

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

Exactitude du rappel de la qualité de vie préopératoire chez les femmes opérées
pour une incontinence urinaire à l'effort

Par

Annick Larochelle, M.D.

Programme de sciences biomédicales

Faculté de médecine

Mémoire présenté à la faculté des études supérieures
En vue de l'obtention du grade de maître ès sciences (M.Sc.)
En sciences biomédicales
Option recherche clinique

Juillet 2008



© Annick Larochelle, 2008

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

Exactitude du rappel de la qualité de vie préopératoire chez les femmes opérées
pour une incontinence urinaire à l'effort

Présenté par:
Annick Larochelle, M.D.

A été évalué par un jury composé des personnes suivantes :

Chantal Dumoulin
Président-rapporteur

William D. Fraser
Directeur de recherche

Stéphane Ouellet
Codirecteur

Patricia Monnier
Membre du jury

Résumé

Objectif

Évaluer la validité et la fiabilité d'une échelle de qualité de vie (QOL) administrée rétrospectivement plutôt que prospectivement auprès de femmes traitées pour une incontinence urinaire à l'effort, par une procédure de TVT (*Tension-free Vaginal Tape*) ou de TOT (*Transobturator Tape*).

Méthodes

Cohorte prospective incluant 152 patientes recrutées au Centre hospitalier universitaire de Montréal. Les échelles IIQ-30 et SF-12 ont été auto-administrées aux sujets avant la chirurgie. Les mêmes questionnaires ont été postés trois mois après la chirurgie afin d'évaluer le rappel de la QOL préopératoire. La validité et la fiabilité d'une mesure rétrospective de la QOL ont été testées par des analyses de corrélation de Pearson (r) et de corrélation intraclass (ICC).

Résultats

Cent treize patientes ont retournés leurs questionnaires postopératoires, pour un taux de réponse de 82,2%. La médiane du temps écoulé entre l'évaluation préopératoire et l'évaluation du rappel en postopératoire était de 14,29 semaines (12,42 – 46,00). Le rappel mesuré selon le IIQ-30 a démontré une excellente validité ($r = 0,64$) et un niveau de fiabilité modéré (ICC = 0,62). Le rappel mesuré avec le SF-12 a démontré une validité modérée ($r = 0,46$ (*physical component summary* ou PCS) et 0,42 (*mental component summary* ou MCS)) et un niveau de fiabilité faible (ICC = 0,44 (PCS) et 0,49 (MCS)).

Conclusion

Les femmes opérées pour une incontinence urinaire à l'effort donnent un rappel adéquat de leur QOL préopératoire, lorsqu'elles sont évaluées par le IIQ-30, trois mois après leur chirurgie. Le SF-12 ne peut être utilisé rétrospectivement de façon fiable.

Mots-clés : rappel, rétrospectif, qualité de vie, incontinence urinaire de stress, TVT, TOT.

Abstract

Objective

To assess the validity and reliability of a retrospective quality of life (QOL) assessment compared to a prospective assessment among women treated for stress urinary incontinence (SUI) with a tension-free vaginal tape (TVT) or a trans-obturator tape (TOT) procedure.

Methods

One hundred and fifty-two patients were enrolled in a prospective cohort study at the Centre hospitalier universitaire de Montréal. The Incontinence Impact Questionnaire (IIQ-30) and the Short-Form Health Survey (SF-12) were self-administered pre-operatively to consenting women scheduled for a TVT or a TOT procedure. At 3 months post-op, the IIQ-30 and SF-12 surveys were mailed to these patients to reassess their pre-operative QOL status. Pearson's correlation coefficient (r) and the intraclass correlation coefficient (ICC) were used to test the validity and reliability of the recalled IIQ-30 and SF-12 scores versus baseline preoperative scores.

Results

One hundred and thirteen patients completed and mailed their post-operative questionnaires addressing the recalled pre-operative QOL, for a response rate of 82.2%. Median time between pre-operative assessment and post-operative recall assessment was 14.29 weeks (range 12.42-46.00). Recall validity was excellent for the IIQ-30 ($r=0.64$), and moderate for the SF-12 ($r= 0.46$ (physical component summary or PCS) and 0.42 (mental component summary or MCS)). Recall reliability was moderate with the IIQ-30 (ICC= 0.62), and poor with the SF-12 (ICC= 0.44 (PCS) and 0.49 (MCS)).

Conclusion

The IIQ-30 can be reliably used in a retrospective manner among women who have undergone surgery for SUI 3 months earlier. The SF-12 cannot be reliably used retrospectively.

Key words: retrospective, recall, quality of life, stress urinary incontinence, TVT, TOT.

Table des matières

Résumé	iii
Abstract	iv
Rôles du chercheur principal et des cochercheurs	x
Introduction.....	11
Chapitre I : Recension de la littérature	13
1 L'incontinence urinaire	13
1.1 Définition.....	13
1.2 Le traitement de l'incontinence urinaire à l'effort.....	13
2 Le concept de la qualité de vie	14
2.1 La qualité de vie reliée à la santé.....	15
3 Les questionnaires de qualité de vie.....	16
3.1 Les questionnaires génériques	17
3.2 Les questionnaires spécifiques	18
4 Construction et validation des questionnaires.....	18
4.1 Construction d'un questionnaire.....	18
4.2 Validation d'un questionnaire.....	19
5 Les questionnaires de qualité de vie spécifiques à l'incontinence urinaire.....	20
6 Utilisation prospective ou rétrospective des questionnaires de qualité de vie ..	22
7 Littérature sur le rappel	23
7.1 Identification des sources de biais potentiels.....	28
8 Pertinence de la recherche.....	31
Chapitre II : Méthodologie	32
1 Objectifs	32
2 Hypothèse.....	32
3 Devis.....	33
4 Population à l'étude.....	33
5 Recrutement	34
6 Collecte des données et suivi	35
7 Analyses statistiques	36
Chapitre III : L'Article	41
Abstract.....	42
Introduction.....	43
Patients & methods.....	44
Results	48
Discussion.....	50
Conclusion	54
Chapitre IV : Discussion	56
Conclusion.....	61
Bibliographie.....	67

Annexe I : Questionnaire sur les données socio-démographiques.....	xi
Annexe II IIQ-30	xiii
Annexe III SF-12.....	xvi
Annexe IV Scatterplots of recalled QOL scores compared to baseline preoperative QOL scores for the IIQ subscales: physical (a), traveling (b), emotional (c), social (d).	xix
Annexe V Patient's recalled scores of preoperative quality of life compared to prospective baseline scores of preoperative quality of life for the SF-12 physical (PCS) (2.1) and Mental (MCS) (2.2) scales.	xx
Annexe VI Univariate linear regression where the difference between Baseline score and recalled score is the dependant variable and patient characteristics are the independent variables for SF-12.....	xxi
Annexe VII Multivariate analysis of age, employment status, education category and income category on recalled score using IIQ-30.	xxii
Annexe VIII Sociodemographic characteristics and baseline QOL scores of incontinent and continent patients.....	xxiii
Curriculum vitae.....	xxiv

Liste des tableaux

Table I : Sociodemographic characteristics and baseline QOL scores between respondents and nonrespondents to the recall survey	62
Table II : Correlation between recalled QOL scores and baseline preoperative QOL scores	63
Table III : Reliability of using recalled QOL scores compared to baseline preoperative QOL scores	65
Table IV : Univariate linear regression where the difference between Baseline score and recalled score is the dependant variable and patient characteristics are the independent variables for the IIQ-30.....	66

Liste des figures

Figure 1 : Scatterplots of patient's recalled QOL scores compared to
baseline preoperative QOL scores for the IIQ total..... 64

Liste des sigles et abréviations

ACL-QOL	Quality of Life Outcome Measure
AVC	Accident vasculaire cérébral
CHUM	Centre hospitalier universitaire de Montréal
CI	Confidence interval
HRQ	Hip Rating Questionnaire
HRQOL	Health-related quality of life
ICC	Intraclass correlation coefficient
IIEF	International Index of Erectile Function
IIQ-30	Incontinence Impact Questionnaire
IKDC	International Knee Documentation Committee
I-QOL	Incontinence quality of life instrument
IRC	Insuffisance rénale chronique
KOOS	The Knee Injury and Osteoarthritis Outcome Score
MCS	Mental component summary
MID ou MICD	Minimal clinically important difference
MPOC	Maladie pulmonaire obstructive chronique
OMS	Organisation mondiale de la santé
PCS	Physical component summary
QALY	Quality-Adjusted-Life-Year
QOL	Qualité de vie ou Quality of life
r	Coefficient de corrélation de Pearson
SAS	Statistical Analysis System
SEM	Standard Error of Measurement
SEP	Sclérose en plaques
SF-36	36-Item Short-Form Health Survey
SF-12	12-Item Short-Form Health Survey
TOT	Transobturator tape
TVT	Tension-free vaginal tape
UCLA PCI-SF	University California Los Angeles Prostate Cancer Index Short Form
WOMET	Western Ontario Meniscal Evaluation Tool

Rôles du chercheur principal et des co-chercheurs

Les chercheurs du projet sont les docteurs Annick Larochelle, William D. Fraser, John Podoba et Stéphane Ouellet.

Le docteur **Annick Larochelle** est le chercheur principal du projet de recherche. Elle est une obstétricienne gynécologue qui s'intéresse aux applications de l'épidémiologie clinique dans le domaine de l'urogynécologie. Elle a un intérêt particulier en incontinence urinaire et sur l'impact de cette condition sur la qualité de vie des femmes. Le Dr Larochelle a conçu le projet et a préparé les documents nécessaires pour la soumission du projet au Comité d'éthique du centre de recherche du pavillon St-Luc du CHUM. Elle a coordonné les responsabilités et les activités de la secrétaire et des infirmières de recherche. Elle s'est aussi assurée de la qualité de la collecte des données et du suivi des sujets auprès de la secrétaire et des infirmières de recherche. Elle a supervisé la saisie des données par les employés du centre de recherche de l'Hôpital Sainte-Justine et a veillé au bon usage du financement reçu. Avec l'aide du Dr Podoba, elle a procédé aux analyses statistiques. Elle est la rédactrice du rapport de recherche.

Le docteur **William D. Fraser** est le directeur de maîtrise. Il est un obstétricien gynécologue ayant une formation surspécialisée en médecine foeto-maternelle et en épidémiologie. Le Dr Fraser s'est assuré de fournir les ressources financières et le personnel qualifié pour la bonne réalisation du projet. Il s'est impliqué dans la rédaction du rapport de recherche.

Le docteur **John E. Podoba** est un épidémiologiste du centre de recherche de l'Hôpital Sainte-Justine. Il a fourni son expertise dans l'élaboration et l'application des analyses statistiques.

Le docteur **Stéphane Ouellet** est le co-directeur de maîtrise. Il est un obstétricien gynécologue du pavillon St-Luc du CHUM qui possède une expertise en incontinence urinaire et en chirurgie pelvienne reconstructive. Le Dr Ouellet a contribué au bon déroulement du projet en veillant sur les activités de la secrétaire et des infirmières de recherche. Il fût disponible sur place pour répondre aux questions des patientes participant au projet.

Introduction

L'incontinence urinaire à l'effort ou l'incontinence urinaire de stress est un problème de santé prévalent. Cette condition affecte 15 à 35% de la population féminine adulte (1). Plus de 200 techniques chirurgicales différentes visant à corriger l'incontinence urinaire de stress ont été décrites (2). Les techniques utilisant les bandelettes de polypropylène par voie vaginale, dont le TVT (*Tension-Free Vaginal Tape*) et le TOT (*Transobturator Tape*), sont les plus récentes et parmi les plus efficaces (3, 4, 5). Nous ne connaissons pas encore très bien l'impact de ces chirurgies sur la qualité de vie des patientes.

Depuis quelques années, le concept de l'efficacité d'un traitement a évolué. On ne cherche plus seulement à connaître les impacts physiologiques et biologiques des thérapies mais également les bénéfices psychosociaux par le biais de l'évaluation de l'impact sur la qualité de vie (QOL) des patients. Pour mesurer l'effet des traitements sur la qualité de vie des malades, plusieurs auteurs ont développé des questionnaires de qualité de vie. Le questionnaire SF-36 (6) et sa version abrégée, le questionnaire SF-12 ont été développés pour évaluer la qualité de vie globale des patients (7). Le SF-12, composé de 12 questions, a été traduit en français et validé auprès de la population québécoise francophone (8). Pour évaluer le problème particulier de l'incontinence urinaire, un questionnaire spécifique visant à mesurer la qualité de vie des femmes atteintes par ce problème a été développé. Le questionnaire IIQ-30 (*Incontinence Impact Questionnaire*) composé de 30 questions a été traduit et validé auprès de la population québécoise francophone (9,10).

Dans le contexte de l'évaluation de l'effet d'un traitement sur la QOL, un questionnaire est habituellement administré aux patients avant le traitement et ensuite après le traitement à un moment prédéterminé. C'est l'approche prospective. Un score de QOL est calculé pour l'évaluation prétraitement et un autre pour l'évaluation post-traitement. La différence de ces deux scores nous permet d'estimer l'évolution de la qualité de vie des sujets traités pendant le période de temps comprise entre les deux moments d'administration du questionnaire. Évidemment, l'approche prospective nécessite deux contacts avec la patiente à des moments différents. Elle exige des ressources financières et des ressources en personnel de recherche. Une façon différente de procéder serait d'administrer les mêmes questionnaires mais seulement après la thérapie, à un moment prédéterminé. Cette

fois, la patiente répond aux questions concernant sa qualité de vie actuelle dans la période post-traitement, et ensuite elle doit répondre une deuxième fois au même questionnaire mais en se rappelant sa condition prétraitement. De cette façon nous obtenons aussi deux scores de qualité de vie, soit le score de QOL post-traitement et le rappel du score de QOL prétraitement basé sur le souvenir du sujet. Encore une fois, la différence entre les deux scores permettrait de déterminer s'il y a eu une amélioration de la qualité de vie depuis l'intervention. Il s'agit de l'approche dite « rétrospective ». Elle aurait l'avantage d'être pratique, rapide et peu coûteuse. Cependant, cette approche ne pourrait être valide que si les mesures recueillies en période post-traitement cherchant à évaluer la période prétraitement reflètent fidèlement la période prétraitement.

L'objectif de cette étude est de valider l'utilisation rétrospective d'une échelle de QOL générique (SF-12) et spécifique (IIQ-30), dans une population recevant un traitement pour l'incontinence urinaire à l'effort. Pour ce faire, il faut évaluer en période postopératoire l'exactitude du rappel de la condition préopératoire. Si cette approche est démontrée comme étant valide, il sera justifié d'administrer les instruments SF-12 et IIQ-30 d'une façon rétrospective. Cela permettrait aux cliniciens de procéder plus facilement à des études cliniques voulant évaluer l'impact des thérapies sur la QOL en réduisant les investissements et les ressources diverses nécessaires.

Chapitre I : Recension de la littérature

1 L'incontinence urinaire

1.1 Définition

L'incontinence urinaire est définie comme une condition où survient une fuite d'urine involontaire objectivement démontrable, et qui devient un problème social ou hygiénique (11). L'incontinence urinaire affecte 15 à 35% des femmes adultes américaines par ailleurs normales et menant une vie active (1). En 1991 en Grande Bretagne, on estime à 9,3% la proportion des femmes souffrant d'incontinence urinaire (12). Au Canada, selon un sondage national réalisé en 1997, 12% des femmes souffraient d'incontinence urinaire (13).

