluation des interventions, en développant et testant dans l'étude de cas un outil pour faciliter cette pratique évaluative. En effet, à partir des exemples tirés de la revue systématique, d'une exhaustive révision de la littérature et en nous appuyant sur le cadre de références de Carroll et collègues, modifié par Hasson, nous avons suggéré des questions pour mener à bien l'évaluation de la fidélité de la mise en œuvre de l'intervention qui prend en considération chaque dimension à évaluer, et avec des exemples qui peuvent orienter le chercheur pour adopter cette pratique évaluative (Tableau 3). Étant donné la nature de l'essai randomisé en grappes, évaluer si le groupe contrôle reçoit partiellement ou complètement l'intervention est nécessaire, afin d'explorer si des contaminations se produisent entre les groupes. Donc, au cadre de référence utilisé, nous ajoutons la composante « groupe contrôle » afin d'en tenir compte dans l'évaluation. Puisque les dimensions sont interdépendantes, elles pourront être évaluées simultanément pour mieux établir la relation entre tous ces facteurs et la fidélité atteinte. Cela révèle en même temps la difficulté de cibler chaque dimension de façon isolée, et de l'importance de faire une évaluation plus articulée qui reflète la réalité de la mise en œuvre. Cela exige une excellente compréhension de l'intervention et des outils de collecte de données adaptés au programme. Donc, la présentation de cet outil est conçue à partir des questions pour aider les chercheurs à réfléchir systématiquement sur la façon de concevoir, et de rapporter l'évaluation de la fidélité de la mise en œuvre de l'intervention en adaptant les questions à leurs propres interventions. Des exemples tirés des études de la révision systématique sont donnés à titre illustratif pour orienter les chercheurs sur l'évaluation et le rapport de chaque dimension dans les interventions qui ont été mises en place dans un contexte déterminé, et qui reflètent la réalité de la fidélité de la mise en œuvre des interventions en santé publique dans le cadre des essais randomisés en grappes. Nous avons testé cet outil dans notre étude de cas afin de valider la compréhension et évaluer son utilité dans l'évaluation des interventions. Nous avons constaté que l'outil est facile à adapter à l'intervention à évaluer, et facilite en même temps l'élaboration des instruments pour la collecte des données. Cet outil repose sur la prise en compte des réalités des interventions complexes qui interagissent dans leur contexte de mise en œuvre sous le cadre des essais randomisés en grappes.

Cependant, le pouvoir d'agir pour favoriser l'appropriation de ce processus d'évaluation devient essentiel, afin d'améliorer cette pratique évaluative par des acteurs concernés (Fetterman, 2005). Donc, pour assurer la compréhension du contenu, pour raffiner et pour valider cet outil dans

le but d'accroître sa validité et son utilisation par d'autres chercheurs, nous avons contacté des collègues qui travaillent dans l'Organisation Mondiale de la Santé (OMS), qui sont experts dans la mise en place et dans l'évaluation des interventions en santé publique. Les collègues de l'OMS ont trouvé cette idée très importante pour s'impliquer dans un groupe de travail pour améliorer l'outil, et pour explorer des options pour le valider à une échelle plus importante. Motivés pour améliorer cette pratique et son rapport dans l'évaluation des interventions de santé publique dans le cadre des essais randomisés en grappes, nous avons contacté des chercheurs du groupe CONSORT pour explorer les options de modifier les lignes directrices, afin que la fidélité de la mise en œuvre de l'intervention soit incluse comme un élément clé à tenir compte pour l'évaluation et le rapport dans ce type d'étude. Le groupe CONSORT a accepté notre argument sur l'importance d'évaluer la fidélité de la mise en œuvre de l'intervention, et d'en tenir compte dans la prochaine révision des guides directeurs dans les prochains mois.

Cet outil favorise la contribution à différents niveaux : (i) aux sciences de l'implémentation pour l'avancement des connaissances concernant la disponibilité d'un outil pratique et fiable à utiliser par les chercheurs qui visent ce type d'intervention; (ii) aux chercheurs qui mènent ce type d'intervention, car cet outil les aidera à mieux concevoir, évaluer et rapporter la fidélité de la mise en œuvre de l'intervention, afin de savoir si les effets mesurés peuvent être attribués aux interventions ou à des facteurs exogènes; (iii) à la conception de l'étude, pour renforcer la validité interne et externe de l'intervention, deux aspects critiques pour répliquer et généraliser toute intervention, ainsi qu'à favoriser la validité théorique et la puissance statistique sur les résultats de l'intervention. En effet, les chercheurs ne devront pas se limiter à l'évaluation d'efficacité, mais devront également explorer les interactions qui se produisent entre l'intervention et le contexte pour renforcer la validité de ses résultats; (iv) pour ceux qui financent des interventions de santé et pour les décideurs qui veulent mettre en place des interventions mieux adaptées au contexte pour améliorer la santé de la population afin d'atteindre les Objectifs de Développement Durable d'ici 2030; et enfin (v) éventuellement, nos résultats pourraient contribuer à améliorer le guide CONSORT concernant l'inclusion de l'évaluation de la fidélité de la mise en œuvre de l'intervention comme un critère essentiel à considérer systématiquement par les auteurs dans le rapport des résultats de ce type d'étude.

Tableau 3. – Outil pour évaluer la fidélité de la mise en œuvre des interventions en santé publique

Adhésion	
Le contenu	<p>Exemples de questions d'évaluation</p> <p>Dans quelle mesure (nom de l'activité) a-t-il été mis en œuvre comme prévu?</p> <p>Dans quelle mesure les principaux éléments d'intervention ont-ils été livrés comme prévu?</p> <p>Quelle proportion des composantes du programme a été livrée par rapport au nombre prescrit?</p> <p>Tous les éléments d'intervention ont-ils été mis en œuvre conformément au protocole d'essai?</p>
	<p>Exemples pratiques des interventions évaluées</p> <p>- To take of blood samples and use rapid diagnostic tests (previous certified training was required):</p> <p><i>“To ensure that chemical sellers correctly conducted the rapid diagnostic tests and adhered to safety and infection control procedures, direct observation of interactions between chemical sellers and clients was carried out by the study team using a checklist on a weekly basis during the first month of evaluation and for a further week mid-way through the trial in both arms” (Ansah et al, 2015).</i></p> <p>-To promote a free distribution of sodium dichloroisocyanurate. Interpersonal communicators (IPCs) visited each household fortnightly and gave them free of charge a box containing 30 tablets along with instructions on use. During each visit, IPCs providing information on the adverse health effects associated with consuming contaminated drinking water and the importance of treating drinking water.</p> <p><i>“Adherence to the intervention was assessed by asking house-holders if they had treated their children’s drinking water and by testing it for RFC [residual free chlorine]. If the caregiver replied that the child’s drinking water was treated, she was asked which method was used. This open-ended question was designed to monitor not only reported use of tablets, but also of other water treatment methods” (Boisson et al, 2013).</i></p> <p>-The intervention consisted of latrine promotion and construction, in accordance with the Government of India’s Total Sanitation Campaign, which combines social mobilisation with a post-hoc subsidy.</p> <p><i>“We measured compliance with the intervention with a survey done at the midpoint of the follow-up period. The survey recorded latrine presence and functionality, reported latrine use, and global positioning system (GPS) location of latrines and households” (Clasen et al, 2014).</i></p>

	<p>-To evaluate the efficacy of the program Keep Moving toward Healthy Heart and Healthy Brain (KM2H2) in encouraging physical activities for the prevention of heart attack and stroke among hypertensive patients The intervention consisted of six sessions and two booster sessions engineered in a progressive manner. The purpose is to motivate and maintain physical activities for the prevention of heart attack and stroke.</p> <p><i>"To ensure intervention fidelity, (1) all telephone counseling must be taperecorded; (2) one fifth of the recordings are randomly selected for assessment; and (3) the Consultation Manual was strictly followed as guidance to the telephone counseling at all times" (Gong, J., et al, 2015)</i></p>
	<p>Exemples de questions d'évaluation</p> <p>Est-ce que tous ceux qui auraient dû participer à l'intervention l'ont fait?</p> <p>Quelle proportion du public cible visé a participé à l'intervention?</p> <p>Quelle proportion de la population cible a participé? Quelle proportion a abandonné ou a été perdue de vue?</p> <p>Quelles sont les caractéristiques de ceux qui ont participé par rapport à ceux qui n'y ont pas participé?</p> <p>Quelles sont les caractéristiques de ceux qui ont reçu tous les éléments d'intervention comme prévu par rapport à ceux qui ont reçu moins d'éléments, ou aucun?</p>
<p>La couverture</p>	<p>Exemples pratiques des interventions évaluées</p> <p>- Proportion of households that have a functional latrine at the end of the implementation period:</p> <p><i>"66% of households in intervention villages were observed to have a completed latrine or one under construction. Coverage was higher among BPL [below poverty line] than APL [above the poverty line] households (68% versus 61%, respectively) and ranged from 54% to 76% across the different NGOs. While 39 (78%) intervention villages achieved over 60% coverage (...), 8 villages (16%) lagged behind with coverage of latrine completed and under construction remaining at less than 40% (Figure 2)."The process documentation team could monitor outcomes such as latrine coverage independently from the implementers, but had to work closely with village motivators and NGOs to collect information on mobilisation activities" (Clasen et al, 2014).</i></p> <p>-Proportion of preterm babies whose mothers received the antenatal corticosteroids in a defined place and period:</p> <p><i>"Of the women identif ed as being at risk of preterm birth, 6109 (98%) of 6214 received antenatal corticosteroids, mainly by ACT kit, with most women receiving a full course (table 3). 14 women received more than one course of antenatal corticosteroids. Of all women who received antenatal corticosteroids in the intervention group, 976 (16%)</i></p>

	<p><i>of 6109 had delivered a less-than-5th-percentile infant”.</i> (Althabe et al, 2015)</p> <p>-Proportion of patients diagnosed with pulmonary drug-sensitive TB that attending intervention clinics and receiving the voucher:</p> <p><i>“Of all 2076 patients who were eligible to receive a voucher for the six to eight months of their treatment, 813 (36.2%) did not receive a voucher at all, and 671 (32.3%) received a voucher for between one and three months”</i> (Lutge et al, 2014).</p>
<p>La fréquence et la durée</p>	<p>Exemples de questions d'évaluation</p> <p>L'intervention a-t-elle été réalisée avec la fréquence et la durée prescrites dans le protocole d'essai?</p> <p>Pour chaque activité programmée, la durée estimée a-t-elle été respectée?</p> <p>Quelle était la fréquence et la durée prévues pour chaque session? Cela a-t-il été respecté?</p> <p>Dans quelle mesure la durée et la fréquence des composantes de l'intervention ont-elles été livrées comme prévu?</p> <p>Exemples pratiques des interventions évaluées</p> <p>-Six tobacco health education events (HEEs), approximately 4–6 weeks apart in between baseline and follow-up assessments, conducted by a trained health educator:</p> <p><i>“In addition, we tracked implementation of the intervention protocol. For the management intervention, senior study staff responsible for this intervention completed process tracking forms for each meeting, including meeting attendees, total meeting time, and topics addressed. Health educators tracked the worker-level intervention; for each health education event, process tracking included the total number of participating workers, average time spent on each session with workers, whether or not specific topics or activities were implemented as per the protocol, and the number of materials displayed at each event. These data were used to track fidelity to the intervention protocol and dose of intervention delivered</i> (Sorensen et al, 2016)</p> <p>-The intervention was a cycle of women’s group meetings led by ASHAs [Accredited Social Health Activists]. It followed rules of participatory learning and action and had a four-phase structure as recommended by World Health Organization</p> <p><i>“Meetings were held fortnightly for the first four months (phases 1 and 2), and every month thereafter, following a sequence described in the appendix. Although the intervention was originally planned for 24 months, it lasted for 31 months (Sept 1, 2010, to March 30, 2013), because the groups chose to hold additional meetings in phase 3”</i>(Tripathy et al, 2016)</p>

-The specifics of intervention activities are not rigidly proscribed but rather develop and evolve in direct response to community priorities, needs and characteristics. Each phase builds on the other, with an increasing number of individuals and groups involved, strengthening a critical mass committed to and able to create social norm change.

“Barriers to program activity during the follow-up period (due to political disturbances and the suspension of activities during political election campaigns) mean that this four-year follow-up equates to approximately 2.8 years of SASA! Programming” (Abramsky et al, 2014).

Facteurs modérateurs

Exemples de questions d'évaluation

La description de l'intervention était-elle claire? La description de l'intervention a-t-elle indiqué les procédures et le matériel requis pour chaque composante?

Y avait-il suffisamment de détails sur l'intervention, et comment cette intervention a-t-elle été administrée pour permettre la reproduction scientifique?

La théorie qui sous-tend l'intervention était-elle clairement décrite? L'intervention a-t-elle un modèle logique qui démontre les effets potentiels de l'intervention?

La théorie et les objectifs de l'intervention, les stratégies utilisées pour exécuter l'intervention, et les rôles et les responsabilités de ceux qui réalisent les interventions ont-ils été décrits?

La complexité de l'intervention

Exemples pratiques des interventions évaluées

-Details of the key components of community mobilisation along with the time frame for each activity are provided in Table 1

“In brief, the approach consisted of initial meetings with community leaders to explain the programme, a baseline assessment of the water, hygiene and sanitation and socio-economic profile of the village, the formation of Village Water and Sanitation Committee (VWSC), and a combination of community-level events and door-to-door household visits to encourage construction and use of toilets. (Clasen et al, 2014)

-The Newhints intervention was developed in close collaboration with the District health management teams of the trial districts

	<p><i>“Newhints is an integrated intervention package (figure 1), based on extensive formative research and developed and implemented in close collaboration with the district health management teams of the trial districts. The core component was training the CBSVs [Community-based surveillance volunteers] in the 49 intervention zones to identify pregnant women in their community and to undertake two home visits during pregnancy and three visits after birth on days 1, 3, and 7. Other components are outlined in figure 1; more detail including the content of each visit can be found in the trial protocol” (Kirkwood et al, 2013).</i></p> <p>-The Behavior Change Communication strategy was designed with the essential nutrition actions framework that focuses on 7 practices as follows: 1) women's nutrition, 2) anemia prevention and control, 3) iodine intake, 4) prevention of vitamin A deficiency, 5) breastfeeding practices, 6) complementary feeding practices, and 7) nutritional care for sick and severely malnourished children</p> <p><i>“The program theory framework and associated program impact pathways were used to design the instruments used in this study. The first step in using the program impact pathways to design the research was to map the inputs, processes, outputs, and outcomes that were expected to lead to impact. Once these were identified, the actors involved with each component, the processes that needed to be in place, and the underlying assumptions that must be met in order to optimize impact were considered, and sets of questions, compiled in the form of semistructured interviews, focus group discussions, and observations, were developed to address each of these (table 1)” (Olney et al, 2015).</i></p>
<p>Les stratégies qui facilitent la mise en œuvre de l'intervention</p>	<p>Exemples de questions d'évaluation</p> <p>Quelles stratégies ont été utilisées pour soutenir la mise en place de l'intervention?</p> <p>Quels facteurs/éléments ont facilité le développement des activités?</p> <p>Quelles stratégies ont été utilisées pour optimiser et standardiser la fidélité de la mise en œuvre de l'intervention?</p> <p>Exemples pratiques des interventions évaluées</p> <p><i>“The most important and modifiable factors to reach each intervention objective were selected from the <u>local evidence</u> collected during the needs assessment, input received from the CPPE [Comprehensive Participatory Planning and Evaluation] and existing <u>literature</u>” (Andrade et al, 2014).</i></p> <p><i>“On the fifth day of <u>training</u>, all trainees were required to take part in the qualification exam. Additional one-day training for telephone counseling was provided for trainees who were assigned to provide the counseling intervention. Trainees who participate in both trainings and passed the two exams were certified as interventionists to deliver the KM2H² [Keep Moving</i></p>

	<p>toward Healthy Heart & Healthy Brain]. <i>For two CHS [community health centers] physicians who did not pass, additional training was provided until they passed the exams</i>” (Gong et al, 2015)</p> <p><i>“To ensure fidelity, in-class assistance involving <u>modelling, coaching, and support and feedback</u> in the implementation of the strategies was provided to each teacher once a month for 4 months for 1 h. A small amount of <u>educational materials</u> was provided to facilitate use of the strategies”</i> (Baker-Henningham et al, 2012)</p> <p><i>“In designing these learning aids, we followed the process for <u>developing materials</u> in Tanzania [38] where “job aids” were developed that are culturally tailored, easy to use, and feasible to distribute. All materials were used in the pilot interventions and we debriefed pregnant WLH [Women Living with HIV], nurses in health clinics, and supervising Peer Mentors in development of the materials to ensure cultural appropriateness”</i> (Rotheram-Borus et al 2014)</p> <p><i>“There was generally <u>good commitment</u> to STRETCH [Streamlining Tasks and Roles to Expand Treatment and Care for HIV] among <u>management, trainers, clinic staff, and patients</u>. Nurses were comfortable with and enthusiastic about the opportunity to be involved more directly in providing life-saving treatment.”</i> (Fairall et al, 2012).</p>
<p>La qualité de la mise en œuvre de l’intervention</p>	<p>Exemples de questions d'évaluation</p> <p>Comment la qualité de la prestation des interventions a-t-elle été contrôlée?</p> <p>Comment la qualité a-t-elle été assurée pour chaque activité mise en œuvre?</p> <hr/> <p>Exemples pratiques des interventions évaluées</p> <p><i>“The study team carried out regular quality control procedures by picking a random sample of test kits from the shop for testing using standard positive blood samples. This was to ensure that the rapid diagnostic test kits continued to perform optimally under the storage conditions in the chemical shop”</i> (Ansah et al, 2015).</p> <p><i>“Implementation of the mass-treatment days was encouraged and monitored independently by project staff and state government staff. ...Monitors knew supervisors would, with 5% probability, randomly restudy selected villages shortly afterwards to ensure the relevant households had all been visited (and to help consolidate monitors’ discipline and methods)”</i> (Awasth et al, 2013).</p> <p><i>“To promote sustainability, planning, and management, the Mchinji district health office was involved. Counsellors were supervised by 24 government health surveillance assistants, who visited quarterly, documenting progress, achievements, and</i></p>