Par ailleurs, il existe différents types d'incontinence urinaire dont l'incontinence à l'effort, l'incontinence d'urgence, l'incontinence par trop plein, l'incontinence fonctionnelle, l'incontinence totale et l'incontinence mixte (combinaison d'une incontinence urinaire à l'effort et d'une incontinence urinaire d'urgence (14)). Près de la moitié des femmes souffrant d'incontinence urinaire présentent des symptômes d'incontinence urinaire à l'effort (49%) (5). L'incontinence urinaire à l'effort, aussi appelée incontinence urinaire de stress, survient lors d'une augmentation de la pression intra-abdominale, par exemple lors d'un effort physique tel que la toux, l'éternuement, le rire ou la marche.

Les coûts financiers occasionnés par ce problème de santé sont élevés. En se basant sur une extrapolation de données américaines, les coûts directs et indirects de l'incontinence urinaire au Canada sont d'environ 2,6 milliards de dollars par an (13, 15).

1.2 Le traitement de l'incontinence urinaire à l'effort

L'approche thérapeutique comporte des options chirurgicales et non-chirurgicales. Une patiente peut d'abord choisir une approche non-chirurgicale telle que la médication, la physiothérapie ou le pessaire d'incontinence (16). Malheureusement, ces alternatives thérapeutiques n'offrent pas les mêmes chances de succès que la chirurgie, surtout face à une incontinence urinaire à l'effort de niveau modéré à sévère (16, 17, 18).

Plus de 200 techniques chirurgicales ont été décrites pour le traitement de l'incontinence urinaire à l'effort (2). Parmi les techniques les plus récentes existent le TVT (*Tension-free*

Vaginal Tape) et le TOT (*Transobturator Tape*) (3, 4, 5). Elles consistent en l'apposition d'une bandelette de tissu synthétique, le polypropylène, par voie vaginale au niveau sous-urétral. Lors de la procédure du TVT, on procède au passage de la bandelette à l'aide de tunneliseurs, à travers l'espace rétropubien. Une cystoscopie doit être effectuée afin d'éliminer tout traumatisme vésical secondaire au passage des tunneliseurs. Lors de la procédure du TOT, la bandelette est passée, toujours à l'aide de tunneliseurs, au travers du trou obturateur. La cystoscopie n'est pas requise. La technique du TVT a été décrite en 1996 par Ulmsten (3) et celle du TOT a été décrite par Delorme en 2001 (4). Il s'agit de chirurgies d'un jour. Elles ont l'avantage d'être moins invasives et de présenter moins de complications que le traitement considéré comme l'étalon-or, soit l'urétroréflexion rétropubienne qui nécessite une laparotomie. L'urétroréflexion rétropubienne de type Burch est probablement l'intervention urogynécologique la mieux évaluée (18). Selon une revue Cochrane, elle permet de corriger l'incontinence urinaire de stress avec succès chez plus de 70% des patientes et ce jusqu'à 5 ans après la procédure (17, 18). Alcalay M et al, ont démontré que ce succès peut se prolonger jusqu'à 10 ou même 20 ans après l'intervention de type Burch (19). Elle est donc considérée comme l'intervention la plus efficace à long terme. Les taux d'efficacité des procédures de TVT et TOT sont aujourd'hui estimés autour de 85-90%, à 5 ans (5, 20, 21). Nilsson a démontré une efficacité de 81 % à sept ans avec la procédure du TVT (22). La procédure du TOT, plus récente que celle du TVT, démontre jusqu'à maintenant une efficacité de l'ordre de 80% à deux ans (20). Face à la nouveauté des procédures de TVT et de TOT, il est pertinent de poursuivre la recherche clinique dans ce domaine afin de connaître leur efficacité à long terme et sous diverses dimensions dont celle de l'amélioration de la qualité de vie suite à la procédure.

2 Le concept de la qualité de vie

La qualité de vie est un concept vaste et difficilement saisissable comportant plusieurs dimensions. Les définitions de la qualité de vie sont diverses et variées. La définition consensuelle reste un défi de taille.

Malgré le désaccord des auteurs sur le sujet, plusieurs d'entre eux apportent des éléments de définition qui se recoupent considérablement. D. Felce et J. Perry (24) indiquent cinq dimensions qui prédominent dans les différents modèles de qualité de vie globale décrits dans la littérature: le bien-être physique, le bien-être matériel, le bien-être social, le

développement et l'activité, et le bien-être émotionnel. Felce et Perry élaborent leur propre modèle en se concentrant davantage sur trois éléments soit les conditions de vie objectives, le sentiment subjectif de bien-être et les valeurs personnelles, et les aspirations. L'ensemble de ces trois éléments en interaction les uns avec les autres et avec des influences externes reflète leur conceptualisation de la qualité de vie globale. La définition de la qualité de vie offerte par ces auteurs apparaît parmi les plus justes et complètes : « Un état de bien-être général qui comprend des indicateurs objectifs et des évaluations subjectives de l'ensemble des états de bien-être physique, matériel, social et émotionnel pris dans un contexte d'activités personnelles dirigées et de développement personnel, sous le regard d'un ensemble de valeurs personnelles. » (24).

2.1 La qualité de vie reliée à la santé

L'intérêt pour l'évaluation de la qualité de vie en médecine s'est développée au cours des dernières décennies parallèlement avec l'évolution de la médecine où l'on a vu naître des traitements de plus en plus innovateurs qui ont permis d'allonger l'espérance de vie mais aussi d'augmenter la proportion de la population au prise avec des conditions médicales chroniques. Autrefois, le rôle traditionnel de la médecine était d'assurer la survie des patients. Maintenant, son rôle n'est plus seulement de prolonger la vie mais aussi de contribuer à améliorer la qualité de vie. Nous avons dû redéfinir le but de nos interventions. Cette évolution a été renforcée par un changement culturel impliquant davantage les patients dans la prise de décision thérapeutique.

La santé est un aspect important de la qualité de vie globale. L'évaluation de la qualité de vie reliée à la santé repose sur la définition de la santé apportée par l'OMS en 1947 : « État de bien-être physique, mental et social complet et non simplement dû à l'absence de maladie ou d'infirmité » (25). Par conséquent, les modèles de qualité de vie spécifiquement reliés à la santé ont voulu traduire l'expérience subjective individuelle reliée directement ou indirectement à la santé, à la maladie, aux handicaps, aux déficits et à l'efficacité des traitements (26). La qualité de vie en médecine clinique peut se définir par la perception qu'a le patient de l'effet de sa maladie et de son traitement sur son fonctionnement. Cette définition reflète bien que le patient est la source principale d'informations dans l'évaluation de la qualité de vie reliée à la santé. Les domaines les plus souvent examinés dans l'étude de la qualité de vie reliée à la santé sont : le

fonctionnement physique, les symptômes, le jugement global sur sa santé, le bien-être psychologique, le bien-être social, le fonctionnement cognitif, les activités, la perception de soi et la satisfaction des soins (27). C'est seulement lorsque le patient est dans l'incapacité physique, mentale, cognitive et/ou psychologique de fournir cette information, que la source deviendra le jugement des proches ou des médecins.

3 Les questionnaires de qualité de vie

La recherche sur la qualité de vie a connu ses débuts après la deuxième guerre mondiale avec l'introduction d'une nouvelle définition de la santé par l'OMS. La recherche sur la QOL est devenue progressivement une discipline cohérente permettant de faire le pont entre les sciences cliniques et les sciences sociales.

Les mesures de la qualité de vie à l'aide de questionnaires ont de nombreuses utilités. D'abord, elles peuvent servir à décrire l'impact d'une maladie sur une population atteinte (28). Elles peuvent aussi aider à orienter l'allocation des ressources en soins de santé (28). La mesure la plus couramment utilisée dans ce contexte est le QALY (*Quality-Adjusted-Life-Year*) (28). Il permet de mesurer les années de vie gagnées suite à une intervention, ajusté pour la qualité de vie présente lors de ces années de vie additionnelles.

Les questionnaires de qualité de vie sont souvent utilisés dans les essais cliniques qui visent à évaluer l'impact d'une intervention. Ils permettent d'étudier l'efficacité d'une thérapie en fonction d'une amélioration de la qualité de vie associée à cette thérapie (28). Les échelles de qualité de vie peuvent aussi être utiles dans l'étude de différents traitements d'une maladie asymptomatique, telle que l'hypertension, en fonction de la tolérance aux effets secondaires. Enfin, ces instruments peuvent être utiles dans la pratique clinique, sur une base régulière, pour évaluer la perspective du patient, ses préférences, ses priorités et aussi pour améliorer la communication médecin-patient.

Il existe divers types d'échelle de qualité de vie qui diffèrent selon leurs objectifs spécifiques. Fitzpatrick (27) décrit sept types d'échelles différents dont l'échelle spécifique à la maladie, l'échelle spécifique au site ou à la région du corps, l'échelle spécifique à une dimension, l'échelle générique, l'échelle sommaire des items, l'échelle individualisée et l'échelle d'utilité. Les instruments les plus couramment utilisés en recherche clinique sont les échelles génériques et les échelles spécifiques.

3.1 Les questionnaires génériques

Les questionnaires de qualité de vie génériques sont construits pour être utilisés auprès de la population générale. Les domaines de la qualité de vie étudiés sont délibérément vastes afin d'évaluer la qualité de vie reliée à la santé de façon globale (28). Étant donné leur aspect plus global, ces instruments sont davantage en mesure de détecter les effets inattendus d'une intervention comparativement à un instrument spécifique qui s'oriente plutôt sur des conséquences connues et anticipées. Ces instruments sont particulièrement utiles lorsqu'il n'existe pas de questionnaire de qualité de vie spécifique pour étudier une condition médicale particulière. Par contre, ils restent des instruments peu sensibles aux changements de la condition médicale par rapport à la maladie étudiée.

Un bon exemple de ce type d'instrument est le SF-36 pour le *36-Item Short-Form Health Survey* (6). Cet instrument a été développé afin d'être utilisé en pratique clinique et en recherche clinique, pour évaluer l'impact des programmes d'intervention dans le domaine de la santé et pour obtenir un reflet de la qualité de vie globale dans la population générale. Il comprend trente-six items regroupés en huit sous-échelles, pour évaluer huit dimensions de la santé globale. Les huit dimensions étudiées sont : les limitations dans les activités physiques, les limitations dans les activités sociales, les limitations dans les activités de rôle en raison de problèmes de santé physique, la douleur physique, la santé mentale globale, les limitations dans les activités de rôle en raison de problèmes de santé mentale, la vitalité et les perceptions sur la santé globale.

Une version courte du SF-36 a été développée, le SF-12 ou *12-Item Short-Form Health Survey* (7). Elle comprend deux sous-échelles plutôt que huit, et douze items plutôt que trente-six. Cet instrument permet une évaluation de la qualité de vie plus rapide et pratique mais tout aussi valide. Une sous-échelle évalue l'aspect physique (*physical component summary* ou PCS) et l'autre l'aspect mental (*mental component summary* ou MCS) de la qualité de vie. Le score de chacune des sous-échelles varie entre 0 à 100. Plus le score est élevé, meilleure est la qualité de vie. Les résultats des sous-échelles physique et mentale sont toujours présentés séparément. Le SF-12 a été traduit en canadien français et a été validé auprès de la population masculine par Lebeau et collaborateurs en 2005 (8). Jusqu'à ce jour, le SF-12 n'a pas été validé auprès de la population canadienne française de sexe féminin.

3.2 Les questionnaires spécifiques

Ces instruments sont conçus pour évaluer la qualité de vie en lien avec des conditions médicales spécifiques. Ils ont l'avantage d'être plus sensibles aux symptômes ou aux problèmes associés à la maladie ou à l'intervention étudiée que les échelles génériques (28). Ces instruments sont plus susceptibles de détecter des changements dans le temps à propos de la condition médicale examinée. Des questionnaires spécifiques ont notamment été développés dans les domaines de l'urologie pour évaluer l'hyperplasie bénigne de la prostate ou les symptômes de l'arbre urinaire inférieur (28), de la rhumatologie pour évaluer l'arthrite et de la pneumologie pour évaluer l'asthme (27). L'acceptabilité de ces instruments auprès de la population étudiée s'est avérée bonne car le sujet d'étude apparaît pertinent aux yeux du patient. Par conséquent, le taux de réponse aux questionnaires est souvent supérieur au taux de réponse à un questionnaire générique. Par contre, il est moins approprié et pertinent d'administrer cet instrument à des individus qui ne sont pas atteints par la condition médicale étudiée car ils n'ont pas l'expérience de la maladie. Aussi, les questionnaires spécifiques ne peuvent pas évaluer les problèmes reliés à une maladie et/ou à son traitement qui n'ont pas été anticipés. Plus un instrument possède un éventail global d'items, plus il existe des possibilités de détecter des effets inattendus (27).

4 Construction et validation des questionnaires

4.1 Construction d'un questionnaire

La construction d'un questionnaire ou d'une échelle nous permet de mesurer des concepts abstraits (une émotion, une opinion, un trait de personnalité, un besoin) de façon plus fiable et précise. Il s'agit d'un long processus comprenant plusieurs étapes. Voici un résumé de ces différentes étapes à suivre selon une synthèse publiée par Paul E. Spector (29).

D'abord, il faut définir le construit d'intérêt, c'est-à-dire le concept abstrait que l'on veut mesurer, de façon claire et précise. Ensuite, on conçoit l'échelle en déterminant le nombre d'items que l'on veut inclure ainsi que le nombre de choix de réponses possible pour chaque item. Un item représente un aspect particulier du construit étudié. Une échelle doit normalement comprendre plusieurs items afin de refléter le mieux possible la complexité du construit. La version initiale doit être testée sur un petit groupe de sujets afin de détecter les items ambigus. Le questionnaire est révisé en fonction des commentaires reçus. Par la

suite, on administre le questionnaire auprès de 100 à 200 répondants pour en faire une analyse complète. On cherche à déterminer si l'ensemble des items choisis représente adéquatement le construit. Finalement, si les premières étapes sont réussies, on peut procéder à la validation du questionnaire par une série d'études de validation afin de s'assurer que l'échelle se comporte telle que prévue.

4.2 Validation d'un questionnaire

Le processus de validation d'un questionnaire implique des techniques de mesure psychométriques afin de vérifier si l'instrument étudié peut être utilisé avec un niveau raisonnable de confiance. Il faut donc tester le questionnaire en fonction de certains critères dont la fiabilité, la validité, la *responsiveness*, c'est-à-dire la capacité de l'instrument à détecter un changement cliniquement important, l'acceptabilité et la faisabilité (27, 28).

La **fiabilité** d'une échelle se réfère à la stabilité de sa mesure, à sa capacité de donner la même mesure en différentes occasions. Elle se définit par sa fiabilité des items (ou cohérence interne) et par sa fiabilité temporelle. La fiabilité des items entre eux traduit l'homogénéité des items et leur capacité à définir le construit étudié. On s'attend à ce que les items aient une bonne corrélation entre eux. Cette fiabilité est mesurée en calculant le coefficient alpha de Cronbach. Plus la valeur est élevée, meilleure est la corrélation. On recherche une valeur supérieure à 0,70. La fiabilité temporelle reflète la stabilité de la mesure sur une courte période de temps (2-14 jours), en assumant que le phénomène étudié reste inchangé. Elle est établie grâce au « test-retest ». Pour ce faire, le coefficient de corrélation intra-classe est l'analyse qui permet le mieux de discriminer les différences dues aux variabilités individuelles et celles causées par la variabilité de l'instrument.