	<p><i>challenges, and by three MaiMwana officers. Health surveillance assistants received initial and yearly refresher training, attended monthly and quarterly meetings, and were given calendars and a supervisory checklist” (Lewycka et al, 2013).</i></p>
La réponse des participants	<p>Exemples de questions d'évaluation</p> <p>Comment les participants perçoivent-ils la pertinence et l'utilité de l'intervention?</p> <p>Quelle était l'attitude des participants par rapport aux activités proposées?</p> <p>Dans quelle mesure les membres du groupe ont-ils participé activement aux activités proposées?</p> <p>Quelles sont les raisons pour lesquelles le public cible participe ou non aux activités proposées?</p>
	<p>Exemples pratiques des interventions évaluées</p> <p><i>“Although the investigators are confident that interventionists followed process and fidelity protocols in delivering TA [Team Awareness], there may be factors that could have contributed to the inability of the study to generate certain hypothesised effects. For instance, level of participant engagement and participation in all of the modules were not measured using an evaluation checklist, instead participants were asked at the end to comment on their experience of the intervention. The use of session ratings is recommended in future replications of TA.” (Burnhams et al, 2015).</i></p> <p><i>“In addition to general “friendly” conversation, peer leaders were instructed to communicate about HIV prevention and testing. As no established best practice existed for health and social media communication, peer leaders talked weekly with the peer leader trainer on how to increase participant engagement. For example, in the first week, peer leaders were instructed to send friendly messages to elicit a basic response from participants and create rapport with them...To monitor intervention content and fidelity, each week, peer leaders returned “response sheets” indicating whether and which participants responded to their contact attempts, coded by date, contact method, topic of content, and participant engagement” (Young et al, 2015).</i></p>
Le recrutement	<p>Exemples de questions d'évaluation</p> <p>Quelles procédures ont été utilisées pour approcher et attirer les participants cibles?</p> <p>Les stratégies de recrutement des participants ont-elles été mises en œuvre comme prévu?</p> <p>Le processus de recrutement a-t-il été appliqué de manière cohérente pour toutes les grappes et pour tous les participants (par exemple personnes / villages / pays)?</p>

	<p>Quels groupes de personnes étaient plus/moins susceptibles d'être recrutés avec succès?</p> <p>Exemples pratiques des interventions évaluées</p> <p><i>“Recruitment and training were considered successful and facilitators’ skills in leading women’s group meetings generally improved over time... Local recruitment was vital in maintaining effective relationships with the local community through familiarity with local customs, knowledge and beliefs” (Fottrell et al, 2013).</i></p> <p><i>“Village motivators were recruited from the project area to facilitate mobilisation activities and coordinate latrine construction logistics in villages” (Clasen et al, 2014).</i></p> <p><i>“The counsellors (four women and four men) were selected among already trained local lay counsellors. These lay health workers were individuals selected by the local communities for voluntary work at the clinics. The main reasons for involving lay counsellors were the low costs and the limited number of local health care workers” (Fylkesnes et al, 2013).</i></p> <p><i>“Patients who had previously had dengue and their families were not involved in setting the research question or the outcome measures (...) but were intimately involved in design and implementation of the intervention and were central in diffusing baseline information to motivate community involvement” (Andersson et al, 2015).</i></p>
<p>Le contexte</p>	<p>Exemples de questions d'évaluation</p> <ul style="list-style-type: none"> • Comment les facteurs contextuels (la restructuration du système de santé, les conditions de travail, le niveau de scolarité des participants, le niveau d'engagement entre les partenaires, etc.) ont-ils influencé les résultats de l'intervention? • Quels facteurs contextuels conduisent à des adaptations bénéfiques de l'intervention? • Quelles ont été les raisons de modifier l'intervention prévue? <p>Exemples pratiques des interventions évaluées</p> <p><i>“Focus group discussions attributed this compliance to political support, sustained and predictable drug supply, minimum additional work for AWC [anganwadi child-care centre] staff (who did not themselves have to report on compliance), helpful staff training meetings, and the knowledge that coverage might well be monitored” (Awasthi et al, 2013).</i></p> <p><i>“Although the intervention might have played some role in the increase, it is likely that other contextual factors were also important: just 1 y after implementation started, 65% of the comparison area births were in a health facility [26]. In part the</i></p>

	<p><i>increase could be due to improved communication and transport: less than 10% of households had a mobile phone in 2007 compared to 48% in 2013. The availability of motorcycle transport also increased substantially, and the focused national antenatal care programme promoting facility delivery might have contributed to the shift. Lastly, facility based quality improvement programmes emphasising facility delivery, including our own in 24 facilities, have been ongoing in the study area, as well as focal home counselling work by different development agencies” (Hanson et al, 2015).</i></p> <p><i>“Our study was conducted in the presence of other ongoing interventions that may have made it difficult to attribute any specific, small changes in nPCR parasite prevalence to our intervention. The absence of detailed concurrent meteorological data makes it impossible to assess the impact of our interventions in terms of the exact timing of the local changes in rainfall and temperature that determine malaria transmission dynamics” (Bousema et al, 2016).</i></p> <p><i>“The results reported below present the key findings emerging from the qualitative evaluation of this trial. The first three sections report on the general acceptability and fidelity of the implementation process. The last four sections highlight four key factors— pharmacy, human resources, clinical support, and local management input— that affected the implementation of STRETCH” [Streamlining Tasks and Roles to Expand Treatment and Care for HIV] (Fairall et al, 2012).</i></p>
<p>Le groupe contrôle</p>	<p>Exemples de questions d'évaluation</p> <p>Les événements se déroulant dans le groupe témoin ont-ils été surveillés?</p> <p>Quelles stratégies ont été utilisées pour surveiller les événements dans le groupe témoin?</p> <p>Le groupe témoin aurait-il pu recevoir ou avoir été exposé à des éléments d'intervention au cours de l'étude qui pourraient avoir un impact sur les résultats?</p> <p>Quelles stratégies ont été adoptées pour éviter la contamination (réception de tout élément d'intervention par les membres du groupe témoin)?</p> <p>Exemples pratiques des interventions évaluées</p> <p><i>Owing to the requirements of the trial design, the media and advocacy activities were restricted to local media channels in order to try to avoid exposing control communities to SASA! [a community mobilization intervention to prevent violence and reduce HIV-risk behaviors] Ideas and materials... Very few respondents in control communities reported any exposure to SASA! materials, activities or multimedia events (2% of men and 1% of women), a reassuring indication that efforts to reduce diffusion</i></p>

of the intervention to control communities were successful.” (Abramsky et al, 2014).

“Mothers from both control and intervention groups were asked if they recalled receiving any messages about child care.” (Singla et al, 2015).

“Nearly 40% of children in control villages also had used Sprinkles MNP [micronutrient powders] because of vendors who sold Sprinkles MNP outside their own villages... Because some children in control villages used Sprinkles MNP during the study period, we conducted an as-treated analysis to assess the impact of the use of Sprinkles MNP on biomarkers of children in both groups combined (Figure 3) (Suchdev et al, 2012).

4.3 Contribution empirique à la pratique évaluative de la fidélité de la mise en œuvre de l'intervention

Plusieurs leçons peuvent être tirées concernant la pratique évaluative de la fidélité de la mise en œuvre de l'intervention (Tableau 4) et sont présentées dans les prochaines lignes. Ces leçons offrent certaines pistes d'action pour améliorer et pour promouvoir la fidélité de la mise en œuvre de l'intervention dans le cadre des essais randomisés en grappes, mais qui peuvent aussi être utilisées dans d'autres types d'études.

Tableau 4. – Points saillants de la pratique évaluative de la fidélité de la mise en œuvre de l'intervention

- Mieux comprendre la théorie de l'intervention (ou théorie du programme) facilite la compréhension du développement d'une intervention, et facilite aussi la création des instruments de collectes de données pour optimiser l'évaluation de la fidélité.

- L'évaluation de la fidélité est une méthode utile pour les essais pilotes, afin de renforcer les interventions, d'améliorer leur durabilité et de les reproduire à grande échelle dans d'autres contextes.

- Les facteurs modérateurs sont des éléments clés sous-étudiés, qui contribuent à une mise en œuvre réussie et à une haute-fidélité de la mise en œuvre.

- Pour aborder l'évaluation de la fidélité de la mise en œuvre de l'intervention, une approche multidisciplinaire, multidimensionnelle et avec des méthodes mixtes s'avère nécessaire pour mesurer, pour mieux comprendre et pour fournir une vision plus complète de la fidélité de la mise en œuvre de l'intervention.

- L'outil d'évaluation de la fidélité de la mise en œuvre de l'intervention, conçu et testé dans le cadre de cette recherche, présente des orientations futures pour la recherche pour améliorer cette pratique évaluative, et pour faire progresser la science de la mise en œuvre.

Concernant l'étude de cas, plusieurs enseignements sont tirés de cette évaluation. La compréhension théorique qui soutient l'intervention a permis de mieux comprendre le développement de l'intervention (Article 2), ainsi qu'apporter des éléments clés pour la construction du modèle logique de l'intervention évaluée. Cela a permis, à son tour, de créer des instruments adaptés pour mener à bien l'évaluation de la fidélité de la mise en œuvre de l'intervention, et de dégager les éléments favorables ou les obstacles pendant la mise en place de l'intervention, qui ont influencé le niveau de la fidélité atteint (Article 3). L'évaluation de l'intervention met en lumière l'importance de concevoir une intervention adaptée au contexte pour favoriser sa réussite, pour faciliter la participation de la population cible et pour atteindre ainsi un plus grand nombre des personnes. Également, s'assurer d'une excellente compréhension de la théorie de l'intervention par les intervenants favorise la mise en place de l'intervention comme elle a été planifiée (Pérez et al., 2020).

En ce qui concerne l'évaluation de la fidélité de la mise en œuvre de l'intervention Tika Vaani, on peut faire ressortir plusieurs pistes d'action pour la pratique évaluative :

Premièrement, l'outil développé pour évaluer la fidélité de la mise en œuvre de l'intervention a été adapté et testé lors de l'intervention. Cela a permis de dégager les différents éléments qui favorisent la mise en place de l'intervention, et donne une certaine assurance quant à l'utilité de cet outil pour développer des instruments de collecte de données adaptés à l'intervention. Cet outil, conçu et testé dans le cadre de cette recherche, constitue le premier pas pour cibler d'autres parties prenantes, afin de construire un guide qui explique en détail son utilisation, et qui permettra d'accroître sa validité et son application par un grand nombre de chercheurs (Guichard, Tardieu, Nour, Lafontaine, & Ridde, 2019).

Deuxièmement, les apprentissages de cette expérience évaluative révèlent des éléments qui vont favoriser la fidélité la mise en œuvre de l'intervention et de son évaluation. Des stratégies pratiques pour promouvoir la fidélité de la mise en œuvre de l'intervention sont présentées ci-dessous, selon le moment de leur utilisation.

Avant la mise en place de l'intervention, la conception de l'intervention est le premier élément central de la fidélité (Bellg et al., 2004; Bond et al., 2000; Dane & Schneider, 1998; Gearing et al., 2011). Les chercheurs doivent avoir une définition claire du contenu de l'intervention et doivent bien définir les composantes essentielles de l'intervention. Ils doivent s'assurer que les intervenants ont une compréhension claire de la théorie de l'intervention, ainsi que de la façon de procéder pour mettre en place l'intervention. Pourtant, la construction d'un modèle logique est essentielle pour clarifier et donner un aperçu visuel des composantes de l'intervention pour les intervenants et bailleurs de fonds ainsi que pour faciliter la mesure de la fidélité de la mise en œuvre de l'intervention. De même, les chercheurs ou les personnes responsables de l'intervention devront développer des ateliers de formation contextualisés aux intervenants et à leurs besoins, ainsi que créer des manuels des fonctions détaillés pour les intervenants et fournir du matériel adapté au public cible. Notamment, la formation des intervenants est un élément essentiel de la fidélité de la mise en œuvre de l'intervention, car le niveau d'instruction, le niveau des compétences à développer, les habilités, les compétences et l'éducation propres des intervenants ont des effets directs sur le déroulement de l'intervention prévue. Une phase de recherche formative devient un élément clé et est essentielle pour développer des interventions adaptées au contexte, pour mieux constituer l'équipe de recherche sur le terrain, mais aussi pour favoriser la compréhension de l'intervention et la mise en place telle que planifiée.

Parallèlement, il est fondamental d'avoir un manuel d'intervention qui décrive les objectifs et les stratégies, y compris la façon dont un programme doit être livré en accord à la théorie du programme. Le rôle, la responsabilité des interventionnistes, l'équipement et le matériel nécessaires, ainsi qu'un protocole de dépannage pour répondre aux problèmes pouvant survenir lors de la mise en œuvre de l'intervention, devront être précisés par les chercheurs dans le manuel, pour mieux guider les intervenants dans la mise en place des activités. Par ailleurs, tester l'intervention et faire une rétroaction entre les participants et les intervenants pour affiner l'adhésion au modèle théorique et pour améliorer l'acceptabilité, la faisabilité et l'efficacité potentielle de l'intervention, est fortement suggéré (B. Borrelli, 2011). Finalement, des lignes directrices définissant la mesure acceptable de fidélité devraient être établies avant la mise en œuvre de l'intervention. Malgré qu'il n'existe pas une mesure standard, il a été suggéré qu'un

niveau de fidélité élevé va au-delà de 80 à 100%, tandis qu'une mesure faible de fidélité est à 50% ou moins (Perepletchikova & Kazdin, 2005).

Pendant la mise en œuvre de l'intervention, des rencontres de rétroaction périodiques avec les intervenants, et la supervision sur le terrain pour appliquer des mesures correctives, afin d'éviter les violations répétées du protocole, sont fortement suggérées. En effet, le suivi est considéré comme le cœur de la fidélité, car il permet de retracer la mesure de la fidélité au cours de l'intervention (Perepletchikova & Kazdin, 2005). Des suivis peuvent se faire pour évaluer, non seulement si les activités de l'intervention sont mises en place, mais aussi pour évaluer les compétences des intervenants et leur capacité d'interagir avec les participants. Une intervention aura probablement plus de succès, si les intervenants reçoivent régulièrement des commentaires, du soutien et des encouragements lors de la mise en œuvre de l'intervention. Ainsi, la supervision des intervenants est une stratégie de base pour faciliter l'adhésion au traitement, et réduire l'écart dérivé par rapport au modèle initial. Également, la fidélité au programme est fortement influencée par l'engagement affiché par le coordinateur du site, qui défend le programme, qui s'assure que les protocoles du programme sont en place, qui permet d'identifier les problèmes et d'aider à résoudre les problèmes de la mise en œuvre (S. F. Mihalic, Fagan, & Argamaso, 2008).

Donc, afin de garantir le respect de la fidélité de la mise en œuvre de l'intervention, il est essentiel d'assurer à la fois une formation avant le départ et une supervision pendant la mise en place de l'intervention. Finalement, le suivi doit se faire autant pour la livraison que pour la réception de l'intervention. En effet, le suivi de la réception de l'intervention est aussi un élément fondamental pour savoir si les participants reçoivent l'intervention pendant la période visée de l'intervention, et permet de connaître le niveau auquel les participants se sont engagés et ont adhéré au contenu des activités suggérés. Des tests de connaissances avant et après l'intervention sont une façon de mesurer la réception de l'intervention. Ces mesures sont essentielles lorsque les participants peuvent par exemple avoir un faible niveau d'éducation (Bellg et al., 2004).

Enfin, *une fois l'intervention terminée*, consulter une personne indépendante du groupe de recherche pour faire l'évaluation rétrospective de la fidélité de la mise en œuvre de l'intervention est pertinent. Cependant, une évaluation de la fidélité de la mise en œuvre de l'intervention pourrait

se faire dans le continuum pour éviter des écarts pendant la phase d'implémentation. Pour faire cette évaluation, une liste de contrôle adapté aux activités de l'intervention pour comparer ce qui a été planifié vs ce qui a été livré aux participants, l'observation directe en temps réel ou l'observation des bandes de vidéos ou audio des activités mises en place, et la consultation d'accomplissement des protocoles sont couramment utilisées. Des notes de terrain des intervenants pour mieux comprendre les enjeux de la mise en œuvre sont aussi un source documentaire importante à considérer (Bellg et al., 2004; Carroll et al., 2007; Gearing et al., 2011; Graham et al., 2016; Hasson et al., 2012; Moncher & Prinz, 1991; Schoenwald et al., 2000; B. Wang et al., 2015).

Donc, l'évaluation de la fidélité de la mise en œuvre de l'intervention demande l'engagement et la collaboration du personnel du programme et d'autres parties prenantes, en plus de l'expertise technique et du leadership de l'évaluateur pour que l'évaluation soit réussie (Orwin, 2000). La description détaillée du programme, l'entretien avec le personnel du programme, l'observation des activités du programme et l'élaboration d'une évaluation modèle de programme devront faire partie des stratégies à considérer au moment de planifier une telle évaluation. Pour Moncher et Prinz, s'appuyer sur des sources externes, sur des rapports des participants et des rapports des intervenants sont des étapes clés à suivre pour vérifier la fidélité de la mise en œuvre de l'intervention (Moncher & Prinz, 1991).

En somme, utiliser une approche multidimensionnelle pour fournir une évaluation plus équilibrée de la fidélité de la mise en œuvre de l'intervention en tenant compte autant des *dimensions structurelles que des dimensions du processus* de la fidélité de la mise en œuvre de l'intervention est nécessaire voire indispensable (Dane & Schneider, 1998; Mowbray, Holter, Teague, & Bybee, 2003). Pour Harne et collègues, *les dimensions structurelles* de la fidélité de la mise en œuvre de l'intervention font référence aux éléments importants ou aux ingrédients actifs de l'intervention. Ces dimensions devront faire l'objet d'une évaluation, soit par observation directe ou soit par une auto-évaluation de l'intervenant, en indiquant si chaque composante s'est produite comme prévu, s'il y a des incohérences ou si l'évènement n'est pas observé (Harn, Parisi, & Stoolmiller, 2013). Construire une grille en tenant compte de toutes les activités de l'intervention, et par la suite, évaluer la fidélité de la mise en œuvre de l'intervention en examinant pour chaque activité planifiée, si celles-ci ont été mises en œuvre telles que prévues (I), telles que

modifiées (M) ou non mises en œuvre (N) est un outil pratique pour enregistrer ces informations (Rebhook et al., 2006). En ce qui concerne *les dimensions du processus*, les auteurs suggèrent d'examiner la qualité de la mise en place des différentes activités de l'intervention, et d'examiner les interactions entre les intervenants et les participants pour mieux comprendre la phase d'implémentation.

Bien que les données sur les *mesures structurelles* soient plus faciles et plus fiables à collecter, *les dimensions du processus*, bien que complexes et plus difficiles à mesurer de manière objective, peuvent fournir des informations essentielles et supplémentaires sur les raisons pour lesquelles une intervention est ou n'est pas efficace (Mowbray et al., 2003). Donc, l'évaluation de la fidélité de la mise en œuvre de l'intervention uniquement basée sur des données quantitatives peut avoir du mal à saisir ces nuances. Les analyses qualitatives et contextuelles offrent une méthode prometteuse pour déterminer dans quelle mesure l'intervention a été réalisée et reçue dans la pratique, et pour obtenir un aperçu réel et complet de la mise en place de l'intervention (Lawton, Mceachan, Jackson, West, & Conner, 2014). Bref, il n'y a pas une mesure unique ou supérieure à d'autres, car ces mesures devront être adaptées en fonction de l'intervention à évaluer. De plus, la triangulation des toutes les données fournies par l'évaluation de la fidélité de la mise en œuvre de l'intervention est essentielle, voire fondamentale, pour renforcer la validité des résultats issus de l'évaluation.

4.4 Recommandations pour de futures recherches

Les pays du Sud font toujours face aux défis en matière de santé publique qui exigent, entre autres, le développement et la mise en place des interventions pour atteindre les Objectifs de Développement Durable. Les interventions de santé publique mises en place dans les pays du Sud sont façonnées par de multiples facteurs qui peuvent influencer sa mise en œuvre. Les partenariats publics-privés actuels en santé publique, dans les pays du Sud, favorisent parfois le détournement des ressources économiques pour des programmes de santé publique, dû à la puissance des acteurs privés (Jacob, 2009). Cette différence de pouvoir entre les parties prenantes et les acteurs pose des défis plus importants pour agir en fonction des réels besoins de la population. Il est crucial que la relation de la santé publique avec les autres sphères soit inclusive et collaborative, en tenant compte

des meilleures preuves disponibles pour améliorer la santé de la population (Glasgow & Emmons, 2007). Favoriser des rencontres entre les différents acteurs clés afin de capitaliser sur l'intelligence interdisciplinaire et plurisectorielle dans le développement, l'implantation et l'évaluation des interventions peut y contribuer.

Malgré qu'un grand nombre de ce qui constitue aujourd'hui des problèmes de santé publique exige des interventions adaptées aux réalités locales, avec la mondialisation, la technologie, les mouvements de biens et de services et la croissance économique qui façonnent la santé (Labonté & Schrecker, 2007), ces interventions peuvent être guidées avec une perspective mondiale, ce qui permet à ces interventions de s'avérer efficaces dans d'autres contextes comparables pour prévenir les maladies et promouvoir la santé pour tous dans le monde. Des actions qui mènent les pays à agir localement, mais à penser à l'échelle globale, devront aller au-delà de la santé publique, si nous voulons nous attaquer aux déterminants transnationaux qui ont des effets sur « les déterminants locaux de la santé », afin de réduire des inégalités persistantes et accomplir l'engagement mondial de « ne laisser personne de côté » (Turcotte-Tremblay, Fregonese, Kadio, Alam, & Merry, 2020). Toutefois, concevoir une intervention réussie nécessite de savoir ce qui fonctionne bien dans un environnement particulier, et de savoir quels sont les problèmes les plus susceptibles de se poser, avant de mettre en place une intervention dans d'autres contextes. Cela implique autant des interventions adaptées au contexte que des évaluations qui puissent fournir des informations légitimes et accroître l'utilisation de leurs résultats par les décideurs.

L'utilité et l'importance des essais randomisés en grappes, comme un outil clé pour fournir des évidences scientifiques, seront renforcées dans la mesure où les études se voient bonifiées par une meilleure compréhension de la mise en œuvre à partir de l'évaluation de leur fidélité, pour atténuer les enjeux auxquels ces interventions sont confrontées. L'efficacité d'une intervention ne peut être isolée du contexte environnemental dans lequel elle intervient. Ainsi, une évaluation qui est transparente, qui tient compte de la mise en œuvre et qui vise les besoins des utilisateurs, afin de fournir les meilleures preuves basées sur l'évidence, s'avère nécessaire et complémentaire pour évaluer l'efficacité d'une intervention.

Pour entamer une telle démarche évaluative, il faut s'assurer d'abord de bien comprendre la théorie de l'intervention en identifiant et en comprenant les éléments essentiels d'un programme, avant de porter un jugement (Brousselle et al., 2011). Cela est encore plus important dans de nouvelles interventions qui se développent, car cette démarche permet de découvrir les éléments qui font une différence dans les résultats. Identifier les éléments de base susceptibles d'être responsables de l'efficacité pourrait laisser plus de choix à certaines caractéristiques, qui peuvent être modifiées sans altérer l'efficacité de l'intervention, et en conservant sa fidélité de sa mise en œuvre. En effet, à partir de la compréhension claire de la théorie de l'intervention et de ses composantes, le modèle logique peut se construire et des outils adaptés peuvent se développer pour favoriser une évaluation réussie, objective et utile pour les chercheurs. Envisager une évaluation de la fidélité de la mise en œuvre de l'intervention devrait tenir compte de cette étape avant d'entreprendre une telle démarche. De même, toute planification de l'évaluation devra tenir compte de l'utilisation de données mixtes, qui facilitent à la fois la mesure et la compréhension du niveau de fidélité atteint.

Le cadre conceptuel utilisé offre un outil potentiel pour évaluer la fidélité de la mise en œuvre d'une intervention, mais ne fournit aucune indication sur la manière d'explorer les relations entre ses éléments clés. L'outil développé, adapté et testé dans l'intervention Tika Vaani, a permis d'élaborer un panorama complet pour évaluer la fidélité de la mise en œuvre de l'intervention, et d'identifier des recommandations avant son passage à grande échelle. Cependant, des recherches futures pouvant affiner, tester et valider cet outil s'avèrent nécessaires pour que ce dernier soit scientifiquement valable pour la science de la recherche d'implémentation, afin que cet outil permette aux chercheurs de bien comprendre si les effets mesurés (ou le manque d'effets) peuvent être attribués aux activités de l'intervention (ou à des lacunes dans la mise en œuvre), et si ces effets mesurés facilitent la prise de décision concernant l'intervention évaluée.

4.5 Forces et limites de la thèse

Dans le champ de l'évaluation, cette thèse se distingue par le sujet de recherche ainsi que par les méthodes mobilisées pour son analyse. En effet, compte tenu de l'importance d'évaluer la

fidélité de la mise en œuvre des interventions en santé publique, les deux volets de cette thèse ont permis d'abord de connaître la réalité théorique et pratique de cette démarche dans des études qui sont citées, comme *gold standard*, pour établir l'effet causal d'une intervention et pour dégager des éléments clés pour son évaluation. Par la suite, cela a favorisé la construction d'un outil et l'application des stratégies à considérer, lors de l'évaluation de la fidélité de la mise en œuvre d'une intervention en santé publique, à travers l'étude de cas. Cette rétroaction fusionnée de résultats riches étend la portée des résultats, autant sur le plan théorique que pratique, concernant cette pratique évaluative et procure également une base solide pour continuer à explorer et à améliorer les lacunes et les défis auxquels se sont confrontés les chercheurs pour entreprendre une telle démarche.

Au niveau méthodologique, l'exhaustive révision documentaire, la rigueur méthodologique de la révision systématique ainsi que le contact avec des chercheurs qui ont évalué des interventions en santé publique sous la conception des essais randomisés en grappes, ont fourni la rigueur théorique pour construire et suggérer l'outil d'évaluation concernant la fidélité de la mise en œuvre de l'intervention. À partir de ce volet, on peut se prononcer sur plusieurs aspects : (i) il s'agit de la première revue systématique publiée portant sur des problèmes méthodologiques liés à l'évaluation de la fidélité de la mise en œuvre des interventions évaluées par des essais randomisés en grappes, ce qui complémente les résultats d'autres études qui ont montré que les informations requises pour assurer la validité interne et externe sont souvent mal rapportées dans ce type d'étude (S. Eldridge et al., 2008; Ivers et al., 2011); (ii) étant donné que plusieurs revues ont des contraintes d'espace, et que les auteurs de l'étude peuvent omettre des informations clés sur l'évaluation de la fidélité de la mise en œuvre de l'intervention, nous avons décidé d'inclure les essais randomisés en grappes, rapportant l'évaluation de la fidélité de la mise en œuvre de l'intervention, dans un document complémentaire, afin d'avoir une image complète du rapport actuel évaluant la fidélité. Notre revue était donc de grande taille, examinant 185 articles distribués dont 21 protocoles d'articles revus par des pairs, 15 protocoles de littérature grise, 65 registres de l'étude, 65 articles revus par des pairs, 6 rapports de littérature grise et 13 documents complémentaires; (iii) un autre point fort de cette revue est que, malgré la difficulté d'évaluer certains risques de biais, nous avons tenté de le faire systématiquement avec un outil d'évaluation détaillé, que nous avons développé sur la base de l'outil Cochrane Collaboration et d'autres sources clés pertinentes, pour évaluer le risque de biais

dans ce type d'étude, présenté en annexe de l'Article 1; (iv) enfin, et au-delà des résultats obtenus de la révision systématique qui dégagent les lacunes concernant la pratique évaluative, nous avons contacté les chercheurs pour mieux comprendre leur expérience pour développer, évaluer et rapporter la fidélité de la mise en œuvre de l'intervention, et nous avons aussi contacté les chercheurs qui n'ont pas considéré la fidélité de la mise en œuvre de l'intervention. Cela nous a confirmé l'importance du sujet de recherche, mais aussi le besoin d'aller de l'avant pour construire un outil qui permettra de guider les chercheurs dans cette démarche évaluative.

Sur le plan pratique, le fait d'être immergé théoriquement dans le domaine de cette pratique évaluative, ainsi que les apprentissages tirés de la revue systématique et de l'étude de cas a facilité : (i) l'adaptation de l'outil pour construire des outils appropriés pour évaluer la fidélité de la mise en œuvre de l'intervention évaluée; (ii) la possibilité de fournir des pistes d'action aux chercheurs pour permettre de planifier, d'élaborer, d'analyser et de rapporter l'évaluation de fidélité de la mise en œuvre d'une intervention en santé publique; (iii) étant donné que cette pratique évaluative exige une approche qui nécessite des données quantitatives et qualitatives qui peuvent limiter les chercheurs à entreprendre une telle démarche, car cela exige des ressources et du temps pour la collecte et l'analyse de données, nos résultats peuvent inspirer les chercheurs à entamer une telle démarche sur la façon de développer une évaluation de la fidélité de la mise en œuvre de l'intervention qui facilite, qui optimise la collecte et l'analyse de données, et qui rend son intégration plus facile à explorer, afin de bien donner un aperçu complet du niveau de la fidélité atteint. À notre connaissance, c'est la première étude qui explore et qui rapporte tous les facteurs qui influencent la fidélité de la mise en œuvre d'une intervention dans le cadre d'un essai randomisé en grappes mis en place dans un pays du Sud.

La richesse de cette évaluation à travers l'étude de cas prend toute sa valeur, grâce à la participation de l'étudiante-chercheuse dans la phase de développement de l'intervention, qui a permis de mieux comprendre la plausibilité de la théorie de l'intervention, le contexte dans lequel l'intervention est mise en place, les raisons qui ont motivé les chercheurs à implémenter une telle intervention, la façon de développer l'intervention et les mécanismes par lesquels l'intervention opère, ainsi que les raisons pour l'évaluer avant de la répliquer à grande échelle. L'utilisation de différentes sources d'information et la triangulation de toutes les données ont permis d'accroître la

validité de construit et de renforcer les résultats de l'évaluation de la fidélité de la mise en œuvre de l'intervention. En même temps, cela a favorisé la construction d'une évaluation objective et à la fois constructive pour élaborer des recommandations et de les examiner avec les concepteurs de l'intervention, afin de déterminer quelles sont les recommandations les plus adaptées au contexte pour les intégrer dans les futures étapes de l'intervention, renforçant ainsi la pertinence et la crédibilité de cette évaluation (Brousselle et al., 2011), ainsi que son utilisation (M. Patton, 2008).

Le fait de participer dans une équipe, où tous sont valorisés, a un rôle important à jouer. La possibilité en tout temps de partager nos connaissances et notre savoir-faire pour mieux comprendre le contexte et l'intervention, pour mieux adapter les instruments de collecte de données et pour mieux comprendre les résultats issus de la recherche, ont favorisé une participation « plus équitable et durable » nécessaire dans la recherche en santé mondiale (Valéry Ridde & Capelle, 2011). Enfin, le fait de participer, du début jusqu'à la fin, à l'évaluation de la fidélité de la mise en œuvre d'une intervention permet de donner un aperçu complet du milieu de la recherche, une capacité réfléchie et collaborative plus grande sur l'implication de développer et d'évaluer des interventions en santé publique, pour atténuer les défis de la recherche en partenariat Nord-Sud (Valéry Ridde & Capelle, 2011).

Bien que cette thèse soit construite avec une rigueur méthodologique qui procure la crédibilité et la validité aux résultats, certaines limites, quelques-unes déjà soulevées dans les résultats de chaque article, sont importantes à aborder, ainsi que les stratégies mise en place pour atténuer leur impact sur les résultats dans la recherche. Ces limites sont abordées ici en accord à chaque volet de recherche :

En ce qui concerne les résultats de notre révision systématique, certains limites peuvent se dégager : (i) pour garantir la disponibilité d'un protocole d'étude, nous avons inclus des études qui ont le mot «regist*» dans l'abstract, pour nous assurer que les études ont un protocole qui considère l'évaluation prévue de la fidélité la mise en œuvre de l'intervention, soit seule, soit en conjonction avec un protocole d'étude déjà publié. Par conséquent, nous n'avons peut-être pas identifié tous les essais randomisés en grappes. Cependant, nous avons examiné 16 études exclues pour évaluer ce risque, mais aucune ne répondait à nos critères d'inclusion. Ainsi, l'effet de cette limitation est

susceptible d'être faible.; (ii) nous n'avons pas pu calculer des résultats pour certaines questions de recherche définies dans notre protocole (M. C. Perez et al., 2016), qui visent les analyses *Intention To Treat* et l'analyse *par protocole* pour dégager l'efficacité de l'intervention pour deux raisons (a) il n'était pas toujours possible de déterminer si ces articles avaient effectué une analyse *Intention To Treat* pour évaluer l'effet de l'intervention, et (b) 58% (38/65) des articles ont rapporté plus d'un résultat principal et les conclusions de ces résultats concernant l'efficacité des interventions ont parfois divergé. Bien que cette absence des résultats ne permette pas d'explorer l'association des résultats de l'intervention liée à l'évaluation de la fidélité de la mise en œuvre de l'intervention, ce manque d'information n'empêche pas de faire découler et de documenter avec certitude l'absence et la divergence dans la planification et le rapport de la fidélité de la mise en œuvre de l'intervention dans ces études.

En ce qui concerne l'étude de cas, le fait d'étudier un seul cas pour évaluer la fidélité de la mise en œuvre de l'intervention accroît la validité interne des résultats à l'intérieur du cas, mais réduit le potentiel de généralisation des résultats (R. K. Yin, 2017). Cependant, le fait d'avoir pris en compte l'évaluation de la fidélité de la mise en œuvre de l'intervention d'autres études, qui ont le même modèle d'étude évaluée, et de contacter les chercheurs responsables de l'étude, permet de renforcer la compréhension de la démarche évaluative et la généralisation théorique de nos résultats. À propos du cas lui-même, quelques limites peuvent être soulevées : d'abord, c'est une intervention évaluée dans une étude pilote qui ne vise pas l'évaluation des effets, sinon sa faisabilité. Donc, l'évaluation n'a pas permis d'explorer l'association de la fidélité de la mise en œuvre de l'intervention aux résultats qui vise l'intervention à grande échelle. Ensuite, l'évaluation de la fidélité de la mise en œuvre de l'intervention a été effectuée dans un cadre de recherche, donc, cette fidélité de la mise en œuvre de l'intervention peut ne pas être complètement comparable à des situations réelles. L'étude pilote a été testée pendant 9 mois, ce qui empêche la détection des variations ou des écarts du protocole qui peuvent se produire dans une étude qui prend plus de temps, étant donné la dépendance au contexte que ce type d'intervention peut présenter. Enfin, les intervenants ont utilisé plusieurs stratégies pour maintenir des normes élevées de fidélité pour la mise en œuvre qui ne sont peut-être pas toutes réalisables à grande échelle. Cependant, la compréhension de la fidélité de la mise en œuvre de l'intervention dans sa phase pilote, sous le

contexte de la recherche et un devis d'évaluation formative, devra être la première étape nécessaire pour valider la théorie du programme et pour évaluer si l'intervention, elle-même, est fiable et possible à grande échelle. Donc, les résultats issus de cette évaluation sont pertinents et d'une grande utilité pour les concepteurs du programme.