La **validité** d'un instrument se définit par sa capacité à mesurer fidèlement le phénomène étudié. Elle se subdivise en validité de contenu, validité de critère et validité de construit. La validité de contenu cherche à déterminer si l'ensemble des items spécifiquement sélectionnés reflètent fidèlement le concept que l'on veut étudier. Pour ce faire, on demande le plus souvent à un panel d'experts de se prononcer sur l'adéquation apparente entre l'instrument proposé et le construit à mesurer. La validité de critère cherche à comparer les mesures obtenues avec l'instrument à valider avec les mesures obtenues d'un

autre instrument considéré comme plus adéquat. La validité de construit porte sur la relation entre les concepts théoriques et leur opérationnalisation. Elle peut être établie à l'aide du *known-group method* ou validité nomologique. Il s'agit de proposer des hypothèses sur les résultats que l'on devrait obtenir à l'aide de l'instrument, dans certaines conditions et pour certaines populations. Par exemple, les résultats d'un questionnaire de qualité de vie sur l'incontinence urinaire devraient différer entre un groupe de patientes continentes et un groupe de patientes incontinentes (27, 29, 30, 31).

La *responsiveness* du questionnaire est sa capacité à détecter les changements cliniques importants. La méthode la plus fréquemment utilisée pour évaluer cette propriété est le calcul de « l'importance de l'effet » (*effect size*). Il s'agit de diviser la différence des scores moyens post-traitement par la déviation standard des scores pré-traitement. Un effet de 0,2 est considéré faible, de 0,5 est considéré modéré et de plus de 0,8 est considéré important (27, 28). L'**acceptabilité** de l'instrument s'évalue simplement et directement par le taux de réponse aux questionnaires. Et enfin, la **faisabilité** d'un questionnaire dépend surtout de la charge de travail et des ressources économiques nécessaires pour la bonne utilisation de l'instrument (27).

5 Les questionnaires de qualité de vie spécifiques à l'incontinence urinaire

L'impact de l'incontinence urinaire sur la qualité de vie des femmes est parfois suffisant pour compromettre les activités sociales et le bien-être psychologique (32). Des évaluations objectives, cliniques et paracliniques (*stress test*, calendrier mictionnel, *pad test*, paramètres urodynamiques) peuvent nous orienter sur la sévérité des troubles urinaires. Par contre, ces mesures ne peuvent pas refléter le retentissement vécu par la patiente. Aussi, on constate que les bénéfices apportés par un traitement ne sont pas seulement d'ordre physiologique ou biologique. De ce fait, des questionnaires de qualité de vie ont été développés dans ce domaine spécifique afin d'évaluer la qualité de vie chez ces femmes (9, 33) et d'évaluer de façon plus globale, complète et réaliste l'effet d'une thérapie chez une population étudiée. Il est même officiellement recommandé, suite à la seconde Consultation internationale sur l'incontinence urinaire, d'inclure des questionnaires de qualité de vie spécifiques aux troubles urinaires dans tous les essais cliniques cherchant à étudier l'impact des traitements sur l'incontinence urinaire (34, 35).

La majorité des questionnaires spécifiques pour l'incontinence urinaire ont été développés en anglais. Quelques uns ont été traduits en français. Pour qu'un questionnaire devienne un outil valide, il doit d'abord subir un processus de validation psychométrique auprès de la population pour laquelle on veut utiliser l'instrument.

D'après une revue de la littérature extensive réalisée par Bonniaud en 2005 (34), les questionnaires sur l'incontinence urinaire disponibles en français sont: *Incontinence Impact Questionnaire* (IIQ) (9), le Ditrovie (36), le Contilife (37), le *Incontinence quality of life instrument* (I-QOL) (38) et le Qualiveen (39). Les trois premiers questionnaires ont été développés spécifiquement pour les troubles urinaires chez la femme. Le I-QOL et le Qualiveen peuvent être administrés aux deux sexes. Le Qualiveen a été validé auprès de patients souffrant de troubles urinaires secondaires à un trauma médullaire ou à la sclérose en plaques. Cet instrument doit être préférablement utilisé auprès de ce type de population présentant le plus souvent des problèmes urinaires spécifiquement reliés à une condition neurologique. Le Ditrovie a été validé auprès de femmes présentant des symptômes d'impériosité, d'incontinence urinaire d'urgence ou d'incontinence urinaire mixte (36). La population de notre étude est composée de femmes présentant principalement une incontinence urinaire de stress.

Le Contilife a été testé auprès de femmes présentant une incontinence urinaire soit d'effort, d'urgence ou mixte (37). Le I-QOL a été développé et validé chez les hommes et les femmes ayant une incontinence urinaire d'effort ou mixte. Par contre, leur validation a été réalisée auprès d'une population francophone européenne. Cette validation ne s'applique pas nécessaire à une population canadienne française car la culture et le vocabulaire diffèrent.

Enfin, seul le questionnaire IIQ a été validé auprès de la population canadienne française. Les résultats démontrent une bonne validité et une bonne fiabilité, similaires aux indices mesurés pour la version anglaise (10, 34). Il a été développé par Wyman et collaborateurs et est conçu pour une auto-administration (9, 40). Cet instrument a été créé afin de mesurer l'impact de l'incontinence urinaire sur les différents aspects de la vie quotidienne des femmes. Il s'agit d'une échelle de type Likert comprenant 30 items. La qualité de vie est évaluée en fonction de quatre sous-échelles : les activités physiques (6 items), les voyages (6 items), les relations sociales (10 items) et la santé émotionnelle (8 items). Chaque sous-

échelle est constituée d'items permettant d'évaluer un aspect plus spécifique de la qualité de vie. Pour chaque item, il existe une possibilité de quatre réponses en intervalle de 1 à 4 traduisant le degré d'impact de l'incontinence urinaire sur l'item étudié, où 1 égale pas du tout et 4 égale beaucoup. Le score total peut varier de 0 à 400. Le score total est obtenu en additionnant le score de chaque sous-échelle. Le score de chaque sous-échelle est calculé en faisant la somme des résultats de chacune des questions de cette sous-échelle; à cette somme on soustrait 1 et ensuite on multiplie par 100/3. Plus le score total est élevé, plus l'incontinence urinaire a un impact négatif sur la qualité de vie.

Suite à l'analyse des différentes caractéristiques des questionnaires français évaluant l'impact de l'incontinence urinaire, on peut conclure que le IIQ-30 est le seul instrument applicable à une population féminine adulte d'origine canadienne française et souffrant principalement d'incontinence urinaire à l'effort.

6 Utilisation prospective ou rétrospective des questionnaires de qualité de vie

Pour évaluer l'efficacité d'une thérapie, par exemple une chirurgie, en fonction de l'impact sur la qualité de vie, les questionnaires sont normalement administrés avant la chirurgie et ensuite à un certain moment après la chirurgie. Après chaque évaluation par questionnaire, un score de qualité de vie est obtenu et on fait la différence entre les deux scores. De cette manière, on détecte de façon prospective un changement de la qualité de vie qui reflète l'impact du traitement chirurgical pour chaque individu (41).

L'approche classique de l'évaluation d'une thérapie implique le suivi d'une cohorte de façon prospective impliquant souvent un groupe contrôle, comme celle qui est décrite ci-dessus. Cette approche est préférable au niveau théorique en raison de la force de l'évidence scientifique qu'elle peut procurer. Par contre, les ressources qu'elle exige sont importantes aux niveaux financier, humain et temporel. Elle n'assure pas non plus l'absence d'erreur ou de biais.

En pratique, une évaluation rétrospective de l'impact d'un traitement sur la QOL serait potentiellement attrayante. Elle impliquerait l'administration des questionnaires après le traitement seulement, pour recueillir l'évaluation de la qualité de vie depuis le traitement et aussi pour recueillir le rappel ou le souvenir que les sujets ont de leur qualité de vie avant le

traitement. Les deux scores seraient ensuite comparés comme on le ferait avec l'approche prospective. L'approche rétrospective a l'avantage d'être plus accessible à l'ensemble des cliniciens qui cherchent à connaître l'efficacité des traitements qu'ils administrent à leurs patients. Cette approche exigerait moins de ressources et d'organisation. Elle serait plus simple, rapide et économique. On pourrait également avoir accès à un plus grand nombre de sujets rapidement.

Cependant, il existe un obstacle à l'utilisation d'une méthode rétrospective. Plusieurs auteurs remettent en question la capacité des sujets de se rappeler adéquatement de leur qualité de vie (41, 42, 43, 44, 45). Plusieurs biais sont évoqués dont le biais de rappel, le biais motivationnel et le biais de mémoire. Pour valider une utilisation rétrospective d'une échelle de QOL dans le contexte de l'évaluation de l'effet d'un traitement, il faut démontrer que le rappel de la condition prétraitement en période post-traitement est équivalent à une évaluation prospective traditionnelle de la condition prétraitement. Si le rappel de la condition prétraitement est adéquat, c'est-à-dire équivalent à une évaluation prospective de la condition prétraitement, le questionnaire ainsi testé peut être utilisé comme outil d'évaluation rétrospective de la qualité de vie.

Dans le domaine de l'incontinence urinaire, l'utilisation d'une échelle spécifique ou globale pour une évaluation rétrospective de la qualité de vie n'a jamais été étudiée. Pour les raisons décrites précédemment, il serait pertinent d'évaluer la validité d'une utilisation rétrospective d'un questionnaire spécifique pour l'incontinence urinaire (IIQ-30) et d'un questionnaire de qualité de vie globale (SF-12).

7 Littérature sur le rappel

Plusieurs publications récentes s'intéressent à la validité d'une évaluation rétrospective de la qualité de vie ou de l'état de santé. L'objectif de ces études était d'évaluer et de comparer les mesures prospectives et rétrospectives dans des domaines spécifiques, tels que la gynécologie, l'orthopédie et l'urologie. Pour la plupart, des devis prospectifs de type « pre-post design » ont été utilisés (45). Chaque auteur rapporte l'exactitude du rappel des sujets de leur condition médicale ou de leur qualité de vie prétraitement.

Mancuso et Charlson (42) ont comparé les évaluations prospective et rétrospective de l'amélioration de la condition médicale suite à un remplacement total de la hanche

(n = 104) en se servant du *Hip Rating Questionnaire* (HRQ). Il s'agit d'un questionnaire spécifique, bien validé, conçu pour une auto-administration et évaluant quatre domaines (la douleur, la marche, le fonctionnement et l'impact sur la santé globale). L'instrument était administré en période préopératoire et ensuite quelques années après la chirurgie, soit en moyenne 2,5 ans après la chirurgie. La comparaison des réponses prospectives et rétrospectives s'est effectuée à l'aide du kappa et du coefficient de corrélation intra-classe. Les auteurs concluent que l'évaluation rétrospective n'a pas donné une mesure adéquate de la condition préopératoire. Le coefficient de corrélation intra-classe était de 0,44 pour le score total du questionnaire. Les auteurs concluent que l'approche rétrospective mène le plus souvent à une surestimation de l'efficacité du traitement. Cette étude utilise un devis clinique simple (*pre-post design*) mais approprié. Les analyses statistiques sont adéquates. Les résultats sont valables mais peuvent être biaisés, étant donnée l'absence de groupe contrôle. Il faut aussi mentionner que le questionnaire utilisé n'évalue pas la qualité de vie mais bien la condition médicale. Il s'agit donc d'une évaluation du rappel de la condition médicale. On peut aussi s'attendre à un mauvais rappel lorsqu'on évalue celui-ci quelques années après l'événement initial; là aussi, plusieurs biais entrent en jeu dont les biais d'histoire et de maturation (se référer à la section 7.1 sur les sources de biais potentiels). Le biais d'histoire s'explique par des événements survenant depuis la chirurgie et modifiant le souvenir de la condition préopératoire. Aussi, l'état psychologique du sujet peut être modifié par la fatigue et le vieillissement, et l'expérience acquise peut influencer les résultats et introduire un biais de maturation. Par exemple, une patiente avec un nouvel état dépressif peut avoir tendance à rapporter davantage de symptômes négatifs.

Litwin et McGuigan (43) ont évalué l'exactitude du rappel de la qualité de vie préopératoire chez les sujets ayant subi une prostatectomie radicale à l'aide d'une étude de cohorte observationnelle (n = 123). Les auteurs ont administré en période préopératoire un questionnaire de qualité de vie global, soit le SF-36, ainsi qu'un questionnaire spécifique, le *UCLA Prostate Cancer Index* (PCI) comprenant 20 items. Ces deux instruments ont été conçus et préalablement validés pour une auto-administration. L'évaluation du rappel s'est effectuée entre 7 et 37 mois après la chirurgie (moyenne de 21,1 mois) à l'aide du questionnaire de qualité de vie global SF-12 et du questionnaire spécifique UCLA PCI-SF comprenant 14 items (version courte du UCLA PCI). De façon générale, les sujets ont eu tendance à se rappeler une qualité de vie préopératoire meilleure que celle rapportée par

l'évaluation prospective. Grâce à une analyse multivariée, on remarque que le jeune âge et un niveau de scolarité plus élevé avaient tendance à diminuer cet effet. Il faut souligner ici que les auteurs n'ont pas utilisé exactement les mêmes questionnaires pour évaluer la qualité de vie préopératoire de base et pour évaluer la qualité de vie postopératoire et le rappel. Les sous-échelles de comparaison étaient par contre semblables étant donné qu'il s'agissait des versions courtes des mêmes questionnaires. Les résultats de l'étude démontrent qu'il existe une différence statistiquement significative entre les évaluations prospective et rétrospective. Des analyses de corrélation ont quand même été réalisées démontrant une corrélation modérée avec des résultats variant entre 0,41 et 0,57.

Legler et collaborateurs (44) ont étudié la capacité des hommes ayant un diagnostic du cancer de la prostate à se rappeler de leur condition médicale précédant le diagnostic (n = 133). Un questionnaire de santé et de qualité de vie spécifique comprenant 17 items et évaluant les fonctions urinaire, sexuelle et intestinale, a été administré quelques jours après le diagnostic pour évaluer la condition pré-diagnostique (45). Six mois après le diagnostic, le questionnaire a été réadministré afin de détecter les changements au niveau des fonctions urinaire, sexuelle et intestinale et d'évaluer le rappel de la condition pré-diagnostique. Malheureusement, aucune information n'est donnée sur l'origine du questionnaire et sur sa validation. Il est probable qu'il s'agisse d'un instrument maison non-validé. Les résultats de l'étude ont démontré qu'il existe un accord assez élevé entre l'évaluation prospective et l'évaluation rétrospective, avec des valeurs de kappa variant entre 0,40 et 0,92. Trois des quatre items évaluant la fonction urinaire avaient des valeurs de kappas de 0,66 à 0,76. Les coefficients de corrélations intra classes étaient de 0,83 pour la fonction urinaire, de 0,62 pour la fonction intestinale et de 0,70 pour la fonction sexuelle. On remarque que le rappel de la fonction urinaire est le plus exact. Dans la discussion, les auteurs font appel à la théorie de la dissonance cognitive pour expliquer la discordance des évaluations rétrospectives et prospectives rapportées par les études antérieures. La théorie de la dissonance cognitive explique la tendance des répondants à surestimer ou à exagérer les bénéfices d'une thérapie, par un processus de rationalisation d'une survie au traitement de leur condition (41). Les auteurs soumettent que malgré le fait que les mesures rétrospectives soient teintées par la dissonance cognitive, les mesures prospectives sont limitées d'autres façons. Premièrement, les réponses de l'évaluation de base peuvent être influencées par des événements perturbateurs survenant simultanément avec la condition

prétraitement ou pré-diagnostique. Deuxièmement, les sujets éprouveraient de la difficulté à demeurer constants dans leurs réponses à travers le temps. Ce phénomène décrit par Howard (46) est nommé le *response-shift bias*. Finalement, Legler et collaborateurs concluent que leurs résultats confirment la validité d'une évaluation rétrospective des changements des fonctions urinaire, sexuelle et intestinale chez les hommes atteints du cancer de la prostate (44). Il est intéressant de constater que l'évaluation rétrospective de la fonction urinaire était la plus fiable avec un excellent coefficient de corrélation intraclasse à 0,83. Par ailleurs, il est à noter que leurs résultats n'ont pas été teintés par l'expérience du traitement chez la majorité de leurs sujets. Seulement 24 sujets avaient reçu un traitement. Si l'intérêt d'un chercheur ou d'un clinicien est plutôt d'évaluer le changement de la qualité de vie en fonction d'une thérapie, ces résultats deviennent moins pertinents.

Karakiewicz et collaborateurs (49) ont testé la fiabilité du rappel de la fonction érectile chez les hommes ayant subi une prostatectomie radicale ($n = 39$). L'évaluation de la fonction érectile s'est effectuée à l'aide de l'*International Index of Erectile Function (IIEF)*. Cet instrument comprend 15 items évaluant la fonction érectile, la fonction orgasmique, le désir sexuel, la satisfaction coïtale et la satisfaction globale. L'index a été administré en préopératoire, à 6 mois et 12 mois postopératoires. Pour chaque moment d'administration, le rappel était évalué. Les résultats ont démontré une bonne correspondance entre l'évaluation prospective et rétrospective de l'évolution de la fonction érectile de la période préopératoire jusqu'à 12 mois postopératoires. Les coefficients de corrélation intra-classe étaient respectivement de 0,65 à 6 mois et 0,75 à 12 mois par rapport à la fonction érectile. Les auteurs suggèrent que l'évaluation rétrospective de conditions plus concrètes telle que la fonction érectile est plus fiable que l'évaluation rétrospective de domaines plus abstraits telle que la fonction orgasmique. Ils concluent que les scores de qualité de vie obtenus de façon rétrospective estiment bien les scores de qualité de vie obtenus de façon prospective. Par contre, il faut rester conscient de certaines sources de distorsion causée entre autres par l'effet *re-coding*, le biais motivationnel et le biais de pratique. L'effet *re-coding* est celui qui survient chez les sujets qui ont reçu de l'information en période préopératoire sur les effets possibles de la chirurgie. Le biais de pratique s'introduit lorsque les sujets sont invités à effectuer plusieurs évaluations, telles que de répondre à des questionnaires qui sont réadministrés plusieurs fois lors de la même étude. Les sujets développent une meilleure compréhension et deviennent aptes à répondre à ces questionnaires, en proportion

avec le nombre d'évaluation requise. Le biais motivationnel a été expliqué préalablement. Les auteurs concluent que, pour diminuer la distorsion de l'évaluation rétrospective de la qualité de vie, les domaines étudiés doivent être de complexité limitée et comprendre un nombre restreint d'items. Enfin, ils estiment que l'évaluation de la qualité de vie doit être réalisée à un moment où la fonction mesurée a atteint un niveau stable. Il s'agit d'une étude de qualité ayant utilisé un devis et des analyses statistiques adéquates.

Enfin, une seule étude a tenté d'évaluer de façon rétrospective, à l'aide de questionnaires, le domaine de l'incontinence urinaire (50). Par contre, on n'évalue pas la qualité de vie mais plutôt la condition médicale. Viktrup et Lose ont évalué l'habileté des femmes en âge de procréer de se rappeler du début de leur incontinence urinaire à l'effort. Trois cents cinq femmes ont été incluses au moment du premier accouchement. Ces femmes ont répondu verbalement à un questionnaire à deux moments différents: quelques jours après l'accouchement et trois mois plus tard par téléphone. Le même questionnaire a été ré-administré par la poste cinq ans plus tard. À ce moment, soixante-huit patientes ont rapporté avoir développé une incontinence urinaire de stress depuis leur premier accouchement. La capacité de ces patientes à rapporter le moment exact du commencement de leur incontinence urinaire de stress a été faible. Quarante-neuf pour cent des patientes ont donné un rappel inexact. Les valeurs de kappa ont varié de 0,02 à 0,38. Les auteurs tentent d'expliquer ces résultats de différentes façons. D'abord, l'intervalle de temps et la signification de l'événement étudié peuvent influencer énormément l'exactitude du rappel (51). Selon l'expérience clinique des auteurs, si la patiente n'a été que peu incommodée par son problème d'incontinence urinaire, il est fort probable qu'elle ne se souvienne pas précisément du moment du début de ce problème. Par ailleurs, il peut être indésirable socialement d'exposer ce problème. Enfin, il est possible que l'utilisation de plusieurs modes d'évaluation, soient l'entrevue au département d'obstétrique, le téléphone à trois mois et la poste à cinq ans, ait pu introduire des biais. Les auteurs n'ont pas précisé quels étaient le ou les questionnaires utilisés, ni s'ils ont été validés.

Une dernière étude se démarque par sa qualité scientifique. Bryant et collaborateurs (52) ont voulu évaluer la capacité des patients à se rappeler de leur fonctionnement et de leur qualité de vie préopératoire, deux semaines après avoir subi une chirurgie du genou. Les

sujets ont été randomisés en deux groupes. Le groupe I a eu une évaluation prospective et rétrospective. Le groupe II ne subissait qu'une évaluation rétrospective. Grâce à ce type de devis, on pouvait détecter un biais de pratique face à l'exposition à l'instrument de mesure chez le groupe I. À chaque évaluation, des questionnaires de qualité de vie d'ordre spécifique (*International Knee Documentation Committee (IKDC)*, *Knee Injury and Osteoarthritis Outcome Score (KOOS)*) et général (SF-36) étaient distribués. Ensuite, si le sujet subissait une reconstruction du ligament croisé antérieur, il devait également compléter le *Quality of Life Outcome Measure (ACL-QOL)* et si le sujet subissait qu'une simple arthroscopie du genou, il devait répondre au *Western Ontario Meniscal Evaluation Tool (WOMET)*. Une différence de score de qualité de vie était considérée importante lorsqu'il existait une différence de 20% entre le score de base en préopératoire et celui du rappel. Pour les cinq questionnaires administrés, les coefficients de corrélation intra-classe variant entre 0,67 et 0,93 indiquaient un rappel adéquat de la qualité de vie préopératoire à deux semaines postopératoire. Il n'y a pas eu de différence statistiquement significative entre le groupe I et le groupe II par rapport au score du rappel de la qualité de vie préopératoire, suggérant l'absence d'effet causé par l'exposition préalable à l'instrument de mesure. Les auteurs concluent que l'on peut raisonnablement substituer des données préopératoires par des données récoltées à deux semaines postopératoires. Ceci permettrait d'améliorer l'efficacité de la collecte des données lors de futures études prospectives ou bien de procéder à des études rétrospectives non-planifiées.

7.1 Identification des sources de biais potentiels

Plusieurs biais peuvent affecter la validité d'une étude comme celle que nous proposons. D'abord, certains biais toucheront la **validité interne** en raison du choix de devis. Face à l'absence de groupe contrôle simultané dans le temps, des biais d'histoire, de l'effet du traitement, de maturation et d'accoutumance au test sont possibles (30). Le biais d'histoire est occasionné par l'apparition d'un événement durant l'étude qui aura une influence sur le score de qualité de vie. Par exemple, le décès d'un être cher à proximité du moment où le sujet doit répondre questionnaire évaluant le rappel de la QOL pré-traitement peut influencer significativement les réponses aux questionnaires. Le biais de l'effet du traitement sur la qualité du rappel est aussi décrit par Herrmann (48). Pour l'évaluer, il faut comparer le rappel d'un groupe avec traitement avec le rappel d'un groupe sans traitement,

de façon randomisée contrôlée. Le biais de maturation implique les effets du vieillissement, de la fatigue et de l'expérience acquise par les sujets entre la période préopératoire et la période postopératoire. Litwin et McGuigan (43) rapporte aussi l'effet du *response-shift biais* où le sujet éprouve une difficulté à demeurer constant dans ses réponses à travers le temps. Lorsqu'un test est administré plus d'une fois à un sujet, il est possible qu'un biais d'accoutumance au test (ou biais de pratique) influence les résultats. Le sujet peut alors mieux répondre (ou encore moins bien répondre) lors des évaluations subséquentes car il connaît le test. Le biais de mesure survient lorsque les instruments de mesure utilisés sont remplacés ou modifiés au cours de l'étude ou bien lorsque l'on change la façon de les administrer. Le biais motivationnel est occasionné par des sujets qui sont très sensibilisés à l'activité de la recherche et essaient plus intensément de se souvenir, de remarquer et de rapporter des symptômes reliés au traitement.

Selon Karakiewicz (49), d'autres aspects d'une étude évaluant le rappel peuvent modifier les résultats. La période de temps entre le traitement et le moment choisi pour évaluer le rappel est un facteur à considérer. Plus le délai est long, plus on s'attend à un rappel difficile et inexact. Le rappel est alors plus sujet aux biais d'histoire ou de maturation. D'après une revue de la littérature décrite dans la discussion de l'étude de Bryant (52), les études ayant évalué le phénomène du rappel sur des périodes plus éloignées du traitement soit entre deux mois et dix ans ont démontré l'inexactitude du rappel des sujets comparativement à des évaluations à intervalles plus rapprochées, soit d'une journée à deux semaines suivant le traitement. La stabilité dans le temps du phénomène observé est un autre aspect impliqué (49). Si le phénomène est instable, les résultats obtenus risquent de ne pas être fiables. Il sera difficile de prouver que les deux types d'évaluation de l'efficacité du traitement sont équivalents.

Par ailleurs, la littérature sur le sujet rapporte des biais spécifiques associés au phénomène du rappel. On parle de sources de distorsion de l'information qui affectent l'exactitude du rappel (41, 42, 48, 49). D'abord, Herrmann (48) rapporte trois sources de distorsion de l'information. Premièrement, le type de question posée peut influencer la qualité du rappel. Une question spécifique risque d'être répondue plus adéquatement et avec plus de validité qu'une question générale. Deuxièmement, la distorsion de l'information peut survenir lors du processus intellectuel nécessaire au rappel. Et troisièmement, l'information peut devenir

biaisée selon le type de devis utilisé. Des biais de rappel, motivationnel, *encoding* et de *group-skill* sont susceptibles de survenir avec un devis où l'on compare un groupe avec traitement, avec un groupe sans traitement. Ces biais sont causés tout simplement par l'effet du traitement. Le biais de rappel serait causé par une fatigue du sujet suite au traitement. Le biais de rappel se rapproche donc du biais de maturation. Avec le biais *encoding*, le sujet connaît les effets de la chirurgie et les enregistre. Cela cause souvent un rappel de plus de symptômes qu'il y en a en réalité. Le biais *group-skill* survient avec l'absence de randomisation entre le groupe traité et le groupe non-traité. L'auteur propose alors un devis « pre-post ou prospectif » qui implique le suivi d'une cohorte qui donne « son rappel » avant le traitement et son rappel après le traitement. On élimine le biais *group-skill* car il n'y a qu'un groupe, mais on introduit le biais de pratique. Hermann propose finalement un devis combinant les deux premiers devis discutés avec l'ajout d'un groupe placebo afin de limiter l'introduction de biais. L'auteur affirme tout-de-même la supériorité de la mesure prospective sur la mesure rétrospective qui dépend trop de la mémoire du sujet concernant son état de santé antérieure. Un rappel rétrospectif d'un changement au niveau des symptômes d'un sujet n'est que le substitut d'une mesure ponctuelle des symptômes avant et après un traitement. Par contre, la mesure rétrospective est utile dans le contexte où le but est d'évaluer les croyances du sujet à propos de l'efficacité du traitement. Partant de cette dernière opinion d'Hermann, l'évaluation de la qualité de vie implique effectivement les croyances du sujet sur l'efficacité d'une thérapie.

Les caractéristiques des sujets peuvent influencer l'exactitude du rappel. Parmi elles, le jeune âge et un bon niveau d'éducation sont des données sociodémographiques connues comme pouvant influencer positivement l'exactitude du rappel de la qualité de vie préopératoire (43). Les caractéristiques de la maladie étudiée peuvent tout aussi bien affecter la qualité du rappel dont la sévérité, la stabilité et l'aspect concret du problème étudié (49, 50).

De son côté, la **validité externe** peut être touchée par d'autres biais. Les résultats ne pourront être généralisés qu'au même type de population. Selon Contandriopoulos (30), le devis quasi-expérimental peut être touché au niveau de sa validité externe par plusieurs biais dont l'attente de l'expérimentateur, des interactions entre les observations et l'intervention et des interactions entre la sélection des sujets et l'intervention. Les

interactions entre la sélection des sujets et l'intervention surviennent en particulier lorsque l'étude est réalisée sur des sujets volontaires. Il est alors difficile de discriminer ce qui est dû au volontariat et ce qui est dû à l'intervention proprement dite.

8 Pertinence de la recherche

Suite à cette revue de littérature, il est plausible que l'évaluation rétrospective de la condition préopératoire des femmes souffrant d'incontinence urinaire à l'effort soit valide. Un intervalle de temps de trois mois est relativement court par rapport au moment de la chirurgie qui constitue en soi un événement marquant. L'incontinence urinaire est une condition médicale concrète et moins sujette à des interprétations subjectives. De ce fait, son évaluation rétrospective risque d'être plus conforme à l'évaluation prospective. Déjà, des auteurs ont démontré que le rappel de la fonction urinaire peut rester adéquat jusqu'à six mois (44). De plus, la population de femmes souffrant d'incontinence urinaire à l'effort désirant un traitement chirurgical consiste en une catégorie de patientes plus sévèrement atteintes. Étant donné la sévérité des symptômes, il s'agit d'une condition significative dans la vie de ces patientes et nous pensons que le souvenir de leur condition préopératoire en sera plus exact et que l'impact des biais potentiels sera relativement faible.

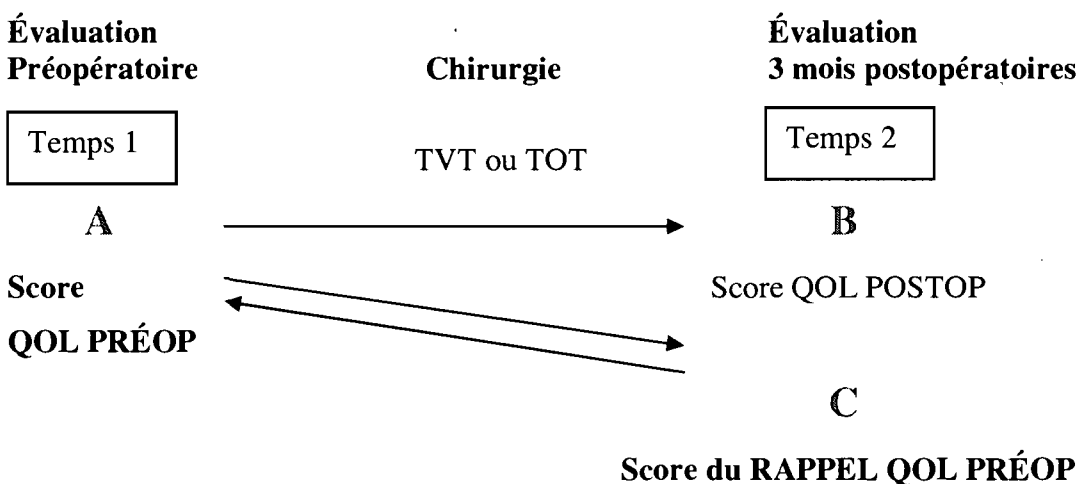
Enfin, si les résultats de notre étude suggèrent que l'approche rétrospective est valide, les résultats de cette étude permettront aux cliniciens d'utiliser les questionnaires IIQ-30 et SF-12 de façon rétrospective afin de connaître l'efficacité de leurs traitements chirurgicaux par rapport à l'amélioration de la qualité de vie de leurs patientes. Cela ajoute une alternative valable à l'évaluation de l'efficacité des traitements.