Dans le même contexte, certaines limites peuvent avoir influencé les résultats de l'évaluation de la fidélité de la mise en œuvre de l'intervention. Par exemple, les enquêteurs communautaires ont fourni le volet de mobilisation communautaire, développant peut-être une relation avec les participants autour de l'intervention et étant affectés par un biais de désirabilité sociale, faussant ainsi les résultats (Althubaiti, 2016). Cependant, les enquêteurs (les intervenants-terrain) ont été impliqués dans le développement de l'intervention, et ils avaient une bonne compréhension de la mise en œuvre de l'intervention et du contexte lui-même, ce qui a permis de mieux se renseigner sur les expériences des participants dans les différentes activités de l'intervention. En même temps, en considérant que cette évaluation est de caractère « formative », sa présence pour la collecte des données a facilité la participation de la communauté pour obtenir des informations pertinentes pour explorer la faisabilité, l'acceptabilité de l'intervention et les modifications nécessaires avant de la mettre en œuvre à grande échelle. De même, les questions aux participants ont été formulées d'une manière qui permettait aux participants de se sentir « acceptés », quelle que soit la réponse, en formulant plusieurs options à choisir et dans un ordre approprié, passant de questions simples et générales à des questions plus spécifiques pour éviter des réponses inexactes ou biaisées. Également, les enquêteurs ont travaillé en équipes différentes, lors de chaque rencontre avec les communautés, pour éviter des biais de communication et pour faciliter la collecte de données de façon optimale et transparente.

En ce qui concerne la démarche évaluative, nous faisons ressortir trois limites importantes :

(i) *La présence limitée sur le terrain* et (ii) *la difficulté à communiquer dans une autre langue* ne permettent pas de faire une évaluation plus approfondie du cas. Bien que notre étude ne vise pas une démarche anthropologique, rester plus longtemps sur le terrain aurait facilité une meilleure compréhension du contexte, et un moyen de découvrir plus d'interactions entre les activités de l'intervention et les facteurs contextuels qui modulent le niveau de fidélité.

Des enjeux par la présence limitée sur le terrain ont été atténués en participant dans le groupe de travail dès le début du développement de l'intervention pour mieux connaître le contexte, et en planifiant un agenda de travail pour assurer la coordination entre la collecte et l'analyse des données, pour réduire les problèmes d'organisation et de gestion du temps, ainsi que pour optimiser au maximum le temps sur le terrain. Pour contrer les limites de la communication, une formation a été donnée à une personne externe pour faire les entretiens aux intervenants-terrain. Pour assurer la qualité des entretiens, nous avons expliqué à l'enquêteur chaque question, et des simulations ont été faites pour assurer la compréhension et le but des entretiens. Le fait même de recourir à une personne externe pour conduire l'entrevue peut réduire les biais de désirabilité sociale, car les personnes interviewées se sentent plus confortables à partager leurs expériences vécues. De même, les entrevues ont été enregistrées pour éviter de perdre de l'information lors de la traduction. Les entretiens ont été traduits par une autre personne qui maîtrisait l'anglais et l'Hindi. La traduction a été validée d'abord par une personne externe de l'équipe de recherche, et ensuite par le superviseur-terrain et par le concepteur de l'intervention, afin d'assurer la fiabilité et la crédibilité des traductions. Enfin, la validation des résultats auprès des intervenants a permis d'assurer la confirmabilité des données traduites. Ces mesures ont permis de minimiser le risque d'erreur d'interprétation lié à la traduction (Squires, 2009).

(iii) Ce projet de recherche s'inscrit dans un projet de recherche de ma directrice qui espère que l'intervention soit fidèle. Ainsi, la position d'évaluatrice-chercheuse qui fait partie d'un groupe de travail, et qui doit produire de façon objective des informations utiles et pertinentes pour les décideurs, n'est pas facile à concilier. En effet, la production des informations réelles plutôt que des informations convenables pour les différentes parties prenantes peut influencer et menacer la validité de l'évaluation (Brousselle et al., 2011). Pour minimiser ce biais, différentes sources telles que des entretiens, des focus groups, de multiples rencontres avec les responsables du programme, et la consultation de différents documents ont été utilisés pour trianguler les informations et pour assurer la validité du processus de recherche évaluative. De même, la tenue d'un journal de bord a permis de noter mes réflexions, de gérer mes impressions subjectives et de relater mon implication personnelle dans le projet. Enfin, le fait d'avoir accès à tous les documents en tout temps à l'aide d'un dossier partagé, et le fait de pouvoir communiquer en toute liberté avec le coordinateur-terrain en tout temps a facilité la collecte et l'interprétation objective des résultats.

Conclusion

La santé publique met de plus en plus l'accent sur l'importance des interventions fondées sur des données probantes pour améliorer la santé de la population et pour réduire les inégalités en matière de santé. Cependant, les interventions en santé publique ont tendance à varier pendant la phase d'implémentation, même si elles sont testées dans le cadre d'essais randomisés en grappes, car ces interventions sont dépendantes du contexte pouvant favoriser des écarts du protocole qui compromettent la validité des résultats.

Les résultats de cette thèse renforcent la pertinence de la fidélité de la mise en œuvre de l'intervention comme un outil essentiel pour ouvrir la boîte noire, afin de mieux comprendre les interactions qui se produisent pendant la phase de la mise en œuvre d'une intervention. Cette évaluation est essentielle pour les interventions bien établies, pour déterminer si les résultats s'expliquent par l'intervention même ou par des facteurs exogènes. Également, cette évaluation est essentielle pour les études pilotes afin d'évaluer la viabilité d'une intervention, de l'affiner et de la répliquer avec les meilleures preuves possibles dans d'autres contextes. Ainsi, cette évaluation contribue à accroître la confiance des décideurs dans leur sélection d'interventions testées par la recherche pour améliorer la qualité des programmes et pour aider à la réplification de l'intervention.

Enfin, les résultats de cette thèse ouvrent la voie pour améliorer la méthodologie évaluative de la fidélité de la mise en œuvre de l'intervention avec l'outil d'évaluation proposé, comme une initiative pour accroître la sensibilisation à la fidélité de la mise en œuvre de l'intervention, et pour mieux guider les chercheurs à entamer une telle démarche évaluative afin de promouvoir sa pratique de façon systématique.

References

- Abry, T., Hulleman, C. S., & Rimm-Kaufman, S. E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *American journal of evaluation, 36*(3), 320-338.
- Agarwal, S., LeFevre, A. E., Lee, J., L'Engle, K., Mehl, G., Sinha, C., & Labrique, A. (2016). Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ, 352*, i1174.
- Akobeng, A. K. (2005). Understanding randomised controlled trials. *Arch Dis Child, 90*(8), 840-844.
- Allison, R., Lecky, D., Town, K., Rugman, C., Ricketts, E., Ockendon-Powell, N., . . . McNulty, C. (2017). Exploring why a complex intervention piloted in general practices did not result in an increase in chlamydia screening and diagnosis: a qualitative evaluation using the fidelity of implementation model. *BMC family practice, 18*(1), 43.
- Allison, R., Lecky, D. M., Town, K., Rugman, C., Ricketts, E. J., Ockendon-Powell, N., . . . McNulty, C. A. (2017). Exploring why a complex intervention piloted in general practices did not result in an increase in chlamydia screening and diagnosis: a qualitative evaluation using the fidelity of implementation model. *BMC Fam Pract, 18*(1), 43. doi:10.1186/s12875-017-0618-0
- Althubaiti, A. J. J. o. m. h. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *9*, 211.
- Anderson, L. M., Petticrew, M., Chandler, J., Grimshaw, J., Tugwell, P., O'Neill, J., . . . Shemilt, I. (2013). Introducing a series of methodological articles on considering complexity in systematic reviews of interventions. *Journal of clinical epidemiology, 66*(11), 1205-1208.
- Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC medical research methodology, 10*(1), 67.
- Augustsson, H., von Thiele Schwarz, U., Stenfors-Hayes, T., & Hasson, H. (2015). Investigating variations in implementation fidelity of an organizational-level occupational health intervention. *International Journal of Behavioral Medicine, 22*(3), 345-355.
- Ayerbe, C., & Missonier, A. (2007). Validité interne et validité externe de l'étude de cas: principes et mise en œuvre pour un renforcement mutuel. *Finance Contrôle Stratégie, 10*(2), 37-62.

- Banerjee, A. V., Banerjee, A., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*: Public Affairs.
- Batt, K., Fox-Rushby, J. A., & Castillo-Riquelme, M. (2004). The costs, effects and cost-effectiveness of strategies to increase coverage of routine immunizations in low-and middle-income countries: systematic review of the grey literature. *Bulletin of the World Health Organization*, 82(9), 689-696.
- Bauman, L. J., Stein, R. E., & Ireys, H. T. (1991). Reinventing fidelity: The transfer of social technology among settings. *American journal of community psychology*, 19(4), 619-639.
- Beard, C., Eisen, R., Barker, C., Garofalo, J., Hahn, M., Hayden, M., . . . Schramm, P. (2016). *Ch. 5: Vectorborne Diseases*. Retrieved from
- Bédécarrats, F., Guérin, I., & Roubaud, F. (2018). The gold standard for randomised evaluations: from discussion of method to political economy.
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Treatment Fidelity Workgroup of the, N. I. H. B. C. C. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*, 23(5), 443-451. doi:10.1037/0278-6133.23.5.443
- Benson, K., & Hartz, A. J. (2000). A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine*, 342(25), 1878-1886.
- Blacklock, C., Bradley, D. C. G., Mikan, S., Willcox, M., Roberts, N., Bergstroem, A., & Mant, D. (2016). Impact of contextual factors on the effect of interventions to improve health worker performance in Sub-Saharan Africa: review of randomised clinical trials. *PLoS One*, 11(1).
- Blakely, C. H., Mayer, J. P., Gottschalk, R. G., Schmitt, N., Davidson, W. S., Roitman, D. B., & Emshoff, J. G. J. A. j. o. c. p. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *15(3)*, 253-268.
- Bland, J. M. (2004). Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol*, 4(1), 21. doi:10.1186/1471-2288-4-21
- Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H.-W. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental health services research*, 2(2), 75-87.
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V., & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ*, 333(7563), 346-349.

- Booth, A., Clarke, M., Dooley, G., Gherzi, D., Moher, D., Petticrew, M., & Stewart, L. (2012). The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Systematic reviews*, 1(1), 2.
- Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent*, 71(s1), S52–63.
- Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A. J., Czajkowski, S., Breger, R., . . . Ogedegbe, G. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of consulting and clinical psychology*, 73(5), 852.
- Bowen, S. (2012). Guide d'évaluation dans le domaine de la recherche en santé. . *Instituts de recherche en santé du Canada*. Consulté en ligne le 20 octobre 2018 en <http://www.cihr-irsc.gc.ca/f/45336.html>
- Bradley, F., Wiles, R., Kinmonth, A.-L., Mant, D., & Gantley, M. (1999). Development and evaluation of complex interventions in health services research: case study of the Southampton heart integrated care project (SHIP). *BMJ*, 318(7185), 711-715.
- Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., Resnick, B. J. R. i. n., & health. (2010). Implementation fidelity in community-based interventions. 33(2), 164-173.
- Breuer, E., Lee, L., De Silva, M., & Lund, C. (2015). Using theory of change to design and evaluate public health interventions: a systematic review. *Implementation Science*, 11(1), 63.
- Brierley, G., Brabyn, S., Torgerson, D., & Watson, J. (2012). Bias in recruitment to cluster randomized trials: a review of recent publications. *J Eval Clin Pract.* , 18(4), 878–886.
- Brooks, D., & Stratford, P. (2009). Les études pilotes: déterminer si elles peuvent être publiées dans *Physiotherapy Canada*. *Physiotherapy Canada*, 61(2), 67.
- Brousselle, A., Champagne, F., Contandriopoulos, A.-P., & Hartz, Z. (2011). *L'évaluation: concepts et méthodes: Deuxième édition*: Les Presses de l'Université de Montréal.
- Brown, C. H., Curran, G., Palinkas, L. A., Aarons, G. A., Wells, K. B., Jones, L., . . . Wallace, A. (2017). An overview of research and evaluation designs for dissemination and implementation. *Annual Review of Public Health*, 38, 1-22.
- Brown, P. A. (2008). A review of the literature on case study research. *Canadian Journal for New Scholars in Education/Revue canadienne des jeunes chercheuses et chercheurs en éducation*, 1(1).

- Brownson, R. C., Fielding, J. E., & Maylahn, C. M. (2009). Evidence-Based Public Health: A Fundamental Concept for Public Health Practice. *30*(1), 175-201. doi:10.1146/annurev.publhealth.031308.100134
- Cambon, L., Minary, L., Ridde, V., & Alla, F. (2013). A tool to analyze the transferability of health promotion interventions. *BMC public health, 13*(1), 1184.
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ, 321*(7262), 694-696. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10987780>
- Campbell, M., Katikireddi, S. V., Hoffmann, T., Armstrong, R., Waters, E., & Craig, P. (2018). TIDieR-PHP: a reporting guideline for population health and policy interventions. *BMJ, 361*.
- Campbell, M. J., Donner, A., & Klar, N. (2007). Developments in cluster randomized trials and Statistics in Medicine. *Stat Med, 26*(1), 2-19. doi:10.1002/sim.2731
- Campbell, M. K., Elbourne, D. R., Altman, D. G., & group, C. (2004). CONSORT statement: extension to cluster randomised trials. *BMJ, 328*(7441), 702-708. doi:10.1136/bmj.328.7441.702
- Campbell, M. K., Mollison, J., Steen, N., Grimshaw, J. M., & Eccles, M. (2000). Analysis of cluster randomized trials in primary care: a practical approach. *Fam Pract, 17*, 192–196.
- Campbell, M. K., Piaggio, G., Elbourne, D. R., Altman, D. G., & Group, C. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ, 345*, e5661. doi:10.1136/bmj.e5661
- Campbell, N. C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., . . . Kinmonth, A. L. (2007). Designing and evaluating complex interventions to improve health care. *BMJ, 334*(7591), 455-459.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci, 2*, 40. doi:10.1186/1748-5908-2-40
- Castro, F. G., Barrera, M., & Martinez, C. R. (2004). The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science, 5*(1), 41-45.
- Champagne, F., & Denis, J. L. (1990). L'analyse de l'implantation: modèles et méthodes. . *The Canadian Journal of Program Evaluation, 5*(2), 47-67.

- Champagne, F., & Denis, J.-L. (1992). Pour une évaluation sensible à l'environnement des interventions: l'analyse de l'implantation. *Service social*, 41(1), 143-163.
- Chen, H.-T. (2004). The roots of theory-driven evaluation. Current views and origins. *Evaluation Roots Thousand Oaks, Calif: Sage*.
- Chen, H.-T., & Rossi, P. H. (1989). Issues in the theory-driven perspective. *Evaluation and Program Planning*, 12(4), 299-306.
- Chen, H.-t., & Rossi, P. H. (1992). *Using theory to improve program and policy evaluations*: Greenwood Publishing Group.
- Chen, H. T. (2014). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective*: Sage Publications.
- Chen, H. T., & Chen, H.-t. (2005). *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*: Sage.
- Clarke, J. (2011). What is a systematic review? *Evidence-based nursing*, 14(3), 64-64.
- Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*, 342(25), 1887-1892.
- Contandriopoulos, A.-P., Champagne F, Potvin L, Denis JL, & P, B. (2005). *Savoir préparer une recherche: la définir, la structurer, la financer*: Montréal, Presses de l'Université de Montréal.
- Contandriopoulos, A.-P., Champagne, F., Denis, J.-L., & Pineault, R. (1993). *L'évaluation dans le domaine de la santé: concepts et méthodes*. Bulletin 33 (1): Les Presses de l'Université de Montréal.
- Contandriopoulos, A.-P., Rey, L., Brousselle, A., & Champagne, F. (2011). Évaluer une intervention complexe: enjeux conceptuels, méthodologiques, et opérationnels. *The Canadian journal of program evaluation= La Revue canadienne d'évaluation de programme*, 26(3), 1.
- Corley, N. A., & Kim, I. (2016). An assessment of intervention fidelity in published social work intervention research studies. *Research on Social Work Practice*, 26(1), 53-60.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and evaluating complex interventions: the new Medical Research Council guidance. *Int J Nurs Stud*, 50(5), 587-592. doi:10.1016/j.ijnurstu.2012.09.010

- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., & Medical Research Council, G. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, *337*, a1655. doi:10.1136/bmj.a1655
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*: Sage publications.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*: Sage publications.
- Dabbs, A. D., Song, M.-K., Hawkins, R., Aubrecht, J., Kovach, K., Terhorst, L., . . . Callan, J. J. N. r. (2011). An intervention fidelity framework for technology-based behavioral interventions. *60*(5), 340.
- Dagenais, C. (2006). Vers une utilisation accrue des résultats issus de la recherche par les intervenants sociaux. Quels modèles de transfert de connaissances privilégiés? *Les Sciences de l'éducation-Pour l'Ère nouvelle*, *39*(3), 23-35.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical psychology review*, *18*(1), 23-45.
- Davidoff, F. (2009). Heterogeneity is not always noise: lessons from improvement. *JAMA*, *302*(23), 2580-2586. doi:10.1001/jama.2009.1845
- de Sardan, J. O., & Ridde, V. (2014). Les spécificités des politiques publiques et des systèmes de santé en Afrique sahélienne. In *Une politique publique de santé et ses contradictions. La gratuité des soins au Burkina Faso, au Mali et au Niger*.
- De Silva, M. J., Breuer, E., Lee, L., Asher, L., Chowdhary, N., Lund, C., & Patel, V. (2014). Theory of change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials*, *15*(1), 267.
- Dehar, M.-A., Casswell, S., & Duignan, P. (1993). Formative and process evaluation of health promotion and disease prevention programs. *Evaluation Review*, *17*(2), 204-220.
- Díaz-Ordaz, K., Kenward, M. G., Cohen, A., Coleman, C. L., & Eldridge, S. (2014). Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines. *Clinical Trials*, *11*(5), 590-600.
- Dobson, D., & Cook, T. J. (1980). Avoiding type III error in program evaluation: Results from a field experiment. *Evaluation and Program Planning*, *3*(4), 269-276.

- Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*.
- Duffy, S. A., Cummins, S. E., Fellows, J. L., Harrington, K. F., Kirby, C., Rogers, E., . . . Waltje, A. H. (2015). Fidelity monitoring across the seven studies in the Consortium of Hospitals Advancing Research on Tobacco (CHART). *Tobacco induced diseases*, 13(1), 29.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4, 3895-3962.
- Dumont, A. (2019). Les essais randomisés en grappe: un exemple en santé maternelle et infantile. In.
- Durlak, J. A., & DuPre, E. P. J. A. j. o. c. p. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *41*(3-4), 327-350.
- Durlak, J. A., & Wells, A. M. J. A. j. o. c. p. (1998). Evaluation of indicated preventive intervention (secondary prevention) mental health programs for children and adolescents. *26*(5), 775-802.
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res*, 18(2), 237-256. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/12729182>
- Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. In: Springer.
- Eisenhardt, K. M. (1989). Building Theories From Case Study Research. *Academy of Management. The Academy of Management Review*, 14(4), 532.
- Eldridge, S., Ashby, D., Bennett, C., Wakelin, M., & Feder, G. (2008). Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ*, 336(7649), 876–880.
- Eldridge, S. M., Ashby, D., Feder, G. S., Rudnicka, A. R., & Ukoumunne, O. C. (2004). Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clinical Trials*, 1(1), 80-90.
- Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L., & Bond, C. M. (2016). Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PLoS One*, 11(3).
- Fassin, D. (1996). *L'espace politique de la santé: essai de généalogie: FeniXX*.

- Fassin, D. (2005). *Faire de la santé publique*: Ecole des Hautes Etudes en Santé Publique.
- Fetterman, D. M. (2005). A Window into the Heart and Soul of Empowerment Evaluation: Looking through the Lens of Empowerment Evaluation Principles.
- Fink, G., Günther, I., & Hill, K. (2014). Slum residence and child health in developing countries. *Demography*, *51*(4), 1175-1197.
- Fixsen, D. L., & Blase, K. A. (2009). Implementation: The Missing Link between Research and Practice. Implementation Brief. Number 1. *FPG Child Development Institute*.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., Wallace, F., Burns, B., . . . Barwick, M. (2005). Implementation research: A synthesis of the literature.
- Flottorp, S., Håvelsrud, K., & Oxman, A. D. (2003). Process evaluation of a cluster randomized trial of tailored interventions to implement guidelines in primary care—why is it so hard to change practice? *Family practice*, *20*(3), 333-339.
- Fontan, J.-M. (2001). L'évaluation de cinquième génération. *Cahiers de recherche sociologique*(35), 7-14.
- Gagnon, Y.-C. (2005). *L'étude de cas comme méthode de recherche: guide de réalisation*: PUQ.
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC medical research methodology*, *13*(1), 117.
- Gautier, L., Pirkle, C. M., Furgal, C., & Lucas, M. (2016). Assessment of the implementation fidelity of the Arctic Char Distribution Project in Nunavik, Quebec. *BMJ Global Health*, *1*(3), e000093.
- Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin Psychol Rev*, *31*(1), 79-88. doi:10.1016/j.cpr.2010.09.007
- Gittelsohn, J., Evans, M., Story, M., Davis, S. M., Metcalfe, L., Helitzer, D. L., & Clay, T. E. (1999). Multisite formative assessment for the Pathways study to prevent obesity in American Indian schoolchildren. *Am J Clin Nutr*, *69*(4), 767S-772S.
- Glasgow, R. E., & Emmons, K. M. J. A. R. P. H. (2007). How can we increase translation of research into practice? Types of evidence needed. *28*, 413-433.

- Graham, L., Wright, J., Walwyn, R., Russell, A. M., Bryant, L., Farrin, A., & House, A. (2016). Measurement of adherence in a randomised controlled trial of a complex intervention: supported self-management for adults with learning disability and type 2 diabetes. *BMC medical research methodology*, *16*(1), 132.
- Grant, A., Treweek, S., Dreischulte, T., Foy, R., & Guthrie, B. (2013). Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials*, *14*, 15. doi:10.1186/1745-6215-14-15
- Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard? *Perspectives in biology and medicine*, *48*(4), 516-534.
- Guichard, A., Tardieu, É., Nour, K., Lafontaine, G., & Ridde, V. (2019). Adapting a health equity tool to meet professional needs (Québec, Canada). *Health promotion international*, *34*(6), e71-e83.
- Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and “ethically important moments” in research. *Qualitative inquiry*, *10*(2), 261-280.
- Gurnani, V., Haldar, P., Aggarwal, M. K., Das, M. K., Chauhan, A., Murray, J., . . . Sudan, P. (2018). Improving vaccination coverage in India: lessons from Intensified Mission Indradhanush, a cross-sectoral systems strengthening strategy. *363*, k4782.
- Hahn, S., Puffer, S., Torgerson, D. J., & Watson, J. (2005). Methodological bias in cluster randomised trials. *BMC medical research methodology*, *5*(1), 10.
- Hamelin, A.-M., Caux, C., Désy, M., Guichard, A., Ouédraogo, S., Tremblay, M.-C., . . . Godard, B. (2018). Développer une culture de l'éthique en recherche interventionnelle en santé des populations. *Éthique publique. Revue internationale d'éthique sociétale et gouvernementale*, *20*(2).
- Handlos, L. N., Chakraborty, H., & Sen, P. K. (2009). Evaluation of cluster-randomized trials on maternal and child health research in developing countries. *Tropical Medicine & International Health*, *14*(8), 947-956.
- Hansen, W. B., Graham, J. W., Wolkenstein, B. H., & Rohrbach, L. A. (1991). Program integrity as a moderator of prevention program effectiveness: results for fifth-grade students in the adolescent alcohol prevention trial. *Journal of studies on alcohol*, *52*(6), 568-579.
- Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—The gold standard for effectiveness research. *BJOG: an international journal of obstetrics and gynaecology*, *125*(13), 1716.

- Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children*, 79(2), 181-193.
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci*, 5, 67. doi:10.1186/1748-5908-5-67
- Hasson, H., Blomberg, S., & Dunér, A. (2012). Fidelity and moderating factors in complex interventions: a case study of a continuum of care program for frail elderly people in health and social care. *Implement Sci*, 7(23), 1–11.
- Havas, S., Anliker, J., Damron, D., Feldman, R., & Langenberg, P. (2000). Uses of process evaluation in the Maryland WIC 5-a-day promotion program. *Health Education & Behavior*, 27(2), 254-263.
- Hayes, R. J., Alexander, N. D., Bennett, S., & Cousens, S. N. (2000). Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Methods Med Res*, 9(2), 95–116.
- Heymann, D. L., & Shindo, N. (2020). COVID-19: what is next for public health? *The Lancet*, 395(10224), 542-545.
- Higgins, J., & Wells, G. (2011). Cochrane handbook for systematic reviews of interventions.
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, 348, g1687. doi:10.1136/bmj.g1687
- Horner, S., Rew, L., & Torres, R. (2006). Enhancing intervention fidelity: a means of strengthening study impact. *Journal for Specialists in Pediatric Nursing*, 11(2), 80-89.
- Hulscher, M., Laurant, M., & Grol, R. (2003). Process evaluation on quality improvement interventions. *BMJ Quality & Safety*, 12(1), 40-46.
- Ibrahim, S., & Sidani, S. (2015). Fidelity of intervention implementation: a review of instruments. *Health*, 2015, 1687.
- IIPS. (2016). National Family Health Survey (NFHS-4), 2015–16: India: . Retrieved from <http://rchiips.org/nfhs/pdf/NFHS4/India.pdf>
- Isaakidis, P., & Ioannidis, J. P. (2003). Evaluation of cluster randomized controlled trials in sub-Saharan Africa. *Am J Epidemiol*, 158(9), 921–926.

- Ivers, N., Taljaard, M., Dixon, S., Bennett, C., McRae, A., Taleban, J., . . . Eccles, M. (2011). Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. *BMJ*, *343*, d5886.
- Jacob, K. (2009). Public health in low-and middle-income countries and the clash of cultures. *Journal of Epidemiology & Community Health*, *63*(7), 509-509.
- Jamal, F., Fletcher, A., Shackleton, N., Elbourne, D., Viner, R., & Bonell, C. (2015). The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example. *Trials*, *16*(1), 466.
- Jean-Louis, D., & Champagne, F. (1990). L'analyse de l'implantation: modèles et méthodes. *The Canadian Journal of Program Evaluation*, *5*(2), 47.
- Jenicek, M. (1997). Epidemiology, evidenced-based medicine, and evidence-based public health. *Journal of epidemiology*, *7*(4), 187-197.
- Johri, M., Pérez, M. C., Arsenault, C., Sharma, J. K., Pai, N. P., Pahwa, S., & Sylvestre, M.-P. (2015). Strategies to increase the demand for childhood vaccination in low-and middle-income countries: a systematic review and meta-analysis. *Bulletin of the World Health Organization*, *93*(5), 339-346.
- Jurg, M. E., De Meij, J. S., Van der Wal, M. F., & Koelen, M. A. (2008). Using health promotion outcomes in formative evaluation studies to predict success factors in interventions: an application to an intervention for promoting physical activity in Dutch children (JUMP-in). *Health promotion international*, *23*(3), 231-239.
- Keith, R. E., Hopp, F. P., Subramanian, U., Wiitala, W., & Lowery, J. C. (2010). Fidelity of implementation: development and testing of a measure. *Implementation Science*, *5*(1), 99.
- Kendall, J. (2003). Designing a research project: randomised controlled trials and their principles. *Emergency medicine journal: EMJ*, *20*(2), 164.
- Kilbourne, A. M., Neumann, M. S., Pincus, H. A., Bauer, M. S., & Stall, R. (2007). Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implementation Science*, *2*(1), 42.
- Kumanyika, S. K., Story, M., Beech, B. M., Sherwood, N. E., Baranowski, J. C., Powell, T. M., . . . Owens, A. S. (2003). Collaborative planning for formative assessment and cultural appropriateness in the Girls health Enrichment Multi-site Studies (GEMS): a retrospection. *Ethnicity and Disease*, *13*(1; SUPP/1), S1-15.

- Labonté, R., & Schrecker, T. (2007). Globalization and social determinants of health: Introduction and methodological background (part 1 of 3). *Globalization and Health*, 3(1), 5.
- Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of evaluation in clinical practice*, 10(2), 307-312.
- Laperrière, A. (1997). Les critères de scientificité des méthodes qualitatives. *La recherche qualitative. Enjeux épistémologiques et méthodologiques*, 1, 365-389.
- Lawton, R., Mceachan, R., Jackson, C., West, R., & Conner, M. (2014). Intervention fidelity and effectiveness of a UK worksite physical activity intervention funded by the Bupa Foundation, UK. *Health promotion international*, 30(1), 38-49.
- Lawton, R., O'Hara, J. K., Sheard, L., Armitage, G., Cocks, K., Buckley, H., . . . Moore, S. (2017). Can patient involvement improve patient safety? A cluster randomised control trial of the Patient Reporting and Action for a Safe Environment (PRASE) intervention. *BMJ Quality & Safety*, 26(8), 622-631.
- Leyrat, C., Morgan, K. E., Leurent, B., & Kahan, B. C. (2018). Cluster randomized trials with a small number of clusters: which analyses should be used? *International Journal of Epidemiology*, 47(1), 321-331.
- Ligthelm, R. J., Borzi, V., Gumprecht, J., Kawamori, R., Wenying, Y., & Valensi, P. (2007). Importance of observational studies in clinical practice. *Clinical therapeutics*, 29(6), 1284-1292.
- Liu, L., Chu, Y., Oza, S., Hogan, D., Perin, J., Bassani, D. G., . . . Dhingra, N. (2019). National, regional, and state-level all-cause and cause-specific under-5 mortality in India in 2000–15: a systematic analysis with implications for the Sustainable Development Goals. *The Lancet Global Health*, 7(6), e721-e734.
- Lorençatto, F., Gould, N. J., McIntyre, S. A., During, C., Bird, J., Walwyn, R., . . . Stanworth, S. J. (2016). A multidimensional approach to assessing intervention fidelity in a process evaluation of audit and feedback interventions to reduce unnecessary blood transfusions: a study protocol. *Implementation Science*, 11(1), 163.
- Lozano, R., Fullman, N., Abate, D., Abay, S. M., Abbafati, C., Abbasi, N., . . . Abdelalim, A. (2018). Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392(10159), 2091-2138.

- Mann, C., Shaw, A., Guthrie, B., Wye, L., Man, M.-S., Hollinghurst, S., . . . Salisbury, C. (2016). Protocol for a process evaluation of a cluster randomised controlled trial to improve management of multimorbidity in general practice: the 3D study. *BMJ open*, *6*(5), e011260.
- Masic, I., Miokovic, M., & Muhamedagic, B. (2008). Evidence based medicine—new approaches and challenges. *Acta Informatica Medica*, *16*(4), 219.
- McArthur, B. A., Riosa, P. B., & Preyde, M. (2012). Treatment fidelity in psychosocial intervention for children and adolescents with comorbid problems. *Child and Adolescent Mental Health*, *17*(3), 139-145.
- Medical Research, & Council. (2008). *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council.
- Mihalic, S. (2004). The importance of implementation fidelity. *Emotional and Behavioral Disorders in Youth*, *4*(4), 83-105.
- Mihalic, S. F., Fagan, A. A., & Argamaso, S. (2008). Implementing the LifeSkills Training drug prevention program: factors related to implementation fidelity. *Implementation Science*, *3*(1), 5.
- Miles, M. B., & Huberman, A. M. (2003). *Analyse des données qualitatives*: De Boeck Supérieur.
- Moberg, J., & Kramer, M. (2015). A brief history of the cluster randomised trial design. *J R Soc Med*, *108*(5), 192-198. doi:10.1177/0141076815582303
- Moher D, Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6* (7), e1000097.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*, *340*, c869. doi:10.1136/bmj.c869
- Moir, T. (2018). *Why is implementation science important for intervention design and evaluation within educational settings?* Paper presented at the Frontiers in Education.
- Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical psychology review*, *11*(3), 247-266.
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, *350*, h1258.

- Morello, R. T., Barker, A. L., Ayton, D. R., Landgren, F., Kamar, J., Hill, K. D., . . . Stoelwinder, J. (2017). Implementation fidelity of a nurse-led falls prevention program in acute hospitals during the 6-PACK trial. *BMC Health Serv Res*, *17*(1), 383. doi:10.1186/s12913-017-2315-z
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American journal of evaluation*, *24*(3), 315-340.
- Mukherjee, J. S., Mugunga, J. C., Shah, A., Leta, A., Birru, E., Oswald, C., . . . Yates, R. (2019). A practical approach to universal health coverage. *The Lancet Global Health*, *7*(4), e410-e411.
- Murtagh, M., Thomson, R., May, C., Rapley, T., Heaven, B., Graham, R., . . . Eccles, M. (2007). Qualitative methods in a randomised controlled trial: the role of an integrated qualitative process evaluation in providing evidence to discontinue the intervention in one arm of a trial of a decision support tool. *BMJ Quality & Safety*, *16*(3), 224-229.
- Muse, P., & Gagnon, Y.-C. (2012). *L'étude de cas comme méthode de recherche, 2e édition*: Presses de l'Université du Québec.
- Napúa, M., Pfeiffer, J. T., Chale, F., Hoek, R., Manuel, J., Michel, C., . . . Sherr, K. (2016). Option B+ in Mozambique: formative research findings for the design of a facility-level clustered randomized controlled trial to improve ART retention in antenatal care. *Journal of acquired immune deficiency syndromes (1999)*, *72*(Suppl 2), S181.
- National Academies of Sciences, E., & Medicine. (2016). *Global health impacts of vector-borne diseases: workshop summary*: National Academies Press.
- Nations-Unies. (2015). Objectifs du Millénaire pour le développement. *Disponible sur https://www.un.org/fr/millenniumgoals/reports/2015/pdf/rapport_2015.pdf*. Retrieved from https://www.un.org/fr/millenniumgoals/reports/2015/pdf/rapport_2015.pdf
- Nebot, M., López, M. J., Ariza, C., Villalbí, J. R., & García-Altés, A. (2011). Evaluación de la efectividad en salud pública: fundamentos conceptuales y metodológicos. *Gaceta sanitaria*, *25*, 3-8.
- Noyes, J., Gough, D., Lewin, S., Mayhew, A., Michie, S., Pantoja, T., . . . Shemilt, I. (2013). A research and development agenda for systematic reviews that ask complex questions about complex interventions. *Journal of clinical epidemiology*, *66*(11), 1262-1270.
- Nutbeam, D. (1998). Health promotion glossary. *Health promotion international*, *13*(4), 349-364.

- O'cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research & Policy*, 13(2), 92-98.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, 78(1), 33-84.
- Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ*, 332(7583), 413-416.
- Olivier de Sardan, J.-P. (1995). La politique du terrain. Sur la production des données en anthropologie. *Enquête. Archives de la revue Enquête*(1), 71-109.
- Orwin, R. G. (2000). Assessing program fidelity in substance abuse health services research. *Addiction*, 95(11s3), 309-327.
- Osrin, D., Azad, K., Fernandez, A., Manandhar, D. S., Mwansambo, C. W., Tripathy, P., & Costello, A. M. (2009). Ethical challenges in cluster randomized controlled trials: experiences from public health interventions in Africa and Asia. *Bulletin of the World Health Organization*, 87, 772-779.
- Oyo-Ita, A., Wiysonge, C. S., Oringanje, C., Nwachukwu, C. E., Oduwale, O., & Meremikwu, M. M. J. C. D. o. S. R. (2016). Interventions for improving coverage of childhood immunisation in low-and middle-income countries. (7).
- Panda, B. K., & Mohanty, S. K. (2019). Progress and prospects of health-related sustainable development goals in india. *Journal of biosocial science*, 51(3), 335-352.
- Paquette, G., Joly, J., & Tourigny, M. (2009). La fidélité de l'implantation d'un programme d'intervention pour des adolescentes agressées sexuellement: une mesure par observation systématique *The Canadian Journal of Program Evaluation*, 24(2), 57.
- Parsaik, A. K., & Singh, B. (2013). Evidence Based Public Health. *Current Topics in Public Health*, 283.
- Patton, M. (2008). Utilization-focused evaluation. SAGE Publications. Inc., California.
- Patton, M. Q. (2002). Qualitative research and evaluation methods. Thousand Oaks. Cal.: Sage Publications, 37-73.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*: Sage publications.

- Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical psychology: Science and practice, 12*(4), 365-383.
- Perez, D., Van der Stuyft, P., Zabala, M. C., Castro, M., & Lefevre, P. (2016). A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. *Implement Sci, 11*(1), 91. doi:10.1186/s13012-016-0457-8
- Perez, M. C., Minoyan, N., Ridde, V., Sylvestre, M. P., & Johri, M. (2016). Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review protocol. *Syst Rev, 5*(1), 177. doi:10.1186/s13643-016-0351-0
- Perez, M. C., Minoyan, N., Ridde, V., Sylvestre, M. P., & Johri, M. (2018). Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in low- and middle-income countries: systematic review. *Trials, 19*(1), 410. doi:10.1186/s13063-018-2796-z
- Pérez, M. C., Singh, R., Chandra, D., Ridde, V., Seth, A., & Johri, M. (2020). *Development of an mHealth Behavior Change Communication Strategy: A case-study from rural Uttar Pradesh in India*. Paper presented at the Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies, Ecuador. <https://doi.org/10.1145/3378393.3402505>
- Peters, D. H., Adam, T., Alonge, O., Agyepong, I. A., & Tran, N. (2013). Implementation research: what it is and how to do it. *347*, f6753. doi:10.1136/bmj.f6753 %J BMJ : British Medical Journal
- Pluye, P., Bengoechea, E. G., & Granikov, V. J. I. J. M. R. A. (2018). A world of possibilities in mixed methods: review of the combinations of strategies used to integrate qualitative and quantitative phases, results and data. *10*, 41-56.
- Pluye, P., L. Potvin, et al. . (2004). Making public health programs last: conceptualizing sustainability. *Evaluation and Program Planning 27*(2): 121-133.
- Pommier, J., & Grimaud, O. (2007). Les fonctions essentielles de santé publique: histoire, définition et applications possibles. *Santé publique, 19*(hs), 9-14.
- Proctor, E. K. (2004). Leverage points for the implementation of evidence-based practice. *Brief Treatment & Crisis Intervention, 4*(3).
- Puffer, S., Torgerson, D., & Watson, J. (2003). Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ, 327*(7418), 785-789.

- Puffer, S., Torgerson, D., & Watson, J. (2003). Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ*, *327*(7418), 785–789.
- R. Makkar, S., Brennan, S., Turner, T., Williamson, A., Redman, S., & Green, S. (2016). The development of SAGE: a tool to evaluate how policymakers’ engage with and use research in health policymaking. *Research Evaluation*, *25*(3), 315-328.
- Ranson, M. K., Sinha, T., Morris, S. S., & Mills, A. J. (2006). CRTs—cluster randomized trials or “courting real troubles”: challenges of running a CRT in rural Gujarat, India. *Can J Public Health*, *97*(1), 72.
- Rebchook, G. M., Kegeles, S. M., Huebner, D., & Team, T. R. (2006). Translating research into practice: the dissemination and initial implementation of an evidence-based HIV prevention program. *AIDS Education & Prevention*, *18*(supp), 119-136.
- Resnick, B., Bellg, A. J., Borrelli, B., De Francesco, C., Breger, R., Hecht, J., . . . Ernst, D. J. A. o. B. M. (2005). Examples of implementation and evaluation of treatment fidelity in the BCC studies: where we are and where we need to go. *29*(2), 46.
- Reviews, C. f., & Dissemination. (2009). *CRD's guidance for undertaking reviews in healthcare*: York Publ. Services.
- Rezze, B. D., Law, M., Eva, K., Pollock, N., & Gorter, J. W. (2013). Development of a generic fidelity measure for rehabilitation intervention research for children with physical disabilities. *Developmental Medicine & Child Neurology*, *55*(8), 737-744.
- Richards, D. A., & Hallberg, I. R. (2015). *Complex interventions in health: an overview of research methods*. London & New York: Routledge.
- Ridde, V. (2007). Réduire les inégalités sociales de santé: santé publique, santé communautaire ou promotion de la santé? *Promotion & Education*, *14*(2), 111-114.
- Ridde, V., & Capelle, F. (2011). La recherche en santé mondiale et les défis des partenariats Nord-Sud. *Canadian journal of public health*, *102*(2), 152-156.
- Ridde, V., & Dagenais, C. (2009). *Approches et pratiques en évaluation de programme*. . Les Presses de l'Université de Montréal, 358.
- Ridde, V., & Dagenais, C. (2019). *Évaluation des interventions de santé mondiale: méthodes avancées*.

- Ridde, V., Druetz, T., Poppy, S., Kouanda, S., & Haddad, S. (2013). Implementation fidelity of the national malaria control program in Burkina Faso. *PLoS One*, *8*(7).
- Ridde, V., & Haddad, S. (2013). Pragmatisme et réalisme pour l'évaluation des interventions de santé publique. *Revue d'épidémiologie et de santé publique*, *61*, S95-S106.
- Ridde, V., Pérez, D., & Robert, E. (2020). Using implementation science theories and frameworks in global health. *BMJ Global Health*, *5*(4), e002269.
- Ridde, V., Pommier, J., & Jabot, F. (2010). Spécificités de l'évaluation des programmes visant la réduction des inégalités. . Dans Potvin, S., Moquet, M.-J. et Jones, C. (dir.), *Réduire les inégalités sociales en santé. Institut national de prévention et d'éducation pour la santé, coll. Santé en action. Saint-Denis*, , 257-269.
- Ritchie, J., & Spencer, L. J. T. q. r. s. c. (2002). Qualitative data analysis for applied policy research. *573*(2002), 305-329.
- Roser, M., Ritchie, H., & Dadonaite, B. (2013). Child & Infant Mortality. *Our World in Data*.
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach*: Sage publications.
- Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *J Epidemiol Community Health*, *56*, 119–127. doi:doi:10.1136/jech.56.2.119
- Rychetnik, L., Hawe, P., Waters, E., Barratt, A., & Frommer, M. (2004). A glossary for evidence based public health. *J Epidemiol Community Health*, *58*(7), 538–545.
- Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. In: British Medical Journal Publishing Group.
- Saetren, H. (2005). Facts and myths about research on public policy implementation: Out-of-Fashion, allegedly dead, but still very much alive and relevant. *Policy Studies Journal*, *33*(4), 559-582.
- Samb, O. M., Ridde, V., & Queuille, L. (2013). Quelle pérennité pour les interventions pilotes de gratuité des soins au Burkina Faso? *Revue Tiers Monde*(3), 73-91.
- Sanson-Fisher, R. W., Bonevski, B., Green, L. W., & D'Este, C. (2007). Limitations of the randomized controlled trial in evaluating population-based health interventions. *Am J Prev Med*, *33*(2), 155-161.

- Scantlebury, A., Cockayne, S., Fairhurst, C., Rodgers, S., Torgerson, D., Hewitt, C., . . . study, R. (2020). Qualitative research to inform hypothesis testing for fidelity-based sub-group analysis in clinical trials: lessons learnt from the process evaluation of a multifaceted podiatry intervention for falls prevention. *Trials*, *21*, 1-10.
- Schoenwald, S. K., Henggeler, S. W., Brondino, M. J., & Rowland, M. D. (2000). Multisystemic therapy: Monitoring treatment fidelity. *Family Process*, *39*(1), 83-103.
- Schorr, L. B. (1998). *Common purpose: Strengthening families and neighborhoods to rebuild America*: Anchor Books.
- Scriven, M. (2008). A summative evaluation of RCT methodology: an alternative approach to causal research. *Journal of multidisciplinary evaluation*, *5*(9), 11-24.
- Severino, J.-M. (2008). La situation des pays en voie de développement. *Les Tribunes de la sante*(4), 31-39.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference/William R. Shadish, Thomas D. Cook, Donald T. Campbell*: Boston: Houghton Mifflin.
- Shanyinde, M., Pickering, R. M., & Weatherall, M. (2011). Questions asked and answered in pilot and feasibility randomized controlled trials. *BMC medical research methodology*, *11*(1), 117.
- Shea, B., Andersson, N., & Henry, D. (2009). Increasing the demand for childhood vaccination in developing countries: a systematic review. *BMC international health and human rights*, *9*(Suppl 1), S5.
- Shrivastwa, N., Wagner, A. L., & Boulton, M. L. (2019). Analysis of State-Specific Differences in Childhood Vaccination Coverage in Rural India. *Vaccines*, *7*(1), 24.
- Sidani, S., & Sechrest, L. (1999). Putting program theory into operation. *American journal of evaluation*, *20*(2), 227-238.
- Silverman, S. L. (2009). From randomized controlled trials to observational studies. *The American journal of medicine*, *122*(2), 114-120.
- Slaughter, S. E., Hill, J. N., & Snelgrove-Clarke, E. (2015). What is the extent and quality of documentation and reporting of fidelity to implementation strategies: a scoping review. *Implementation Science*, *10*(1), 129.

- Sondaal, S. F. V., Browne, J. L., Amoakoh-Coleman, M., Borgstein, A., Miltenburg, A. S., Verwijs, M., & Klipstein-Grobusch, K. (2016). Assessing the Effect of mHealth Interventions in Improving Maternal and Neonatal Care in Low- and Middle-Income Countries: A Systematic Review. *PLoS One*, *11*(5), e0154664.
- Squires, A. (2009). Methodological challenges in cross-language qualitative research: A research review. *International journal of nursing studies*, *46*(2), 277-287.
- Stake, R. E. (2008). Qualitative case studies.
- Steckler, A. B., Linnan, L., & Israel, B. (2002). *Process evaluation for public health interventions and research*: Jossey-Bass San Francisco, CA.
- Stetler, C. B., Legro, M. W., Wallace, C. M., Bowman, C., Guihan, M., Hagedorn, H., . . . Smith, J. L. (2006). The role of formative evaluation in implementation research and the QUERI experience. *Journal of general internal medicine*, *21*(2), S1.
- Stevens, K. R. (2001). Systematic reviews: the heart of evidence-based practice. *AACN Advanced Critical Care*, *12*(4), 529-538.
- Suchman, E. A. (1967). Principles and practice of evaluative research. *An introduction to social research*, 327.
- Sullivan, G. M. (2011). Getting off the “gold standard”: randomized controlled trials and education research. *Journal of graduate medical education*, *3*(3), 285-289.
- Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., . . . Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol*, *10*, 1. doi:10.1186/1471-2288-10-1
- The World, & Bank. Country and lending groups. Data & statistics: country classification (2016). In Washington, D.C: . Available at: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519> Accessed 3 July,2016.
- Titler, M. G. (2008). The evidence for evidence-based practice implementation. In *Patient safety and quality: An evidence-based handbook for nurses*: Agency for Healthcare Research and Quality (US).
- Tizio, S. (2004). Etat de santé et systèmes de soins dans les pays en développement: La contribution des politiques de santé au développement durable. *Mondes en développement*(3), 101-117.
- Turcotte-Tremblay, A.-M., Fregonese, F., Kadio, K., Alam, N., & Merry, L. (2020). Global health is more than just ‘Public Health Somewhere Else’. *BMJ Global Health*, *5*(5), e002545.

- Upshur, R. E. (2002). Principles for the justification of public health intervention. *Can J Public Health, 93*(2), 101–103.
- Van Belle, S., Wong, G., Westhorp, G., Pearson, M., Emmel, N., Manzano, A., & Marchal, B. (2016). Can “realist” randomised controlled trials be genuinely realist? *Trials, 17*(1), 313.
- van Breukelen, G. J., & Candel, M. J. (2018). How to design and analyse cluster randomized trials with a small number of clusters? Comment on Leyrat et al. *International Journal of Epidemiology*.
- Vishwakarma, M., Shekhar, C., Dutta, M., & Yadav, A. (2019). Gaps in infant and child mortality among social groups and its linkages with institutional delivery and child immunization using census and National Family Health Survey (2015-16). *Journal of Public Health, 1*-11.
- von Thiele Schwarz, U., Hasson, H., & Lindfors, P. (2015). Applying a fidelity framework to understand adaptations in an occupational health intervention. *Work, 51*(2), 195-203.
- Wang, B., Stanton, B., Deveaux, L., Poitier, M., Lunn, S., Koci, V., . . . Li, X. J. I. s. (2015). Factors influencing implementation dose and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program. *10*(1), 44.
- Wang, H., Bhutta, Z. A., Coates, M. M., Coggeshall, M., Dandona, L., Diallo, K., . . . Gething, P. W. (2016). Global, regional, national, and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1725-1774.
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New directions for evaluation, 1997*(76), 41-55.
- Wells, M., Williams, B., Treweek, S., Coyle, J., & Taylor, J. (2012). Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials, 13*(1), 95.
- WHO. (2016). *World health statistics 2016: monitoring health for the SDGs sustainable development goals*: World Health Organization.
- WHO. (2017). Tracking universal health coverage: 2017 global monitoring report.
- WHO. (2018). Immunization coverage- Key challenges. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/immunization-coverage>

Wynn, B. O., Dutta, A., & Nelson, M. I. (2006). *Challenges in program evaluation of health interventions in developing countries*: Rand Corporation.

Yin, R. (2009). Designing Case Studies: Identifying your case (s) and establishing the logic of your case study. *Case Study Research: Design and Methods*, 5, 25-66.

Yin, R. K. (2017). *Case study research and applications: Design and methods*: Sage publications.

Annexes

Annexe 1 Grille pour l'extraction de données-Revue systématique

Identification de l'étude				Identification évaluateur			
Évaluation de la fidélité de la mise en œuvre de l'intervention concernant l'adhésion rapportée dans le protocole				Évaluation de la fidélité de la mise en œuvre de l'intervention concernant l'adhésion rapportée dans l'article principal de l'essai			
Dimension + méthode utilisée	Emplacement*			Dimension + méthode utilisée	Emplacement*		
	a	b	c		d	e	f
Contenu							
Couverture							
Durée							
Fréquence							
Commentaires				Commentaires			
Facteurs modérateurs évalués							
Dimension + méthode utilisée	Emplacement*			Dimension + méthode utilisée	Emplacement*		
	a	b	c		d	e	f
Complexité de l'intervention							
Stratégies facilitant la mise en œuvre							

Qualité de la mise en œuvre							
Réponse des participants							
Recrutement							
Contexte							
Groupe contrôle							
Commentaires				Commentaires			

* Emplacement de la dimension rapportée

- a) Registre de l'essai
- b) Protocole publié
- c) Littérature grise
- d) Document principal
- e) Évaluation du processus
- f) Littérature grise

Annexe 2 Grille structurée pour les petites rencontres communautaires

Article 3: Additional file 4 Small group meeting format

First small group meeting

Meeting date:

Meeting starting time:

State name: Uttar Pradesh

District Name: Hardoi

Block Name: Bawan

Village Name:

Village code:

ASHA Name:

ASHA phone no.

AWW Name:

AWW phone no.

TikaVaani team members name

A. Activities in Village				
1	Meeting with community workers	Does task done	Remarks Note: Please comment	If yes then why and what modification was done
a.	ASHA	Yes – 1 No – 2		Yes – 1 No – 2

b.	Aanganwadi worker	Yes - 1 No - 2		Yes - 1 No - 2	
c.	Aanganwadi helper	Yes - 1 No - 2		Yes - 1 No - 2	
d.	Were newborn children information collected from Aanganwadi worker or ASHA?	Yes - 1 No - 2		Yes - 1 No - 2	
2	Visit of selected households	Does task done	Remarks Note: Please comment	Was the task modified	If yes then why and what modification was done
a.	Was the team member visited target household to invite them to the small group meeting?	Yes - 1 No - 2		Yes - 1 No - 2	
b.	Were mobile numbers collected from the selected households?	Yes - 1		Yes - 1 No - 2	

		No – 2			
c.	Was the meeting place clam and peaceful?	Yes – 1 No – 2		Yes – 1 No – 2	
3	Did discussion have with the houses to choose a place for a small group meeting?	Yes – 1 No – 2		Yes – 1 No – 2	
4	Wall painting in the village	Yes – 1 No – 2		Yes – 1 No – 2	
5	Poster pasting in the village	Yes – 1 No – 2		Yes – 1 No – 2	
B. Small group activities					

Activities		Does task done	Remarks Note: Please comment				Was the task modified				If yes then why and what modification was done
1.	Was introduction activity done before starting the meeting?	Yes - 1 No - 2					Yes - 1 No - 2				
2.	Was TikaVaani number demonstrated?	Yes - 1 No - 2					Yes - 1 No - 2				
3.	Were TikaVaani slips distributed to everyone in the meeting?	Yes - 1 No - 2					Yes - 1 No - 2				
4. Selected households participation in the meeting											
1	Selected households code in the village (Use baseline survey code)	HH Code _____	HH Code _____	HH Code _____	HH Code _____	HH Code _____	HH Code _____	HH Code _____	HH Code _____	HH Code _____	HH Code _____
	Mother	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2	Yes - 1 No - 2

	Father	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2
	Grandmother	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2
	Grandfather	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2
	Aunty more than 12 years old (Devrani/Jaithani/nanad)	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2
	Uncle more than 12 years old (Tau/chacha)	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2
	Other family members	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2	Yes – 1 No – 2
		Yes – 1			Remarks				
5. Community workers participation		No – 2			Note: Please comment				
ASHA		Yes – 1 No – 2							
Aanganwadi worker		Yes – 1 No – 2							
Aanganwadi helper		Yes – 1 No – 2							

New activity	
6. Was new activity added ?	Yes-1 No- 2
7. If yes, what activity added and why:	
8. Did you find any difficulties to implement any activity?	Yes-1 No- 2
9. If yes, note the difficulties and activity	
10. What strategies have you used to ensure that programmed activities are followed?	
Please comment:	
11. Note here about other information that you consider important:	
12. Group discussion General points	
1) Ask to the group participants who all listened to stories on Tika Vani?	