Chapitre II : Méthodologie

1 Objectifs

L'**objectif principal** est d'évaluer l'exactitude du rappel de la qualité de vie préopératoire, trois mois après une chirurgie de type TVT ou TOT chez les femmes souffrant d'incontinence urinaire à l'effort.

Représentation schématique de l'objectif principal:



Question de recherche:

Une évaluation rétrospective de l'impact d'une chirurgie sur la qualité de vie est-elle équivalente à une évaluation prospective ?

Ou

Score de QOL préop – score de QOL postop = Score du rappel de QOL préop – Score de QOL postop

Ou

$$A - B = C - B \rightarrow A = C ???$$

L'**objectif secondaire** est de valider le principal questionnaire utilisé, soit le IIQ-30, au sein de la population à l'étude.

2 Hypothèse

Le score du rappel de la qualité de vie préopératoire (C) des patientes opérées pour une incontinence urinaire de stress est équivalent au score de la qualité de vie préopératoire obtenue de façon prospective (A).

3 Devis

Dans le contexte d'une recherche dans le domaine des sciences sociales, on parle d'une stratégie de recherche développementale. Il faut déterminer si l'utilisation rétrospective de questionnaires de qualité de vie spécifique et globale, ici IIQ-30 et SF-12, peut nous permettre d'évaluer de façon fiable l'efficacité des chirurgies correctrices de l'incontinence urinaire. Selon Herrmann (48), un devis simple souvent utilisé pour étudier le rappel est le *pre-post design or prospective design*. Il s'agit d'un devis quasi-expérimental appelé en français l'étude avant-après (30). Ce devis élimine le *group skills bias*. Cependant, il reste assujéti à d'autres biais dont les biais motivationnel, de pratique et *encoding*.

Dans le contexte d'une stratégie de recherche développementale, nous évaluerons la fiabilité et la validité de l'instrument principal utilisé soit le IIQ-30, par une revalidation dans la population à l'étude.

4 Population à l'étude

L'étude est basée sur une cohorte prospective de sujets. La population cible est constituée de femmes parlant et lisant le français et qui sont en attente d'une chirurgie de type TVT ou TOT pour traiter un problème d'incontinence urinaire à l'effort. Le diagnostic d'incontinence urinaire à l'effort a préalablement été établi par le médecin qui a déposé la demande d'admission pour la chirurgie de TVT ou TOT. Il pouvait aussi bien s'agir de patientes ayant un diagnostic d'incontinence urinaire à l'effort pure ou un diagnostic d'incontinence urinaire mixte, mais avec une prédominance d'incontinence à l'effort. Cependant, le type spécifique d'incontinence urinaire à l'effort (pure ou mixte) n'a pas été intégré dans la base de données. Ce groupe de sujets est identifié comme étant le groupe « **incontinentes** ». Des **critères d'exclusion** ont été choisis afin de contrôler l'introduction de biais.

- Troubles cognitifs pouvant nuire à la bonne compréhension des questionnaires,
- Patientes en attente d'une chirurgie de type TVT ou TOT, conjointement avec une chirurgie visant à corriger un prolapsus génital,
- Diagnostic psychiatrique : Trouble dépressif ou anxieux sous médication, schizophrénie,

- Maladie chronique grave : IRC sous dialyse, insuffisance cardiaque, MPOC, ou autre pathologie pouvant affecter la qualité de vie,
- Incontinence fécale ou aux gaz,
- Handicap majeur dont la paralysie, la surdité, la cécité,
- Comorbidité neurologique pouvant affecter la continence : SEP, AVC, Parkinson, néoplasie cérébrale, trauma de la moelle épinière, démence.

Nous nous sommes guidés sur la méthodologie employée lors du projet de validation du questionnaire IIQ-30 au Québec. L'équipe de S. Beaulieu (10) avait exclu les sujets atteints de troubles psychiatriques ou de maladies chroniques en raison de leur trop grande influence sur le score de qualité de vie. Pour les mêmes raisons, l'incontinence fécale et/ou aux gaz et les handicaps majeurs dont la paralysie des membres, la cécité et la surdité ont été ajoutés aux critères d'exclusion.

Un deuxième groupe de sujets est sélectionné afin de procéder à la validation du questionnaire IIQ-30 dans la population cible et de répondre à l'objectif secondaire de l'étude. Afin d'examiner la validité de construit des questionnaires, il faut sélectionner un groupe de sujets ne souffrant pas d'incontinence urinaire. Afin de mieux se comprendre et d'éviter des confusions, ce groupe est identifié comme le groupe « **continentes** ». Les critères d'inclusion sont : femmes âgées entre 18 et 60 ans; ayant la capacité de parler et de lire le français; ne souffrant pas d'incontinence urinaire au moment du recrutement par l'infirmière de recherche. Les critères d'exclusion sont : incapacité de lire les questionnaires ou d'écrire, incapacité de comprendre les questionnaires en raison de troubles cognitifs, incontinence urinaire, fécale ou aux gaz, prolapsus génital, diagnostic psychiatrique, maladie chronique grave ou handicap majeur tel que défini précédemment.

5 Recrutement

L'échantillon a été sélectionné selon un **échantillonnage non-probabiliste de type accidentel**, autant pour le groupe « incontinentes » que pour le groupe « continentes » (30).

Les sujets constituant le groupe « incontinentes » ont été sélectionnés à partir de la population de patientes en attente d'une chirurgie de type TVT ou TOT, via le bureau d'admission du pavillon St-Luc du Centre hospitalier universitaire de Montréal (CHUM).

Ces patientes ont été approchées par une infirmière de recherche au moment de leur rendez-vous à la clinique préopératoire. Si elles acceptaient de participer à l'étude, elles signaient trois copies du formulaire de consentement et une copie leur était remise.

Les sujets du groupe « continentales » ont été sélectionnés à partir de la population de patientes des cliniques de gynécologie générale, consultant leur médecin pour un simple suivi annuel, dans ce même centre hospitalier. Ces patientes ont été approchées par le personnel médical de la clinique: infirmières, résidents ou médecins. Si elles acceptaient de participer à l'étude, elles signaient trois copies du formulaire de consentement et une copie leur était remise.

6 Collecte des données et suivi

À l'intérieur de cette étude, il n'y a que deux sources de données. Il s'agit principalement d'informations fournies par les sujets. Pour vérifier certaines données, une revue des dossiers hospitaliers a également été réalisée. Les informations fournies par les sujets ont été recueillies par l'utilisation de questionnaires de qualité de vie conçus pour une auto-administration. Les caractéristiques des répondants et des non-répondants ont été récoltées lors de l'évaluation préopératoire grâce à l'ajout d'un questionnaire s'informant des données sociodémographiques des sujets (Annexe I).

Au total trois différents questionnaires ont été administrés aux sujets : deux questionnaires de qualité de vie, IIQ-30 (Annexe II) et SF-12 (Annexe III), et un questionnaire récoltant les données sociodémographiques. Ces échelles ont été administrées autant aux patientes du groupe « incontinentes » pour permettre de répondre à la question de recherche, qu'aux patientes du groupe « continentales » pour effectuer la validation du questionnaire IIQ-30.

Suite à la signature du formulaire de consentement, les sujets du groupe « **incontinentes** » ont répondu par écrit aux trois questionnaires fournis par une infirmière de recherche. Les deux questionnaires de qualité de vie SF-12 et IIQ-30 leur ont été réadministrés trois mois après la chirurgie, avec l'ajout d'une copie pour évaluer spécifiquement le rappel. Les questionnaires évaluant le rappel de la qualité de vie postopératoire sont les mêmes questionnaires que ceux administrés en préopératoire, par contre les questions ont été adaptées au niveau des temps de verbe afin d'éviter la confusion chez le sujet. Ils ont été envoyés par la poste avec une lettre de présentation et une enveloppe de retour pré-adressée

et préaffranchie, quelques jours avant la date précise du 3 mois post-opératoires. La lettre de présentation permettait de rappeler aux patientes leur participation à l'étude et de leur indiquer la date précise où elles devaient répondre aux questionnaires. Une secrétaire de recherche a effectué des rappels téléphoniques pour les non-répondants, deux semaines après l'envoi postal des questionnaires. Les sujets ayant renvoyés un ou des questionnaires incomplets ont été contactés par téléphone par une infirmière de recherche afin de compléter adéquatement leurs questionnaires postopératoires. Les sujets ayant répondu tardivement aux questionnaires post-opératoires **et ce plus de quatre mois après la date de chirurgie** étaient éliminés. La période de temps entre le moment du recrutement, qui correspond au moment de l'évaluation préopératoire, et le moment de la chirurgie pouvait varier de façon indéterminée. Pour la majorité des sujets, cette période de temps était inférieure à deux semaines. Par contre, certaines chirurgies ont pu être retardées pour différentes raisons pouvant aussi bien découler d'une décision du chirurgien ou de la gestion hospitalière. Les sujets dont la chirurgie a été retardée de plus de six mois par rapport au moment du recrutement ont été éliminés.

Les sujets du groupe « **continentes** » ont répondu aux questionnaires IIQ-30 et SF-12, au moment de leur recrutement. Cette évaluation servira à établir la validité de construit du questionnaire IIQ-30 (voir section 7.3).

7 Analyses statistiques

La majorité des analyses ont été effectuées à l'aide du logiciel SAS version 8,02 sauf pour le calcul de la taille d'échantillon. Tous les tests ont été réalisés de façon bivariée et le niveau statistiquement significatif a été fixé à 0,05.

7.1 Calcul de la taille d'échantillon

La taille d'échantillon nécessaire a été calculée à l'aide d'un test d'équivalence pour deux moyennes corrélées avec distribution normale, en utilisant le programme PASS 2005 (53).

Voici la formule utilisée par ce programme :

$$N = \sigma^2 (1/k + 1/1 - k) (Z_\alpha + Z_\beta)^2 / \delta^2$$

où N est le nombre total de sujets, σ^2 est la variance, k et $(1-k)$ représentent les proportions de l'échantillon total pour chaque groupe de sujets, Z_α et Z_β sont les quantiles supérieurs de la distribution normale correspondant aux valeurs α et β prédéterminées, et δ est la différence minimale que l'on veut détecter.

Une moyenne du score de qualité de vie total du IIQ-30 de 171 avec une déviation standard de 101, provenant de l'étude d'Albo et al (54), ont été utilisées dans le calcul.

Dans le contexte d'une étude d'équivalence, il fallait déterminer à l'avance quelle est la différence entre le score de qualité de vie préopératoire de base et le score du rappel de la qualité de vie préopératoire évaluée en postopératoire que l'on allait considérer inacceptable (δ). La seule étude dans la littérature discutant de cette différence est celle de Bryant et al (52). Une différence de 20% ou moins entre le score de qualité de vie de base et celui du rappel ne constituait pas une différence cliniquement significative, selon les auteurs. Donc, si les scores de qualité de vie de base et celui du rappel ne différaient pas de plus de 20%, ils étaient considérés équivalents. Se servant de la même hypothèse que Bryant (52) et des données de l'étude d'Albo (54), le nombre de sujets nécessaires calculé, avec une valeur α fixée à 0,05 et une valeur β fixée à 0,20, tout en conservant une corrélation minimale de 0,5 entre le score de qualité de vie de base et celui du rappel, serait de 79. Si la corrélation est meilleure, le nombre de sujets nécessaires sera moindre. En estimant que la pire corrélation que nous pourrions obtenir serait de 0,5, nous considérons que la taille d'échantillon minimale sécuritaire serait de 79. Le même calcul a été appliqué au questionnaire SF-12. Le nombre de sujets nécessaires étant largement inférieur à celui nécessaire pour le IIQ-30, nous avons planifié l'étude qu'en fonction du calcul effectué pour le IIQ-30.

Le calcul de la taille d'échantillon nécessaire pour répondre au deuxième objectif de l'étude est discuté dans la section 7.4.

7.2 Analyses descriptives

D'abord, le taux de réponse aux questionnaires a été calculé en faisant le rapport entre les sujets ayant retourné leurs questionnaires moins de quatre mois après la chirurgie sur le total des sujets ayant donné leur consentement à l'étude. La moyenne (ou la médiane selon

le type de distribution) de chacune des variables sociodémographiques a été calculée. Elles incluent l'âge, l'ethnie, l'état civil, la situation de vie, le tabagisme, le niveau d'éducation, l'emploi, le revenu et les comorbidités. Les données sociodémographiques ainsi que les scores de QOL préopératoire de base des répondantes et des non-répondantes ont été comparés.

7.3 Analyses reliées à la question de recherche

La moyenne du score de QOL préopératoire et celle du rappel ont été comparées en fonction du critère d'équivalence préalablement établi correspondant à une différence de moins de 20% entre le score de base et celui du rappel pour chacun des questionnaires (52). Le test de t de Student n'étant pas l'analyse statistique adéquate dans le contexte d'une étude d'équivalence, il n'a donc pas été utilisé pour effectuer cette comparaison.

Le but de l'étude est d'évaluer l'exactitude du rappel de la qualité de vie préopératoire. La validité et la fiabilité sont les deux aspects qui permettent le mieux d'estimer la précision d'une mesure. En se référant à l'étude de Bryant et al (52), la validité et la fiabilité du rappel ont été évaluées distinctement.

Pour évaluer la validité du rappel, le score de qualité de vie en période préopératoire a été identifié comme étant la variable indépendante (x) et le score du rappel de la qualité de vie a été identifié comme étant la variable dépendante (y) de façon arbitraire. Une régression linéaire par analyse de corrélation de Pearson a été réalisée de façon à étudier la force de l'association entre x et y. La force de l'association entre x et y était considérée faible lorsque que $r < 0,3$, modéré si $r = 0,3 - 0,45$, importante si $r = 0,45 - 0,6$ et élevé si $r > 0,6$ (55).

La fiabilité du rappel, qui concerne sa reproductibilité chez un même sujet, a été évaluée grâce au coefficient de corrélation intra-classe (ρ). L'analyse de corrélation de Pearson permet d'évaluer la force de l'association entre deux variables mais pas nécessairement leur accord ou leur concordance. L'analyse du coefficient de corrélation intra-classe permet d'évaluer la concordance et de tester la fiabilité de l'instrument. Il existe six types de formules pour calculer le coefficient de corrélation intraclasse (56). La formule choisie dépend de l'objectif de l'étude, du devis utilisé et du type de mesures réalisées. Chaque type de coefficient de corrélation intraclasse se définit par deux chiffres entre parenthèses.

Le premier chiffre détermine le modèle (1, 2 ou 3) et le deuxième chiffre détermine la forme. La forme peut utiliser une seule mesure (l) ou la moyenne de plusieurs mesures (k) comme unité d'analyse. Dans le contexte de notre étude, la formule (3,1) a été considérée comme étant la plus appropriée. Elle utilise le modèle 3 qui est suggéré lorsque l'on veut tester la fiabilité «intra-évaluateurs» ou l'*intrarater reliability*. Ces analyses ont été effectuées avec le programme SAS grâce aux formules développées par Douglas Steinley & Phillip Wood (57). La fiabilité était considérée faible avec un coefficient $\rho < 0,5$, modérée avec un $\rho = 0,5 - 0,75$ et bonne si le $\rho > 0,75$ (58).

Enfin, des analyses de régression ont été effectuées afin d'explorer l'influence de certaines variables sociodémographiques et l'influence de l'intervalle de temps entre le moment du rappel et le moment de l'évaluation préopératoire sur l'exactitude du rappel. La différence entre le score mesuré en préopératoire et le score du rappel était considérée comme la variable dépendante (A – C). Les caractéristiques sociodémographiques ainsi que la durée de temps entre le rappel et l'évaluation préopératoire constituaient les variables indépendantes.

7.4 Analyses reliées à la revalidation du questionnaire IIQ-30

Actuellement, il n'existe pas de consensus sur la taille d'échantillon nécessaire pour procéder à la validation d'un questionnaire. Les auteurs Tinsley & Tinsley (59) (Tinsley HEA, Tinsley DJ. Uses of Factor Analysis in Counseling Psychology Research. *Journal of Counseling Psychology* 1987; 34: 414-24) ont proposé d'obtenir 5 à 10 sujets par item. Pour établir une fiabilité temporelle d'un questionnaire, Nunnally (60) (Nunnally JC. *Psychometric Theory* 1978, 2nd ed. New York: McGraw-Hill) recommande également l'utilisation de 5 sujets par item. Le questionnaire IIQ-30 comprend quatre sous-échelles dont la quantité des items varient entre 6 et 10 pour un total de 30 items. L'étude de référence en ce qui concerne la validation du IIQ-30 a utilisé un groupe de 162 sujets (9). Beaulieu et al ont procédé à la validation du IIQ-30 auprès d'une population féminine canadienne-française à l'aide d'un groupe de 34 sujets (10). Pour procéder à la validation du IIQ-30 dans le contexte de notre étude, nous avons prévu tester la fiabilité des items, ainsi que la validité de critère sur un minimum de 79 patientes incontinentes. Étant donné l'absence de consensus, et se référant aux publications antérieures sur la validation du IIQ-

30, nous avons estimé que la sélection additionnelle de 50 patientes dans le groupe « continentales » serait adéquate pour établir la validité de construit.

D'abord, la fiabilité des items a été évaluée en calculant le coefficient alpha de Cronbach (61).

Classiquement, l'évaluation de la validité de critère se fait en comparant les résultats aux questionnaires avec des mesures cliniques. Cependant, au cours de notre étude, aucune mesure clinique n'a été effectuée. La validité de critère (ou validité concomitante) a été testée en comparant les résultats du IIQ-30 (10) avec ceux d'un deuxième questionnaire considéré comme l'« étalon-or » en matière d'évaluation de la qualité de vie, le SF-12 (7). Il suffit d'établir un degré de convergence entre les deux échelles en effectuant une analyse de corrélation. Le degré de convergence est considéré faible lorsque que $r < 0,3$, modéré si $r = 0,3-0,45$, importante si $r = 0,45-0,6$ et élevé si $r > 0,6$ (55).

Afin de procéder à la validité du construit, le questionnaire IIQ-30 a été administré à un groupe de sujets continents (groupe appelé « continentales »), sans autre morbidité pouvant influencer les réponses aux questionnaires. Ces patientes sont celles qui ont été recrutées lors d'un rendez-vous pour un suivi annuel. Il s'agit de prouver que le questionnaire IIQ-30 est capable de discriminer entre ces deux groupes de patientes très différentes. Une analyse de corrélation a été réalisée afin d'établir la validité de construit du questionnaire IIQ-30.

Chapitre III : L'Article

ACCURACY OF RECALL IN HEALTH-RELATED QUALITY OF LIFE ASSESSMENT AMONG WOMEN OPERATED ON FOR STRESS URINARY INCONTINENCE

**Annick Larochelle, MD, John E. Podoba, PhD ,
Stéphane Ouellet, MD, William D. Fraser, MD, PhD.**

Abstract

Objective

To assess the validity and reliability of retrospective quality of life (QOL) assessment compared to prospective assessment among women treated for stress urinary incontinence (SUI) with a tension-free vaginal tape (TVT) or a trans-obturator tape (TOT) procedure.

Methods

One hundred and fifty-two patients were enrolled in a prospective cohort study at the Centre hospitalier universitaire de Montréal (CHUM). The Incontinence Impact Questionnaire (IIQ-30) and the Short-Form Health Survey (SF-12) were self-administered pre-operatively to consenting women scheduled for a TVT or TOT procedure. At 3 months post-op, the IIQ-30 and SF-12 surveys were mailed to these patients to reassess their pre-operative QOL status. Pearson's correlation coefficient (r) and the intraclass correlation coefficient (ICC) were used to test the validity and reliability of the recalled IIQ and SF-12 scores versus baseline preoperative scores.

Results

One hundred and thirteen patients completed and mailed their post-operative questionnaires addressing the recalled pre-operative QOL for a response rate of 82.2%. Median time between pre-operative assessment and post-operative recall assessment was 14.29 weeks (range 12.42-46.00). Recall validity was excellent for the IIQ-30 ($r=0.64$) and moderate for the SF-12 ($r= 0.46$ (physical component summary or PCS) and 0.42 (mental component summary or MCS)). Recall reliability was moderate with the IIQ-30 (ICC= 0.62) and poor with the SF-12 (ICC= 0.44 (PCS) and 0.49 (MCS)).

Conclusion

The IIQ-30 can be reliably used in a retrospective manner among women who have undergone surgery for SUI 3 months earlier. The SF-12 cannot be reliably used retrospectively.

Key words: retrospective, recall, quality of life, stress urinary incontinence, TVT, TOT

Introduction

Urinary incontinence is a prevalent problem affecting 15 to 35 percent of adult American women with otherwise normal, productive lives (1). The personal impact of urinary incontinence may be enough to compromise an individual's daily living, social interactions and self-perception (1, 32, 40). Over 200 operative techniques have been described for the treatment of stress urinary incontinence (SUI) (2). The search for an effective and minimally-invasive solution led to development of the tension-free vaginal tape (TVT) and trans-obturator tape (TOT) procedures (3, 4, 5). These techniques consist of prolene sling placement under the mid-urethra by the vaginal approach with introducers. The introducers are inserted through the retropubic space for the TVT procedure or through the obturator foramen for the TOT procedure. Little is known about the influence of these new surgical approaches on patients' quality of life (QOL).

Traditionally, treatment effectiveness was solely evaluated in relation to its physiological or biological impact. In recent years, interest has been growing in the assessment of the impact of these therapies on QOL. Specific instruments have been developed to evaluate QOL among women affected by urinary incontinence such as the Incontinence Impact Questionnaire-30 (IIQ-30) (9). These instruments are constructed for use in a prospective manner where the questionnaire is administered before treatment and again at a determined time interval after treatment. The impact on QOL is estimated by subtracting the post-treatment QOL score from the pre-treatment QOL score.

A more rapid, accessible and practical QOL approach to assessment could be retrospective where questionnaires are administered only during the post-treatment period. At that time, 2 copies of the same questionnaires are administered to the patient: one copy measures post-treatment QOL and the second measures recall of pre-treatment QOL. In this way, it is possible to estimate the impact of treatment on QOL by calculating the difference between the post-treatment QOL score and the recall score. While the approach would be more efficient in terms of time and costs, the validity of such a retrospective design depends on the patients' memory and, consequently, recall accuracy.

The accuracy with which health-related quality of life (HRQOL) is recalled has been questioned. Numerous biases have been described that can affect the reliability of

retrospective QOL or health condition assessment (41, 42, 43, 48, 49, 50, 52). Recall accuracy depends on multiple factors including study design, time delay, patient characteristics such as age and level of education, and disease characteristics such as severity, stability and concrete aspects of the disease. While some authors have reported good recall accuracy among patients with prostate cancer (44, 49), or among patients who have undergone knee surgery (52), others have discerned inadequate recall among patients with prostate cancer (43) or after total hip replacement (42). Viktrup and Lose (50) evaluated the ability of 305 primiparas to recall the onset of SUI, 5 years after their first delivery. Forty-nine percent of women recalled the onset of SUI inaccurately. At present, the accuracy of retrospective QOL assessment among surgically-treated incontinent women is unknown.

Considering the importance of improving research resource allocation and lessening patient burden, the primary objective of this study was to ascertain the women's ability to recall their pre-treatment QOL after a TVT or TOT procedure. The secondary objective was to validate the IIQ-30 in our population.

Patients & methods

Design & population

Between August 2005 and December 2007, we conducted a prospective cohort study of female patients scheduled for TVT or TOT by their gynecologist or their urologist at the Centre hospitalier universitaire de Montréal (CHUM).

The inclusion criteria were: 1) ability to speak, read and write French and 2) scheduled for a TVT or TOT procedure to treat SUI. Exclusion criteria were: 1) a cognitive disorder that could affect patients' ability to complete the questionnaires; 2) concomitant genital prolapse surgery; 3) presence of an active psychiatric disorder; 4) any severe, chronic disorder that could influence QOL (dialysis in chronic renal failure, chronic heart failure, chronic obstructive pulmonary disease); 5) anal incontinence; 6) a major handicap (hemiplegia, quadriplegia, blindness, deafness); 7) a neurological disorder that could possibly have an impact on bladder function (multiple sclerosis, Parkinson's disease, spinal trauma, cerebral tumour, stroke, dementia).

Initially, we intended to exclude patients with major comorbidities because they could influence QOL. The socio-demographic questionnaire contained a specific question about a past or present co-morbidity experience such as diabetes, myocardial infarction, stroke, amputation, leg circulation problems, asthma, stomach ulcer, kidney disease, major depression, epilepsy, alcoholism and drug abuse; however, the questionnaire did not effectively discriminate between patients who had a past co-morbidity and those with a present co-morbidity. Forty-five percent of subjects had a past or present history of one of these co-morbidities. All of them were included in the analysis because this variable did not have an impact on recall according to linear regression analysis.

Another group of subjects was recruited to fulfill the secondary objective which was to revalidate the IIQ-30 questionnaire. To establish construct validity by the “known-group” method, a group of continent patients was required. These subjects were enrolled among healthy patients coming for their annual gynecologic visit to the CHUM.

Recruitment & data collection

Patients scheduled for either a TVT or TOT procedure were first identified by a research nurse at the admission office. The subjects were invited by the research nurse to participate at the time of their pre-operative assessment. All patients completed 3 questionnaires: a socio-demographic data questionnaire, followed by the IIQ-30 and Short-Form Health Survey (SF-12).

Then, a few days before the three-month post-op date, 2 copies of the IIQ-30 and 2 copies of the SF-12 were mailed to the participants with a cover letter and a pre-addressed stamped return envelope. We planned a waiting period of 12 weeks between surgery and recall assessment to limit treatment effect biases without being too far from the pre-operative baseline measurement. One copy of each questionnaire evaluated the recall of pre-operative QOL and the other copy of each questionnaire evaluated the new, post-operative QOL. Subjects who did not respond within 2 weeks of the mailing received a phone call reminder. Those who sent incomplete questionnaires were quickly contacted by phone and asked to complete the unanswered questions. A subject was excluded from the study if the questionnaires were not returned within 1 month.

The group of continent subjects completed the IIQ-30 and SF-12 questionnaires after consenting to participate in the study.

The study was approved by the research ethics committee and scientific evaluation committee. All patients gave a written informed consent before enrollment.

Measurement instruments

The IIQ-30 is a self-administered measure developed by Wyman et al. (40) to specifically assess the impact of urinary incontinence on a woman's HRQOL. It has been translated and validated in the French-speaking Canadian population (10). The IIQ is a 4-point Likert-type scale with 30 items. QOL is evaluated in relation to 4 subscales: physical activity (6 items), travel (6 items), social relationships (10 items) and emotional health (8 items). The total score varies from zero to 400. Higher scores reflect a poorer QOL.

The SF-12 (7) is a short form of the SF-36 (6), which is the most-widely used generic instrument. These measures are intended to capture a very broad range of aspects of health status and the consequences of illness and are therefore relevant to a wide range of patients. The SF-12 has 2 scales instead of the 8 found in the SF-36 and 12 items instead of 36. One scale evaluates the physical dimension (physical component summary or PCS) of QOL and the other gauges the psychological dimension (mental component summary or MCS) of QOL. The score of each subscale varies from zero to 100, with higher scores representing better QOL. There is no combination of PCS and MCS scores in a total score. The SF-12 has been translated and validated in the male French-speaking Canadian population (8).

Statistical analysis

Bryant et al. (52) considered the recall score as equivalent to the baseline pre-operative score if the difference was less than 20% of the mean baseline pre-operative score. Based on data reported by Albo et al. (54) for the IIQ-30, we estimated that the number of subjects required to achieve a minimum correlation of 0.5 with a power of 80% would be 79.

The response rate was defined as the ratio of the number of responding subjects to those who consented to participate in the study. Socio-demographic data on responders and non-responders were compared. Change in QOL status after surgery was verified by comparing IIQ pre-operative to post-operative scores by a t test. To evaluate recall accuracy, we first

compared the means of baseline pre-operative scores and recalled pre-operative scores by a t test. Pearson's correlation coefficient (r) was used to assess the validity of the recall scores compared to baseline pre-operative scores for analysis of the IIQ-30 and its 4 subscales, as well as the SF-12. Arbitrarily, we identified the recall scores as the dependent variable (y), and the baseline pre-operative scores as the independent variable (x). The association was considered low if $r < 0.3$, moderate if $r = 0.3-0.45$, good if $r = 0.45-0.6$ and high if $r > 0.6$ (53). The intraclass correlation coefficient (ICC) and its 95% confidence interval (95% CI) assessed the reliability and reproducibility of retrospective QOL assessment (56). Reliability was considered poor if the coefficient was below 0.5, moderate if the value was between 0.5 to 0.75, and good if it was greater than 0.75 (58).

Regression analyses were conducted to examine variations in recall accuracy in relation to patient characteristics and the time interval between pre-operative and the post-operative QOL assessment. Patient characteristics included age, living situation, marital status, employment status, smoking status, education category, income category and co-morbidity. First, univariate linear regression was undertaken to assess the association between the independent variables, defined as patient characteristics and the time interval to recall, and the dependent variable, defined as the difference between pre-operative baseline score and the recall score. Then, multivariate linear regression models tested the effect of significant independent variables on the dependent variable.

The IIQ-30 was subjected to psychometric testing which focused on reliability and validity. No consensus exists on the ideal sample size when validating a questionnaire (31). Validation studies performed to validate the IIQ-30 questionnaire used between 34 and 162 subjects. Our study used the group of incontinent patients to establish item reliability and criterion validity, and we arbitrarily recruited an additional group of 51 continent patients to establish the construct validity.

Cronbach's alpha was quantified as an estimation of item reliability or internal consistency of the item combinations. Values of ≥ 0.70 were considered acceptable. Criterion validity was established by comparing the IIQ measure with the SF-12 measure, considered as the "gold-standard", using Pearson's correlation coefficient. Construct validity was established by the "known-group method". The IIQ questionnaire was administered to a continent group and to an incontinent group of subjects. The continent group sample size was arbitrarily

decided to be 50. The correlation between these 2 groups in relation to IIQ performance was computed by Pearson's correlation coefficient.

For all tests, statistical significance was set at $\alpha=0.05$. All tests were two-sided. The data were analysed by SAS version 8.02 (SAS Institute) except for sample size calculation which was computed by the PASS 2005 program (53).

Results

Of the 152 consenting patients, 39 were excluded. Among the excluded patients, 27 failed to return the questionnaires 3 months after surgery, resulting in a response rate of 82.2%. Five women did not have their surgery within 6 months of recruitment, 5 returned incomplete questionnaires 3 months after their surgery, 1 woman was lost to follow-up by our research team and another woman underwent concomitant prolapse surgery with her TOT procedure. One hundred and thirteen patients were available for final analyses. Among them, 108 subjects had a TOT procedure and 5 had a TVT procedure, reflecting a decision by the urologists early in the study to completely convert to the TOT procedure. All analyses were done again after excluding subjects who had undergone a TVT procedure and no significant difference was detected. Mean age of the respondents was 55.5 years (range 30-77). The variable "time duration between baseline pre-operative assessment and recall assessment" showed a skewed distribution and her median was 14.3 weeks (range 12.4-46).