<p>Number of target household who listened to stories by receiving calls _____</p> <p>Number of target household who listened to stories by dialing TV number _____</p> <p>Please comment _____</p>
<p>2) Ask to all who all dialed TikaVaani number?</p> <p>Number of target households who dialed TikaVaani number _____</p> <p>Other households who dialed TikaVaani number _____</p> <p>Please comment: _____</p>
<p>3) If dialed number then reason for dialing the number?</p> <p>Please write__</p>
<p>4) If not listen to TikaVaani stories or not dialed TV number then reason</p> <p>Please write__</p>
<p>5) If someone listen to TV stories by dialing or receiving call then ask which stories?</p> <p>1- Pneumonia-write number _____</p> <p>2- VHND- write number _____</p> <p>3- FAQ2 – write number _____</p>

Please comment:___
6) If they listened to the stories then what did they learn from the stories? Please write___
7) The information you got over Tika Vani platform, did you hear that information first time, or you were already know about it? Please write___
8) If they already know the information, then ask did they know complete information or partial information? Please write___
9) Ask to the participants, do you believe on information given on TikaVaani? Please write___
10) Ask to the participants, How were the stories? Please write___
11) If they liked the stories then ask reason? Please write___

12) If they did not like the stories then reason
Please write ____
13) Did they face any problem in listening the stories or using Tika Vani number? And ask what problems?
Please write ____
14) If they faced problem then write problem?
Please write ____
13. Pneumonia capsule
1) How many of you have heard about Pneumonia?
Please write ____
2) Ask the participants, If anyone of you had pneumonia in your household or nearby your household if yes then what had happened or how did you know that it was pneumonia?
Please write ____
<i>Please play pneumonia reminder capsule and then start discussion</i>
3) <i>Ask, have you ever listened to played information on Tika Vani?</i>
Please write ____
4) What are the symptoms of Pneumonia?
Please write ____

5) What are the benefits of knowing Pneumonia symptoms?
Please write ___
6) What were the three steps in the story ?
Please write ___
7) How does pneumonia affect the body of children?
Please write ___
8) How does Pneumonia spread?
Please write ___
9) What are the four measures to prevent pneumonia?
Please write ___
10) Ask, after listening story on Pneumonia on TV platform what measures they have taken to fight with Pneumonia? If someone took some measures, then ask everyone to clap their hands so that other can get inspiration.
Please write ___

14. VHND capsule

1) How many of you know about VHND and what do you know?

Please write ____

2) If you go to VHND, Why do you go and how is your experience of VHND?

Please write ____

Please play VHND reminder capsule then start discussion

3) *Did they listened this capsule or story on Tika Vani ?*

Please write ____

4) What information have you got from the story?

Please write ____

5) On which days VHND organized in the villages according to story? On which day VHND organized in your village?

Please write ____

6) What facilities are given in the VHND?

Please write ____
7) What are the benefits of having VHND in the village?
Please write ____
8) Who all present in VHND to provide health facilities to the people?
Please write ____
9) Ask, after listening story on VHND on TV platform, has there been any change in going on vaccination day? what change?
Please write ____

Meeting end time: _____

Annexe 3 Grille d'entretien pour guider les rencontres communautaires

Article 3: Additional file 3 Qualitative Tools

Community

Group Discussion -Instructions

Study Team Roles

Moderator: (name)

Role: Guide the discussion in the group. Be neutral, do not take a position.

Observer : (name)

Role: The observer does not intervene in the discussion. He/She is in charge of the audio recording participants' answers.

Moderator's Introduction

Good afternoon and welcome to our session. Thank you for taking the time to join us to talk about Tika Vaani Intervention.

Village name:

Date:

Number of participants:

* **Women:**

* **Men:**

INTRODUCTION

My name is ----- and He/She is ----- . We are employees of Jagriti Foundation and represent the study team. We would like to know what you like about Tika Vaani, what you do not like, and how Tika Vaani might be improved.

You were invited to join this discussion because Tika Vaani was offered in this village. We are having discussions like this in all the villages that received Tika Vaani.

Please feel free to share your point of view even if it differs from what others have said. There are no wrong answers but rather differing points of view. We learn from all the different points of view, and sometimes criticism can be the most helpful for learning.

We are going to record the discussion and take notes to be sure that we do not miss any information. Rest assured, your answers will be kept fully confidential.

Questions to guide the discussion

1. To improve the health of children, Tika Vaani give information through phones and also through community meetings. Is getting information through phones and community meetings is a good method or any other method is better?
2. is it important for families to get the information to improve the health of their children?
3. Those who used TikaVaani mobile phone service, have they faced any problem?
4. If yes, what was the biggest problem you faced in using Tika Vaani service? (do not provide options)

5. Was demonstration of Tika Vaani number, getting information through mobile phones was clear or understandable?

6. we have information through stories, main messages and direct instruments

6a: was information provided on Tika Vaani was clear and understandable?

6b: which method of information dissemination was liked by you the most (stories, main message and direct instruments)

6c: Which is good method for learning?

6d: Do you like to get information through mobile or through community meetings or both?

7. we called two times every week on Monday and Saturday at around 1PM in the afternoon and around 8PM in the night to give information on mobile

Did you find the timing of calls and getting two times calls in a week OK/RIGHT?

How to improve this?

8. we called the parents of children whose vaccination was due, to remind them about vaccination. to remind about vaccination we called 4 days before vaccination, one day before vaccination and in the morning of vaccination?

8a was reminding families on child vaccination through calls helpful?

8b. how can we improve it more?

9. Information given in the community meeting was helpful?

10. For the people who have not used TikaVaani number or never participated in community meetings?
 - 10a. would you please tell us why you people did not call on TikaVaani number or never used it?
 - 10b. if you have young children, then can we do something so that you start using TikaVaani?
11. We have see in few cases, mothers would like to listen the stories on TikaVaani but they did not have phone, their husbands had phones. We can we get the support of family members so that mother can listen to this?
12. If in future, you want to get information on child's health or would like to get any information on any subject like how to make ORS or symptoms of Penumonia or when to do vaccination, you can call on TikaVaani number and listen the information. So do you think this service is helpful?
13. Is there any other topic on which you want to listen messages/information?
14. What do you expect from us to make TikaVaani better?

Annexe 4 Guide d'entretien pour guider les entrevues avec les travailleurs de la santé

Community workers (ASHA/ AWWs, AWW Helpers)

Interview Guide

Village name _____

Professional information

- Job title
- Number of years in the profession

Introduction: Thank you for agreeing to be interviewed. This consent form was necessary for us to ensure that you understand the purpose of your involvement and that you agree to the conditions of your participation.

I. General Information

1. During this year did you receive training to improve your work skills?
2. How did you feel about it?

*If the person has received training

-From whom have you received training?

-In which subjects you have received training?

(exemple : vaccination, drinking water, hand washing, childhood nutrition, etc.)

-Was the training useful or helped you to improve your work skills?

-Would you like to attend these types of training again in future?

3. Would you like to receive training to improve your current work skills?

*If yes

- what topic do you like receive ?
- how would you like to receive training?

*If No

- why you do not want to receive training?

4. Would you like to receive information free of cost to improve your work skills by mobile phone?

- If NO Why?

II. Tika Vaani Intervention Acceptance

5. Have you heard about the Tika Vaani intervention? from who?

6. What do you know about Tika Vaani Intervention?

7. Do you listen the information capsules on the Tika Vaani platform?

- If yes, with whom do you listen?
- If no, what is the main reason for not listening?

8. Have you made a call to the Tika Vaani platform?

***If No**

- What is the main reason why you have not called TV?
- Is there something that would help you to begin using the platform?
- Go to the question 14.

***If Yes** go to the question 9

II. Experiences and difficulties with the Tika Vaani platform

9. Please tell me in detail about your experiences using the Tika Vaani platform

For example:

How did you find the use of the platform?

Did you have any difficulty communicating or following instructions in the platform?

Did you find the information easy or difficult to understand?

10. Do you trust the information given on TikaVaani? Why?

11. Do you see any difference in the community after TikaVaani intervention?

12. Do you have any recommendations to improve the content?

13. Do you have any recommendations to improve accessibility or ease of use?

14. Which additional topics or information would you like to be included in the Tika Vaani platform?

III. Perceptions and satisfaction the Tika Vaani intervention

15. Is Tika Vaani helping you in your work, for example, by improving your knowledge, skills, or self-confidence?

*If yes

➤ Could you give any examples?

*If no

➤ Why?

16. Did you learn anything new from the Tika Vaani intervention?

17. Would you recommend Tika Vaani platform to members of the community?

➤ Why?

18. Would you recommend the platform to other health workers?

➤ Why?

19. In your opinion, what can be done to improve this intervention?

20. Do you have any additional comments or suggestions?

Thank you for your participation. Your comments and ideas are greatly appreciated and will be very useful to improve Tika Vaani.

Annexe 5 Grille d'entretien pour guider les entrevues et les focus groups avec les intervenants-terrain

Tika Vaani Team

INFORMED CONSENT FORM

Information and Purpose

You are invited to participate in this interview in order to know if the TV intervention was implemented as planned. The researcher is also interested in the factors that moderate the implementation the Tika Vaani intervention and how Tika Vaani might be improved before applying it at a larger scale. The purpose of this meeting is to gain a better understanding about your experiences implementing Tika Vaani, including what you like about Tika Vaani, what you do not like and if you have any recommendations to improve this intervention.

Your Participation

Participation consists of a one-on-one interview and mini discussion groups with TV team, lasting approximately one-hour. You may pass on any question that makes you feel uncomfortable. You also have the right to withdraw from this interview at any time. In the event you choose to withdraw all information you provide (including tapes) will be destroyed and omitted from the final paper

Benefits and Risks

The benefit of your participation is to contribute information to improve Tika Vaani intervention before applying it at a larger scale. There are no risks associated with participating in the study.

Confidentiality

Our discussion will be audio taped to help accurately capture your insights in your own words. The tapes will only be used for the purpose of this study, however, your name will not be recorded on the tape. Your name and identifying information will not be associated with any part of the written report of the research. All of your information and our discussion will be kept confidential. The researcher will not share your individual responses with anyone other than the research supervisor. If you have any questions or concerns about this study, you may contact me at the e-mail address : mc.perez.osorio@umontreal.ca

Signature of participant

I have read the information and consent form. My questions or concerns were resolved. By signing below you agree that you have read and understood the above information, and would be interested in participating in this study.

Signature _____ Date _____

Signature of the person who obtains consent

I explained to the participant the research project and this consent form and I answered the questions he / she asked me.

Signature _____ Date _____

Introduction: Thank you for agreeing to be interviewed. This consent form was necessary for us to ensure that you understand the purpose of your involvement and that you agree to the conditions of your participation.

I. Acceptability and satisfaction concerning the TV intervention

1. When did you join the Tika Vaani Team?
2. What parts of your job do you find most challenging?
3. What do you find most enjoyable?
4. Do you feel satisfied with the resources you had available to you to deliver the intervention (for example, transport, speakers, pamphlets, paint)?

II. Strategies to facilitate implementation

5. Did you receive training on how to use the TV platform?
 - If yes, it was useful for you?
 - Please tell me more about this training
6. Did you receive training on how to contact the community?
 - If yes, it was useful for you?
 - Please tell me more about this training
7. Did you have manuals or guides to plan your work?
 - If yes, it was useful for you?
 - Please describe how you used these manuals or guides
8. Did you face any obstacles to promote the TV platform?
If yes

- Could you please give me some examples?
- Please tell me about the strategies that you used to overcome them

III. Quality of delivery

9. Did you receive feedback from a supervisor during the implementation phase?
 - How did you feel about it?
10. Do you feel that there is anything else that supervisors could have done to help you in your work to implement TikaVaani?
 - Why?

IV. Participant responsiveness

11. What are some of the positive and negative comments you received from participants about the TV platform?
12. In your opinion, was it easier to promote the participation of people to attend demonstrations and meetings in some villages than in others?
 - Why?
13. Can you tell me about your experiences in keeping participants' attention during community demonstrations and meetings?
14. What techniques did you use to keep participants attention during community demonstrations and meetings?

V. Recruitment

15. Do you feel that the intervention was able to reach the target population?
 - Why?
16. Was the recruitment process consistently applied across all villages?
17. In your opinion, which of these activities were **most appreciated** by the community:
 - Community meetings
 - Teaching the Tika Vaani platform
 - Information capsules in Tika Vaani platform

- Others....

18. In your opinion, which of these activities were **least appreciated** by the community:

- Community meetings
- Teaching the Tika Vaani platform
- Information capsules in Tika Vaani platform
- Others....

19. How did people respond to you when you moved about the community? Were people hostile, or welcoming or neutral?

20. Was this reaction similar towards the beginning and the end of the pilot study?

- If there are any differences, please explain why you think there was a difference?

VI. Context

21. In your view, are there any other important factors related to the local context that might have helped or hindered implementation of the Tika Vaani intervention?

(For example, political issues or policies, leadership, economic factors, disease outbreaks?)

22. Are there any other challenges you faced that we haven't discussed? Or anything else you would like to add?

Thank you for your participation. Your comments, suggestions and ideas are highly appreciated and will be very useful to improve Tika Vaani.

Perception concerning the TV intervention (Discussion group)

In order to know if the field workers understand what is the theory that underlies the intervention, a small group discussion will be held to discuss with the field workers the following questions:

1. Could you tell me about the goals of the Tika Vaani intervention?

2. Do you think these goals are achievable?
 - Why?

3. Please describe the activities undertaken to achieve these goals

4. In your opinion, which activities and approaches facilitated the implementation of the Tika Vaani intervention?
 - Why do you think these strategies helped? Why?

5. In your opinion, what could be done to improve this intervention?
Which activities would you like to add or exclude? (For example, should there be more or less: visits to the community, information capsules, different topics, different target groups, etc.)

Note: A participative approach (interview + discussion group) will create consensus on what is required to improve the intervention and initiate the necessary change/modification, clarify the specific theoretical constructs if it is necessary, in addition, to facilitate assessment of implementation fidelity.

Annexe 6 Grille pour évaluer la fidélité de la mise en œuvre de l'intervention

Article 3: Additional file 2: Data collection tool to evaluate intervention fidelity*

Content: Task (planned)	Task (completed)		Coverage	Frequency	Duration	Comment
	YES (1)	NO (0)	(I) Implemented as planned (M) Modified (C) Cancelled	(I) Implemented as planned (M) Modified (C) Cancelled	(I) Implemented as planned (M) Modified (C) Cancelled	Please note if necessary: (indicate the date) a. What difficulties did you face in implementing this activity? b. What strategies did you use to ensure that planned activities were followed?
(Please describe each activity component)						
ADAPTATION: Report if any activity was added during the course of implementation and give reasons						
OBSERVATION: Other pertinent information						

*For each planned activity, assess whether it was: (I) Implemented as planned, (M) Modified or (C) Cancelled

Annexe 7 Liste de control GRAMMS pour le rapport des études utilisant des méthodes mixtes

Article 3: Additional file 1. Good Reporting of A Mixed Methods Study (GRAMMS) checklist*

Guideline	Section: page
Describe the justification for using a mixed methods approach to the research question	Methods Section: Study Framework to assess Implementation fidelity Lines: 241-244
Describe the design in terms of the purpose, priority and sequence of methods	Methods Section: Data integration Lines: 338-341
Describe each method in terms of sampling, data collection and analysis	Methods Section: Data collection Data analysis Lines: 249-335
Describe where integration has occurred, how it has occurred and who has participated in it	Methods Section: Data integration Lines: 338-341; Fig 1
Describe any limitation of one method associated with the present of the other method	Discussion Section: Strengths and Limitations Lines: 701-704
Describe any insights gained from mixing or integrating methods	Discussion Section: Strengths and Limitations Lines: 695-701

*O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy. 2008;13: 92-98.

Annexe 8 Certificats d'éthique de la recherche



Comité d'éthique de la recherche du CHUM
Pavillon R, 900 rue St-Denis, 3^e étage
Montréal (Québec) H2X 0A9

Le 11 janvier 2017

Mme Mira Johri
Chercheuse
CRCHUM
Professeure titulaire
École de santé publique, Université de Montréal
Axe de recherche : santé mondiale
courriel : mira.johri@umontreal.ca

Objet :	16.084 – Approbation FINALE CÉR restreint
	Élever la voix de la communauté pour améliorer les services de vaccination auprès des enfants, en Inde

Madame,

Nous accusons réception, en date du 10 janvier 2017, des précisions et corrections demandées ainsi que des documents suivants en vue de l'approbation finale du projet mentionné en rubrique :

- Formulaire d'information et de consentement anglais (questionnaires), version 1 datée du 30 juillet 2016
- Formulaire d'information et de consentement hindi (questionnaires), version 1 datée du 16 août 2016
- Formulaire d'information et de consentement urdu (questionnaires), version 1 datée du 18 septembre 2016
- Formulaire d'information et de consentement anglais (discussions), version 1 datée du 30 juillet 2016
- Formulaire d'information et de consentement hindi (discussions), version 1 datée du 16 août 2016
- Formulaire d'information et de consentement urdu (discussions), version 1 datée du 18 septembre 2016
- Lettre d'approbation, Institutional Committee for Ethics and Review for Research, IIMR, Jaipur, datée du 07 janvier 2017
- Permission Letter-District Program Officer (DPO) Hardoi SDS, datée du 07 septembre 2016
- Permission Letter- Chief Medical Officer (CMO) Hardoi, datée du 29 septembre 2016
- Permission Letter- District Magistrate (DM) Hardoi, datée du 21 septembre 2016
- formulaire 20 complété

Le tout étant jugé satisfaisant, il nous fait plaisir de vous informer que la présente constitue l'approbation finale de votre projet de recherche, **valide pour un an à compter du 11 janvier 2017**.

Vous devrez compléter le formulaire de renouvellement que nous vous ferons parvenir annuellement. De même, vous devrez soumettre pour approbation préalable, toute demande de modification ou document de suivi requis par le comité d'éthique conformément à ses Statuts et Règlements et ce via Nagano.

Veillez noter que le projet de recherche ne pourra débuter avant que vous n'ayez reçu la lettre d'autorisation de réaliser la recherche dans les murs de l'établissement. De même, lorsque cela s'applique à votre situation, le projet ne peut débuter tant que le contrat n'est pas finalisé et dûment signé.

Le comité suit les règles de constitution et de fonctionnement de l'Énoncé de Politique des trois Conseils (ÉPTC 2) et des Bonnes pratiques cliniques de la CIH.

Pour toute question relative à cette correspondance, veuillez communiquer avec le secrétariat du comité par téléphone ou courriel: ethique.recherche.chum@ssss.gouv.qc.ca – 514 890-8000, poste 14485.

Vous souhaitant la meilleure des chances dans la poursuite de vos travaux, nous vous prions d'accepter, Madame, nos salutations distinguées.



Camille Assemat
Vice-présidente
Comité d'éthique de la recherche du CHUM



Indian Institute of Health Management Research, (IIHMR)
 1, Prabhu Dayal Marg, Sanganer Airport, Jaipur - 302011, India Phone +91-141-3924700 Fax +91-141-3924738; ihmr@iihmr.org

Institutional Committee for Ethics and Review of Research

August 2015-July 2017
 Format for Approval

Dated: 7/1/2017

Institutional Committee for Ethics and Review of Research (ICERR)

Chairperson

Dr V Mohan Katoch,
 Former Secretary,
 Department of Health Research, MoHFW Former DG-ICMR, New Delhi

Members

Dr Ashoo Grover
 ICMR, New Delhi

Mr M M Ranjan
 Senior Advocate, Rajasthan High Court, Jaipur

Dr Ganesh Narayan Saxena
 Retd Professor, SMS Medical College and Hospital, Jaipur

Dr Rajeshwari Gupta
 Retd Professor, SMS Medical College and Hospital, Jaipur

Ms Kamaljit Yadav,
 Subodh Public School, Sanganer, Jaipur

Dr Barun Kanjilal
 Professor, IIHMR, Jaipur

Dr Suresh Joshi
 Professor, IIHMR, Jaipur

Dr Major Vinod K SV
 Associate Professor, IIHMR, Jaipur

Dr Arindam Das, Associate Professor, IIHMR, Jaipur

Mr Azeem ur Rehman, General Manager-Coordination BCT, Jaipur

Dr Saurabh Kumar, Assistant Professor, IIHMR, Jaipur

Mr Rahul Sharma, Assistant Professor, IIHMR, Jaipur

Ex-officio Member
 Dean (Research), IIHMR, Jaipur

Member Secretary
Dr Nutan P. Jain
 Professor, IIHMR, Jaipur

	Serial No. of Institutional Committee for Ethics and Review of Health Management Research office	Decem-ber 2016	1
1	Project Title	Dialled in: Tapping community voice to improve child immunization services in India (Version 3)	
2	Name of Faculty In-charge/ Project coordinator	Dr Alok Mathur	
3	Date of Submission to the Committee	2	3 1 1 2 0 1 6
4	Date of Submission to the Agency		
5	Review Category of the Proposal		
	Expedited Review	✓	
	Full review		
6	Date of Review	3	1 1 2 2 0 1 6
7	Reviewer's Name	Dr V M Katoch, Dr Suresh Joshi, Dr Vinod K SV, Dr D K Man- gal, Mr Umesh Sarswat, Mr Azeem ur Rehman, Dr Nutan P Jain	
8	Reviewer's Feedback (Please tick the following, if you are satisfied)		
8.1	Respect for person	✓	
8.2	Fair subject selection	✓	
8.3	Informed consent	✓	
8.4	Maintaining privacy	✓	
8.5	Maintaining confidentiality	✓	
9	Final Comment		
	Approved without suggestions		
	Approved with suggestions	✓	
	Sent back for revision and re-submission		
	Not approved		
10	Signature of the Member Secretary		

Annexe 9 Note de synthèse des résultats de la révision systématique

Tool to facilitate assessment of intervention fidelity in Public health interventions in low- and middle-income countries

Myriam Cielo Pérez ^{a,b}; Nanor Minoyan ^{a,b}; Valéry Ridde ^{c,d}; Marie-Pierre Sylvestre, ^{a,b}; Mira Johri, ^{a,e}

ABSTRACT

Cluster randomised trials (CRTs) of complex public health interventions may be particularly at risk of experiencing protocol deviations and non-adherence. These may compromise the validity of their findings. Evaluation of intervention fidelity (IF) the CRTs of public health interventions in LMICs is required. However, we have found that Fidelity assessment is not currently a systematic practice in CRTs of public health interventions carried out in LMICs.

Key messages

- Fidelity assessment is not a systematic practice in CRTs of public health interventions carried out in LMICs.
- Guidelines for CRTs should promote fidelity assessment to improve the transparency of reports and to help draw firm conclusions about the outcomes of interventions.
- Researchers conducting CRTs of public health interventions in LMICs require a practical tool to facilitate assessment of intervention fidelity.

INTRODUCTION :

Cluster randomised trials (CRTs) have become a key instrument to evaluate public health interventions, particularly in low- and middle income countries (LMICs).

CRTs may be particularly at risk of experiencing protocol deviations and non-adherence, which may compromise the validity of their findings.

Fidelity assessment is an aspect of process evaluation that aims to understand and measure to what extent the interventions are implemented as intended¹.

Current reporting guidelines for CRTs do not offer specific guidance on how to perform implementation fidelity assessment.

Fidelity assessment may be particularly important when assessing CRTs of public health interventions, as they tend to be complex and constituted by multiple components allowing greater potential for variation during implementation²

Dimension for Intervention Fidelity assessment*

	Content
	Active ingredients of the intervention.
	Coverage
	Degree to which all persons who met study inclusion criteria received the intervention
Fidelity	Frequency
Components	Whether the intervention was delivered with the regularity or frequency planned by its designers
	Duration
	Whether the intervention was delivered with the duration planned by its designers

*Carroll et al. (2007) refined by Hasson³ (2010)

APPROACHE

- We undertook a methodologically-oriented systematic review of current practices related to assessment of intervention fidelity within CRTs of public health interventions in LMICs.
- CRTs of public health interventions in LMICs that planned or reported fidelity assessment in either the trial protocol or the main trial report were included.
- We considered that studies evaluated fidelity if they either proposed methods to assess, or reported results related to, the evaluation of at least one of the four key fidelity components: (1) content, (2) coverage, (3) frequency, and (4) duration.
- The MEDLINE/PubMed, CINAHL, and EMBASE databases were queried from January 2012 to May 2018.
- CRTs reporting a registration number in the abstract were included.

RESULTS

• Among 90 CRTs identified, 25 (28%) did not plan or report assessing IF; the remaining 65 studies (72%) assessed at least one key IF dimension.

• IF assessment was planned in 40% (36/90) of trial protocols and reported in 71.1% (64/90) of trial reports.

• The proportion of overall agreement between the trial protocol and trial report concerning occurrence of IF assessment was 66.7% (60/90). It was computed as $(a + d)/N$ -See Table 1.

Table 1 Intervention fidelity assessment among CRTs of public health interventions in LMICs (n=90)‡

		Trial Report		Total
		Positive	Negative	
Trial Protocol	Positive	35 (a)	1 (c)	36 a + c
	Negative	29 (b)	25 (d)	54 b + d
Total		64 a + b	26 c + d	90 N (a + b + c + d)

‡(a) number of studies assessing IF in both in the trial protocol and the trial report

(b) number of studies assessing IF in the trial report but not in the trial protocol

(c) number of studies assessing IF in the trial protocol but not in the trial report.

(d) number of studies that did not assess IF in either the trial protocol or the trial report.

N represents the set of recent CRTs of public health interventions in LMICs that have registered the study protocol in a publicly available trial registry. For each CRT in N, we determine whether IF was assessed in the registered (or published) protocol or in the trial report.

CONCLUSION

- Public health initiatives in LMIC require high-quality evaluations to determine whether programs work and how to improve them.
- There is considerable heterogeneity in the practice of fidelity assessment in CRTs of public health interventions carried out in LMICs. Failure to assess intervention fidelity may weaken the ability of researchers to gauge internal and external validity of studies.

RECOMMENDATIONS

To enhance the quality of future CRTs:

- CRT reporting guidelines should encourage fidelity assessment.
- Scientific journals should encourage researchers to conduct and report IF assessment in CRTs.
- The authors could be encouraged to plan fidelity assessment during the design of the study and reporting it in their trial register or in an online resource.
- Development of a practical fidelity assessment tool is required to guide researchers in how to conduct IF assessment.

To develop a practical tool, we will conduct an exploratory qualitative study in collaboration with researchers using this design to better understand the issues they have faced in planning and/ or reporting fidelity assessment, or the reasons they did not perform/report it.

SOURCES CONSULTED OR RECOMMENDED:

For a detailed description of the methods, see the published protocol available at: 10.1186/s13643-0160351-0.

- Perez MC, Minoyan N, Ridde V, Sylvestre MP, Johri M. Comparison of registered and published intervention fidelity assessment in cluster randomised trials of public health interventions in lowand middle-income countries: systematic review protocol. Syst Rev. 2016;5 (1):177.

References

1. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation science* : IS. 2007;2 (1):40.
2. Medical Research, Council. *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council; 2008.
3. Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci*. 2010;5 (1):67.

Author affiliations

^a Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Québec, Canada.

^b Département de médecine sociale et préventive, École de santé publique (ESPUM), Université de Montréal, Québec, Canada.

^c Institut de Recherche en Santé Publique Université de Montréal (IRSPUM), Montréal, Québec, Canada.

^d Institut de recherche pour le développement (IRD), Marseille, France.

^e Département de gestion, d'évaluation, et de politique de santé, École de santé publique, Université de Montréal, Québec, Canada.

For more information, please contact

Myriam Cielo Pérez: mc.perez.osorio@umontreal.ca

Annexe 10 Synthèse du panorama des chercheurs concernant l'évaluation de la fidélité des études incluses dans la révision systématique

Après avoir invité les auteurs principaux de chaque étude identifiée dans notre revue systématique, 30% des chercheurs (27/90) ont répondu à notre enquête, dont 78% (21/27) ont répondu aux questions, et, parmi eux, 76% (16 /21) des participants ont plus de 10 ans d'expérience dans la réalisation des essais randomisés en grappes, 22% (6/27) ont indiqué qu'ils ne souhaitaient pas participer

Question	Interprétation	Exemple des réponses des chercheurs participants
1. Connaissez-vous le concept d'évaluation de la fidélité de la mise en œuvre de l'intervention?	100 % (21/21) des participants sont familiers avec le concept. Parmi eux, 47% (10/21) ne sont pas très sûrs de leur compréhension correcte ou si la façon de l'évaluer ou de le signaler est adéquate.	<i>" Yes, but I guess my understanding is not comprehensive enough "</i> <i>"I am familiar with the concept of interventions being delivered or implemented as intended but was not familiar with the specific term 'intervention fidelity'"</i> <i>"Yes, empirically. Not sure I recall formal training on the topic."</i>
2. Considérez-vous l'évaluation de la fidélité de la mise en œuvre de l'intervention comme nécessaire ou non dans un essai randomisé en grappes?	100 % (21/21) des participants sont d'accord avec la nécessité d'évaluer la fidélité de la mise en œuvre de l'intervention dans les essais randomisés en grappes.	<i>"Yes, I think it is particularly important in this context given treatment/intervention implementation might be susceptible to variation by cluster"</i> <i>"Yes, absolutely, at least the most basic indicators will need to be mentioned as it is important to see the result in relation to what was delivered"</i> <i>"It is surely very important, as it may directly affect the outcome."</i> <i>"Yes, it would be useful, though the method to assess the fidelity will vary and depend on the nature of interventions, the context and resources. Useful document/tools will be useful. The usefulness will depend on the quality of the tool"</i>

3. Si vous avez fait une évaluation de la fidélité de la mise en œuvre de l'intervention dans votre étude: quels succès ou défis avez-vous rencontrés?

Parmi eux, 14% (3/21) considèrent que l'évaluation de la fidélité de la mise en œuvre de l'intervention ne devrait pas être un critère obligatoire des guides des essais randomisés en grappes, ou qu'il n'est pas obligatoire de générer des guides supplémentaires pour tenir compte de cette pratique.

De façon générale, quelques chercheurs trouvent que le fait de surveiller de proche l'intervention a aidé à obtenir un haut niveau de la fidélité de la mise en œuvre de l'intervention, et n'ont pas eu de difficulté pour l'évaluer.

“Do NOT try to generate a checklist. Encourage awareness of the issue and the need for a NARRATIVE description of fidelity”

“It necessary to be concerned about it and necessary to report on it. The last thing it needs is another mindless checklist assessment”

“I'd say useful, stopping short of saying necessary or mandated in CONSORT guidelines kind of way.”

« No challenges, « we had no issues of fidelity - program ran virtually exactly as we designed it »

Des raisons évoquées

« very close and intense monitoring of the small team implementing »

« we monitored data quality on the intervention implementation »

« There were a series of analyses and research studies as a part of the programme evaluation framework »

« Monthly meetings of the steering committee to assess overall study implementation and address any issues. »

« Attendance of intervention personal with booster training sessions ».

Des enjeux soulevés

4. Si vous n'avez pas évalué la fidélité de la mise en œuvre de l'intervention dans votre étude, quelles étaient les principales

D'autres mentionnent plusieurs aspects, dont la difficulté de la mise en place de l'intervention telle que prévue, ainsi que des difficultés dans l'évaluation de la fidélité de la mise en œuvre de l'intervention au niveau conceptuel et sur la façon de la mesurer.

100 % (21/21) des participants ont manifesté avoir évalué la fidélité de la mise en œuvre des interventions. Cependant, plusieurs ne savaient pas s'ils l'avaient évaluée de façon appropriée.

« we faced to keep the initial intervention design in these trials: finding professionals (nurses and teachers) who had a good basic level of training, logical reasoning and were willing to live far from urban centres; keeping trained staff; migration of people; heterogeneity in children's initial knowledge and level; delays when implementing part of the programme ».

« Understand very well the concept about fidelity assessment, how I need to do that in our study and explain why this issue is important »

« Estimating coverage has been challenging. Ranging from needing large population-based surveys to see household coverage, to venue-based surveys to see if clientele coverage... Sometimes there is a disconnect between who in a community the intervention is meant for and who received it ».

« Assessment of fidelity is expensive, resource intensive and time consuming ».

« A particular challenge is to give a qualitative researcher enough understanding of what is useful to prevent generation of a lot of qualitative data that don't allow us to get an impression of delivery quality ».

« I believe we have attempted to measure fidelity in all cluster randomized trials I can recall. Certain aspects may not have been effectively measured, particularly content ».

« We did assess fidelity. In fact, both trials had a quite strict and intense monitoring system. Yet, our strategy to empower the intervention managers to do most of the monitoring might be questionable ».

raisons de ne pas le faire?

5. Si nous devions développer un guide méthodologique simple pour aider les chercheurs à développer un plan pour évaluer la fidélité de la mise en œuvre de l'intervention, serait-il utile et si oui, sous quel format?

85% (18/21) des participants ont trouvé cette idée très utile, tandis que 5% (1/21) ne l'ont pas trouvée utile, et 10% (2/21) n'ont pas répondu à la question.

*** Serait-ce utile pour vous?**

“Yes, it would be useful, though the method to assess the fidelity will vary and depend on the nature of interventions, the context and resources. Useful document/tools will be useful. The usefulness will depend on the quality of the tool”

“Yes, absolutely! Something that is easily accessible and straightforward, with appropriate source information, would be great...”

« No this would be dreadful »

Quel format serait le plus utile?

« It would be helpful to have a reference document, but also clear and to-the-point slides that could be shared with study partners (particularly those that are outside of the U.S. and potentially nonnative English speakers) ».

« A simple checklist of interventions with one indicator or proxy indicator for fidelity/adherence...

and a guide that allows you to define well how to use it in a study »

« A standard format that is adaptable to local context will be useful ».

« a peer-reviewed journal publication would be most useful. It helps to cite in grant application as it has high credibility ».

6. Avez-vous des recommandations spécifiques à partager avec nous qui pourraient nous aider à développer un guide pour l'évaluation de la fidélité de la mise en œuvre de l'intervention?

Des recommandations visant des aspects pratiques et méthodologiques pour cette pratique évaluative ont été fournies.

« It can also help in additional analysis of CRT data. It will be useful in quasi- or observational evaluation designs. However, it will not help in design, conduct or analysis of CRT because we do not want to deviate from the gold standard where causality can be established »

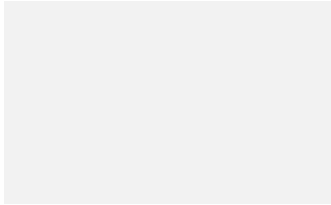
« Often the intervention has many different factors and is multidisciplinary. It will be good to have a tool to accommodate the wide range ».

« the guide should help develop a score or index which can be consistently measured and represented as a (continuous or ordered) number. The methods can be mixed/qualitative but without a numeric indicator for IF, it will be tough to use in most analysis of eRCT data »

« A guide to assess fidelity could include the following: -introduction on the importance of fidelity assessment in different study designs, - methods that can be used, with advantages and disadvantages of each method, -Sample size estimation for fidelity assessment -Checklist for each method that is adaptable and can be converted into a questionnaire -How should the findings of the assessment be used. Can these be used to give feedback to the intervention delivery team to strengthen the intervention »

« Hope that guidelines will be made user-friendly. If the guide is complicated, nobody will want to use it »

« The IF guideline should define fidelity from “practical achievability” perspective. For example, I would not agree that making entire village open defecation free is aspired fidelity because realistically 80-90% open defecation can be reduced by the most successful intervention. So the top13 ladder of fidelity should be “demonstrated” proven in pilot/ideal settings at least. The guidelines



CRT: cluster randomised trial
IF: intervention Fidelity

can include IF assessment tools for different sectors – education, agriculture, Water-Sanitation, immunization, maternal health, etc. Use gradually technology and use exploratory measures to record this is very important (like interview, discussion groups, observation)”.