Most patients were Caucasian, living with a spouse or a partner, and were current smokers. There were certain socio-demographic and baseline QOL mean score differences between respondents and non-respondents (Table I). Non-respondents were more likely to be unmarried, to have a job and to have an income over \$50,000 CDN. IIQ-30 baseline QOL mean scores were slightly lower for non-respondents, with the exception of the traveling subscale, reflecting better pre-operative QOL compared to the respondents. Three months after surgery, the mean post-operative IIQ score was significantly below the pre-operative IIQ score ($p<0.001$) demonstrating an improvement in QOL. The mean difference between the pre-operative score and the post-operative score was 144.1 (187.4-43.3).

The mean of the baseline pre-operative IIQ-30 score was 187.3 (range 171.3-203.2). The mean of the recalled IIQ-30 score was 210.9 (range 192.9-228.9). The difference between the mean baseline score and the mean recall score was 23, representing 12.6% of the mean baseline score. This is well within the 20% equivalence range that was our a priori equivalence definition. The mean of the SF-12 physical scale baseline score was 46.7, and the mean of the recall score was 43.8 for a difference of 3.0 (6.4% of the baseline score), while the mean of the SF-12 mental scale baseline score was 46.5 and the mean of the recall score was 45.1, for a difference of 1.4 (3.1% of the baseline score). The difference between the baseline and recall scores was statistically significant for the SF-12 physical scale ($p=0.02$), but not for the SF-12 mental scale ($p=0.16$).

The strength of the linear relationship between the baseline pre-operative scores and the recall scores was good to excellent for the IIQ-30 (r between 0.58 and 0.64) indicating good recall validity (Table II). The correlation coefficient for each IIQ-30 subscale was over 0.5, but traveling and emotional scales scored higher than 0.6 (Annexe IV). Figure 1 presents linear regression scatter plots for the IIQ-30 total. The association between baseline pre-operative scores and the recalled scores was moderate for the SF-12 mental scale ($r=0.46$) and physical scale ($r=0.42$) suggesting moderate validity (Annexe V). Recall reliability was moderate with the IIQ-30 and poor with the SF-12 (Table III).

Realizing that the time interval between pre-operative assessment and recall assessment was quite variable and might influence recall, Pearson's correlation coefficient and ICC were computed after excluding subjects who responded more than 16 weeks after the pre-operative baseline assessment. The r and ICC were similar whether or not subjects responding more than 16 weeks after recruitment were excluded from the analysis.

Regression analyses were performed to assess the impact of patient characteristics and the time interval to recall on recall accuracy (Table IV). By univariate analysis, age, employment status and income were found to significantly influence recall accuracy when correlated to recall using the IIQ-30. None of the predictive factors examined were significant for the SF-12 (Annexe VI).

To establish a predictive model for the IIQ-30 recall score, education category was incorporated into the multivariate analysis in addition to age, employment status and

income category, because it was close to being a statistically significant predictive factor in univariate analysis (Annexe VII). Few models give R^2 around 0.1000. One model maintained R^2 at 0.1023 with only 2 predictive factors, age and income category, but only age was a statistically significant predictor ($p=0.02$). This model explains 10% of variance of the recall score and was selected as the best predictive model since it explains the same amount of variance while being the most parsimonious. Nevertheless, this model is a poor predictor.

The IIQ-30 showed good psychometric performance in our population. Item reliability was excellent with Cronbach's alpha coefficient of 0.94 for IIQ-30 total, of 0.84 for the physical subscale, of 0.86 for the traveling subscale, of 0.87 for the social subscale, and of 0.87 for the emotional subscale. Criterion validity was tested by Pearson's correlation coefficient comparing IIQ-30 and SF-12 baseline scores. The correlation was poor with an $r=-0.25$ between the IIQ-30 and SF-12 physical scale and an $r=0.10$ between the IIQ-30 and SF-12 mental scale. There was no association between the IIQ-30 and SF-12 results, indicating that these 2 instruments are not evaluating the same QOL dimension: the IIQ-30 is a specific instrument, and the SF-12 is a generic instrument. Construct validity was evaluated by the known-group method. Socio-demographic characteristics of the incontinent and the continent groups are shown in Annexe VIII. The IIQ-30 baseline mean scores of incontinent patients did not correlate well with the IIQ-30 mean scores of continent patients ($r=0.01$), attesting good IIQ-30 construct validity.

Discussion

Urinary incontinence has far-reaching psychosocial effects upon otherwise healthy women (32). This condition makes women anxious, embarrassed and unwilling to undertake a wide range of activities that form major component of QOL. The primary reason surgeons undertake procedures for urinary incontinence is to improve women's QOL. Thus, QOL assessment becomes essential in the evaluation of treatment effectiveness. Traditionally, data are collected prospectively to limit biases; however, this approach is time-consuming and requires additional personnel and financial resources while placing an unnecessary burden on the patients. Post-treatment assessment is more economical and efficient. Our results support the hypothesis and indicate that a retrospective assessment could be a

reasonable substitute for a prospective QOL evaluation when using the IIQ-30 in women surgically treated for SUI. We noticed that the recall score was higher than the preoperative baseline score, however, this difference was still less than 20% of the baseline score which was our pre-established criteria for clinical equivalence. Unfortunately, the minimal clinically important difference has not been identified for IIQ-30 so far. Corcos et al (62) published on IIQ-30 cut-off scores in 2002 but the scoring method employed was not following the one described by original paper on IIQ-30 (9) and consequently did not correspond to our methodology. Pearson's correlation coefficient and the ICC varied between 0.54 and 0.64, confirming IIQ-30 validity and reliability when used retrospectively. The SF-12 showed less recall accuracy with Pearson's correlation coefficient and ICC lower than 0.5.

The accuracy and validity of a retrospective QOL assessment can be affected by sources of recall distortion. In our study, keeping in mind that higher IIQ scores indicate poorer QOL, retrospective assessment of pre-operative QOL tends to suggest lower pre-operative QOL compared to the results obtained by prospective evaluation. The mean IIQ-30 baseline score was 187 and the mean recall score was 210. Thus, the correlation between prospective and retrospective assessment was good but not excellent. Litwin and McGuigan (43) reported a similar observation among men treated for prostate cancer by radical prostatectomy. The phenomenon can be explained by biases affecting the subjects' intellectual processes when trying to remember pre-operative QOL (49). Motivational bias occurs when subjects, who are very sensitive to research activities, try harder to note and report symptoms. Encoding bias occurs with subjects who know and record treatment effects. The cognitive dissonance theory suggests that subjects who are particularly impressed by the treatment experience tend to exaggerate treatment benefits (41). These recall errors are also a reflection of the subjects' belief about their condition and treatment. Their effect would be that a retrospective assessment would give a better QOL improvement following a treatment than a prospective QOL assessment. In the context where we aim to assess the subjects' satisfaction in relation to their treatment, the impact of these biases could be seen as marginal.

The response shift phenomenon can also explain our results, that is, as time progresses and subjects adapt to new impairments, they may experience an altered frame of reference that

distorts recall (46, 47, 63). Subjects may proceed to an unconscious recalibration of their own reference system subsequent to treatment when living through new experiences. The chronic nature of these impairments may also lead them to take on an inexact perspective of their baseline QOL. Some response shift theorists consider the recalled score as the only valid measure of pre-operative status because it is rated according to the same reference system as the post-operative or final outcome rating. This theory is interesting because of the fact that it supports and even favours retrospective measurement.

Some authors have discussed the influence of time on QOL recall accuracy (33, 44). In general, the longer the time frame, the worse is the accuracy of recall. Legler et al (44) reported good recall accuracy of urinary, sexual and bowel functions after a 6 month time-frame with an ICC of 0.83 for urinary functions. To measure the effect of time on recall accuracy, we also examined the effect of excluding subjects who completed the recall questionnaires more than 16 weeks after the pre-operative baseline evaluation. Regression analysis did not reveal any significant effect of time on recall accuracy. In our medical practice, we routinely organize a follow-up visit 6 weeks after surgery. It would have been useful to know the recall accuracy at that point in time, to minimize financial, temporal and personnel resources necessary to study the effectiveness of therapy. We assume that if retrospective QOL assessment is valid 3 months after a TVT or TOT procedure, it is probably more accurate at 6 weeks after surgery.

The characteristics of the medical problem under investigation can affect recall precision. Urinary incontinence is a concrete condition with little ambiguity. Karakiewicz et al (49) suggested that retrospective evaluation of more concrete and quantitative HRQOL domains, such as erectile function, may be less impacted than conceptually more complex domains, such as orgasmic function. The poor performance of the SF-12 as an instrument for retrospective QOL assessment reflects this fact. The SF-12 scale as a generic questionnaire has been developed to evaluate global QOL and, consequently, requires subjects to recall a more general condition implying more abstract concepts.

Subject' characteristics can influence recall accuracy as well. Litwin and McGuigan (43) have already demonstrated that age and education can impact recall accuracy. Our results revealed that age was the only factor that significantly affected recall precision. The older

the patient, the greater was the difference between the baseline pre-operative QOL score and the recall QOL score.

Our study had other limitations. First, other factors, such as urinary incontinence severity, could have influenced recall accuracy but were not measured. For example, to measure the impact of urinary incontinence severity, patients would have to go through more investigations, including a stress test, a pad test and/or urodynamic study. This additional clinical evaluation would have needed more resources and was not our primary objective. The intensity of the urinary incontinence described by usual clinical parameters most likely varied from one subject to another in our population, but was not measured or controlled.

Second, another possible impact on recall accuracy could have been the preop QOL score itself. Patient's perception of distress influence QOL score. Langford CF et al demonstrated that among women with SUI, those having worst pre-treatment QOL were having higher expectations of treatment (64). Is it possible that patient's treatment expectations could influence the recall of preop QOL? And consequently, do the level of preop QOL could be an index of recall accuracy?

Third, the effect of treatment on recall precision is another variable which we would have liked to control. A study design including a control group without treatment would have been helpful. From an ethical and practical point of view, it is difficult to convince patients suffering from urinary incontinence to participate in a study that implies a possibility of not being treated.

To our knowledge, no previous study has evaluated pre-operative QOL recall accuracy among women operated on for SUI. The IIQ-30 questionnaire has demonstrated good psychometric performance for traditional prospective use in our population. The study's first objective was to validate 2 QOL instruments, a generic questionnaire (SF-12) and a specific questionnaire (IIQ-30), when administered retrospectively. The validity and reliability of retrospective assessment with the SF-12 and IIQ-30 were tested by Pearson's correlation coefficient and ICC. Additional analyses have been conducted to identify factors that affect recall. The sample size provided good statistical power. Moreover, the study included subjects with past and/or present co-morbidities, and subjects of different

ages (range 30 to 77 years old) which made the study population representative of the general population and improved external validity.

Conclusion

The IIQ-30 questionnaire, when used retrospectively within 3 months of a TOT/TVT procedure, provides a valid and relatively accurate estimate of pre-operative QOL.

Consequently, this retrospective QOL assessment can substitute for the prospective QOL assessment to study the impact of a therapy. In this way, recruitment effectiveness and data collection can be made easier and improve research resource allocation. The SF-12 instrument is not considered reliable when administered retrospectively.

Chapitre IV : Discussion

L'incontinence urinaire a un impact suffisant pour compromettre les activités sociales et le bien-être psychologique des femmes et par conséquent affecter significativement leur qualité de vie (32). L'amélioration de la qualité de vie des patientes devient la première motivation des chirurgiens traitant des femmes souffrant d'incontinence urinaire à l'effort. Ainsi, l'évaluation de l'impact d'une chirurgie sur la QOL des patientes à l'aide de questionnaires validés est une mesure essentielle à la connaissance de l'efficacité du traitement. La collecte des données en période préopératoire exige un investissement en temps additionnel ainsi que des ressources financières et en personnel qualifié supplémentaires. On demande également aux patientes une attention et une participation additionnelles. Nos résultats supportent que l'on puisse raisonnablement substituer une évaluation prospective par une évaluation rétrospective de l'impact sur la QOL d'une chirurgie traitant l'incontinence urinaire à l'effort. Toutefois, le score moyen du rappel était supérieur à celui du score préopératoire de base. Par contre, cette différence reste inférieure à 20% du score préopératoire de base et reste donc cliniquement équivalent selon notre critère pré-établi. Malheureusement, il n'y a pas d'intervalle de score minimal décrit dans la littérature pouvant traduire une différence clinique significative pour le IIQ-30. Corcos et al (62) ont publié sur les *cut-off scores* du IIQ-30 en 2002. Cependant, la méthode de cotation utilisée par les auteurs n'a pas suivi celle recommandée par l'article de référence du IIQ-30 (9) et donc, leurs résultats ne peuvent pas être utilisés pour mieux interpréter nos résultats. Des coefficients de corrélation de Pearson supérieurs à 0,5 et des coefficients de corrélation intraclasse variant entre 0,54 et 0,62 indiquent que le rappel de la QOL préopératoire est suffisamment valide et fiable en utilisant l'instrument IIQ-30. Par contre, l'exactitude du rappel de la QOL préopératoire est plus faible avec l'instrument SF-12 qui démontre des coefficients de corrélation de Pearson et coefficients de corrélation intraclasse inférieurs à 0,5.

Des sources de distorsion de l'information peuvent affecter la mémoire du sujet, l'exactitude du rappel et la validité d'une évaluation rétrospective de la qualité de vie. Nous nous retrouvons face à une évaluation rétrospective de la QOL préopératoire qui tend à rapporter une QOL inférieure à celle évaluée de façon prospective. Le score préopératoire de base du IIQ-30 était à 187 et celui du rappel à 210. Avec l'instrument IIQ-30, plus le

score est élevé, pire est la qualité de vie. Malgré ce fait, la corrélation entre l'évaluation prospective et rétrospective était bonne mais pas excellente. Litwin et McGuigan (43) ont rapporté une observation similaire chez les patients opérés pour une prostatectomie radicale. Ce phénomène est possiblement causé par des biais affectant le processus intellectuel utilisé inconsciemment par les patientes lorsqu'elles essaient de se rappeler leur QOL préopératoire. Parmi ces biais, le biais motivationnel survient lorsque des sujets très sensibles à l'activité de la recherche essaient de se souvenir plus intensément des symptômes reliés au traitement. Le biais *encoding* où les sujets connaissent les effets du traitement et les enregistrent, et le biais de l'effet du traitement (ou de la dissonance cognitive) où les sujets sont particulièrement affectés par l'expérience du traitement peuvent aussi expliquer en partie ces résultats (41). Ces distorsions de l'information sont le reflet des croyances du sujet sur sa condition et son traitement. Leurs effets seront de rapporter une amélioration de la QOL plus importante que si l'évaluation de la QOL avait été réalisée de façon prospective. Dans la mesure où notre évaluation cherche à connaître la satisfaction du sujet par rapport à son traitement, l'impact de ces biais sur la valeur d'une évaluation rétrospective peut être perçu comme étant marginale.

Le phénomène du *response shift* peut aussi expliquer nos résultats (46, 47, 63). Il s'agit de l'effet du temps et de l'adaptation du sujet à ces incapacités qui modifient le système de référence du sujet et la perception de son état. Le sujet procède à une espèce de réajustement inconscient de son échelle de références lorsqu'il est plongé dans des expériences plus récentes. Les incapacités chroniques peuvent aussi conduire les sujets à adopter une perspective « décalée » de leur qualité de vie préopératoire. Les théoriciens du *response shift* considèrent que le rappel de la condition préopératoire est la seule mesure valide car elle est mesurée dans le même système de référence que l'évaluation du résultat postopératoire. Cette théorie est intéressante du fait qu'elle soutient et même favorise une évaluation rétrospective de l'impact d'une thérapie sur la qualité de vie même si elle n'est pas équivalente à une évaluation prospective.

La période de temps entre le moment de l'évaluation de la QOL préopératoire et le moment de l'évaluation du rappel de la QOL préopératoire a varié entre 12,4 et 46 semaines, pour une médiane de 14,3 semaines. Cette donnée a été présentée sous forme de médiane étant donné que sa distribution ne suivait pas une loi normale. Un délai de trois mois entre la

chirurgie et l'évaluation du rappel permettait de limiter le biais de l'effet du traitement sans trop s'éloigner du moment préopératoire. Plusieurs auteurs ont décrit l'influence du temps sur la qualité du rappel (33, 44). Plus le délai est long, plus le rappel deviendra inexact. Legler (44) a démontré un rappel adéquat des fonctions urinaire, sexuelle et intestinale suite à une période de 6 mois avec un coefficient de corrélation intraclasse de 0,83 pour la fonction urinaire. Pour évaluer l'effet du temps sur la qualité du rappel, une deuxième analyse des données a été réalisée en excluant les sujets ayant eu un délai de plus de 16 semaines entre l'évaluation préopératoire initiale et l'évaluation du rappel. L'analyse de régression linéaire n'a pas démontré une influence significative du temps sur la précision du rappel. Dans notre pratique médicale habituelle, nous assurons un suivi postopératoire six semaines après une chirurgie. Il aurait été utile de connaître l'exactitude du rappel à ce moment dans la perspective où l'on recherche toujours à minimiser les besoins financier, temporel et personnel nécessaires à l'étude de l'efficacité d'une thérapie. Nous pouvons supposer que si le rappel de la QOL préopératoire était adéquate trois mois après la chirurgie, il est fort probable que le rappel évalué à six semaines serait d'une validité équivalente ou même supérieure.

Les caractéristiques de la maladie étudiée influencent aussi l'exactitude du rappel. L'incontinence urinaire à l'effort est une condition concrète où il n'y a pas d'ambiguïté. Karakiewicz (49) suggère que l'évaluation de condition plus concrète telle que la fonction érectile est plus fiable que l'évaluation rétrospective de fonction plus abstraite telle que la fonction orgasmique. Cependant, un sujet ne peut pas donner un rappel adéquat s'il ne comprend pas le ou les symptômes visés par la question. Des travaux de recherche en psychologie ont démontré que les questions spécifiques étaient répondues avec beaucoup plus de validité et de fiabilité que les questions d'ordre général (49). La faible performance du SF-12 en tant qu'outil d'évaluation rétrospective de la qualité de vie reflète ce fait. L'instrument SF-12, étant un questionnaire générique développé pour évaluer la qualité de vie globale, exige des sujets de se rappeler une condition plus générale, plus complexe et impliquant davantage des concepts abstraits.

Les caractéristiques des sujets peuvent également influencer les résultats. Litwin et McGuigan (43) ont déjà démontré que l'âge et l'éducation pouvaient affecter l'exactitude du rappel. De notre côté, l'âge était le seul facteur affectant significativement le rappel.

Plus une patiente était âgée, plus elle avait tendance à se rappeler une qualité de vie inférieure à celle qu'elle avait en réalité.

Parmi les limites de l'étude, il faut mentionner que d'autres facteurs ont pu influencer l'exactitude du rappel mais non pas pu être mesurés tels que la sévérité de l'incontinence urinaire. Pour mesurer l'impact de la sévérité de l'incontinence urinaire, il aurait fallu examiner les patientes, procéder à un *stress test*, à un *pads test* et/ou à un bilan urodynamique. Cette évaluation clinique additionnelle aurait demandé beaucoup plus de ressources et l'évaluation des facteurs affectant la qualité du rappel n'était le premier objectif de l'étude. Malgré tout, l'intensité du problème d'incontinence urinaire de notre population devait varier d'un sujet à l'autre mais cette variable n'a pas été mesurée et donc n'a pas pu être contrôlée.

Ensuite, il se pourrait que le niveau du score de QOL pré-opératoire ait pu influencer l'exactitude du rappel. La perception individuelle de la détresse peut influencer le score de QOL. Langford CF et al ont démontré que parmi les femmes souffrant d'incontinence urinaire à l'effort, celles qui rapportaient un score de QOL inférieur avaient des attentes plus élevées par rapport au succès de leur traitement (64) (Langford CF, Elmissiry MM, Ghoniem GM. Do women have realistic expectations of treatment for stress urinary incontinence? *Neurourol Urodynam* 2008; 27: 480-4.). Suite à cette réflexion, se peut-il que le niveau d'attente face au succès d'un traitement puisse influencer le rappel de la qualité de vie préopératoire? Par conséquent, est-ce que le score de QOL préopératoire pourrait devenir un facteur affectant l'exactitude du rappel?

L'effet du traitement sur la qualité du rappel est également une variable que l'on aurait pu vouloir contrôler. Il aurait fallu utiliser un devis avec groupe contrôle sans traitement. En pratique, il est difficile de convaincre des patientes souffrant d'incontinence urinaire à l'effort de participer à une étude où il est possible qu'elles ne puissent pas pouvoir bénéficier du traitement désiré.

À notre connaissance, aucune étude n'a évalué l'exactitude du rappel de la QOL préopératoire chez les femmes opérées pour une incontinence urinaire à l'effort. D'abord, l'instrument IIQ-30 a été validé pour une utilisation prospective traditionnelle et il a démontré des propriétés psychométriques satisfaisantes attestant sa validité auprès de la

population à l'étude. Cette étude se voulait principalement une étude de validation de deux instruments, un questionnaire générique (SF-12) et un questionnaire spécifique (IIQ-30), lorsqu'ils sont utilisés de façon rétrospective. L'objectif premier n'était pas de revalider un nouvel instrument mais plutôt de valider une nouvelle utilisation d'un instrument déjà validé. Ainsi, la validité et la fiabilité des questionnaires IIQ-30 et SF-12 utilisés rétrospectivement ont été testées rigoureusement avec des analyses de corrélation de Pearson et des analyses de corrélation intraclasse. Des analyses supplémentaires ont été effectuées afin d'identifier et de contrôler les facteurs confondants. Cette étude comporte un nombre important de sujets lui donnant une puissance statistique. Par ailleurs, l'étude a inclus des sujets avec comorbidités actuelles et/ou antérieures, et des sujets d'âge variable (30-70 ans). Ainsi, la population étudiée fût une meilleure représentation de la population générale et a permis d'améliorer la validité externe de l'étude.

Conclusion

L'instrument IIQ-30, lorsqu'administré rétrospectivement trois mois après une chirurgie de TOT ou de TVT, peut donner une évaluation valide et fiable de la qualité de vie préopératoire. Par conséquent, il est possible d'améliorer l'efficacité du recrutement et de la collecte des données dans l'étude de l'impact des thérapies sur la qualité de vie des sujets. L'instrument SF-12 ne peut être utilisé rétrospectivement de façon fiable.

Table I : Sociodemographic characteristics and baseline QOL scores between respondents and non-respondents to the recall survey

Patient characteristics	Respondents (n=113)	Non-respondents (n=27)
Age at baseline (years)		
Mean	55.6	51.6
Range	(30-77)	(36-77)
Race		
White	110 (97.4%)	26 (96.3%)
Marital status		
Married	53 (46.9%)	8 (29.6%)
Living situation		
Living with a spouse or a partner	85 (75.2%)	21 (77.8%)
Smoking status		
Current smoker	68 (60.2%)	17 (63.0%)
Education		
Did not complete high school	34 (30.1%)	11 (40.7%)
Completed high school	26 (23.0%)	5 (18.5%)
Completed college or university	53 (46.9%)	10 (37.0%)
Employment situation		
Full or part time job	73 (64.6%)	23 (85.2%)
Income*		
Less than \$20,000	18 (15.9%)	4 (14.8%)
Between \$20,000 – 50,000	33 (29.2%)	7 (25.9%)
More than \$50,000	45 (39.8%)	14 (51.8%)
Co-morbidities**	51 (45.1%)	14 (51.8%)
IIQ baseline scores (mean)		
Total	187.4	171.8
Physical	57.1	51.0
Traveling	48.7	48.6
Social	40.2	35.3
Emotional	41.5	36.9
SF-12 baseline scores (mean)		
PCS	46.6	48.4
MCS	46.4	42.7

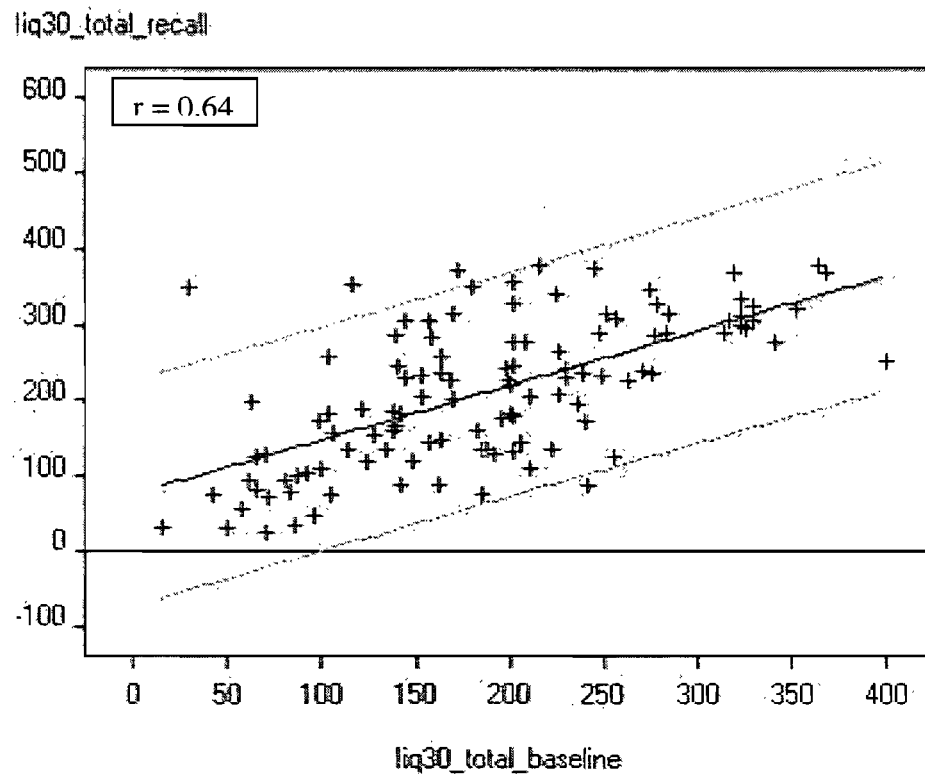
* 17 (15.0%) patients did not respond to this question

** Subjects with one or another past or present co-morbidity: diabetes, myocardial infarction, stroke, amputation, leg circulation problems, asthma, stomach ulcer, kidney disease, major depression, epilepsy, alcoholism and toxicomania

Table II : Correlation between recalled QOL scores and baseline pre-operative QOL scores

Questionnaires	Pearson's <i>r</i>
IIQ-30	
Total	0.64
Physical	0.54
Traveling	0.63
Social	0.58
Emotional	0.63
SF-12	
PCS	0.46
MCS	0.42

Figure 1 : Scatterplots of patients' recall QOL scores compared to baseline pre-operative QOL scores for the IIQ total



r: Pearson's correlation coefficient

Table III : Reliability of using recalled QOL scores compared to baseline pre-operative QOL scores

Questionnaires	ICC (95% CI)
IIQ-30	
Total	0.63 (0.48 to 0.71)
Physical	0.55 (0.39 to 0.66)
Traveling	0.62 (0.47 to 0.71)
Social	0.58 (0.41 to 0.67)
Emotional	0.63 (0.48 to 0.71)
SF-12	
PCS	0.46 (0.27 to 0.57)
MCS	0.48 (0.32 to 0.61)

ICC: Intraclass correlation coefficient

PCS: Physical component summary

MCS: Mental component summary

Table IV : Univariate linear regression where the difference between baseline IIQ-30 score and recalled score is the dependant variable and patient characteristics are the independent variables for the IIQ-30

IIQ-30	β	Standard error (SE)	p-value	R ²
• Age	1.29	0.48	0.0081	0.0614
• Living situation	-2.30	11.36	0.8396	0.0004
• Marital Status	-8.70	10.49	0.4090	0.0061
• Employment status	-22.28	10.78	0.0411	0.0371
• Smoking status	9.46	10.69	0.3780	0.0070
• Education category	-10.23	6.02	0.0922	0.0253
• Income category	-16.50	7.43	0.0288	0.0499
• Co-morbidity	1.53	10.55	0.8846	0.0002
• Duration to recall	-0.38	1.15	0.7442	0.0010

Bibliographie

1. Klutke JJ, Klutke CG. *Tension Free Vaginal Tape Procedure*. AUA Update Series 2003, Lesson **10**; 22 : 74-80.
2. Wall LL. "Urinary stress incontinence" dans *TeLinde's operative gynecology*, 8^e éd., Rock JA, Jones HW, Philadelphie, Lippincott-Raven, chapter 36, p.1087-134, 1997.
3. Ulmsten U, Henriksson L, Johnson P, Varhos G. An ambulatory surgical procedure under local anesthesia for treatment of female urinary incontinence. *Int Urogynecol J* 1996; 7: 81-6.
4. Delorme E. Transobturator urethral suspension: mini-invasive procedure in the treatment of stress urinary incontinence in women. *Prog Urol* 2001; 11(6): 1306-13.
5. Delorme E, Droupy S, deTayrac R, Delmas V. La bandelette trans-obturatrice (Uratape®). Un nouveau procédé mini-invasif de traitement de l'incontinence urinaire de la femme. *Prog Urol* 2003; 13 : 656-9.
6. Ware JEJ, Sherbourne CD, The MOS 36 Short Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30: 473-83.
7. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; 34(3): 220-233.
8. Lebeau T, Perrotte P, Valiquette L, Bénard F, McCormack M, Saad F, Karakiewicz PI. Validation of prostate cancer index and SF-12 short forms. *Can J Urol* 2005; 12 (6): 2873-9.
9. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fand JA for the Continence Program in Women (CPW) Research Group. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. *Quality of Life Research* 1994; 3: 291-306.
10. Beaulieu S, Collet JP, Tu LM, Macrammalla E, Wood-Dauphinee S, Corcos J. Performance of the Incontinence Impact Questionnaire in Canada. *Can J Urol* 1999; 6(1) : 692-9.
11. Bates CP, Bradley WE, Glen ES, & coll.. The standardization of terminology of lower urinary tract function. *J Urol* 1979; 121: 551.
12. Brocklehurst JC. Urinary incontinence in the community-analysis of a MORI poll. *BMJ* 1993; 306: 832-4.
13. Angus Reid poll. *Urinary Incontinence in the Canadian adult population*. Sondage réalisé par la Fondation d'aide pour les personnes incontinentes (Canada). Toronto : Ont. : Angus Reid Group; 1997.
14. Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmstem U et al. The standardization of terminology of lower urinary tract function. Report from the Standardization Subcommittee of the International Continence Society. *Neurourol Urodynam* 2002; 21 : 167-78.

15. Wagner TH, Hu TW. Economic costs of urinary incontinence. *Urology* 1998; 51: 355-356.
16. Robert M, Ross S, Farrell SA, Easton WA, Epp A, Girouard L et al. Conservative management of urinary incontinence. *J Obstet Gynaecol Can* 2006; 28(12) : 1113-8.
17. Robert M, Farrell SA, Easton WA, Epp A, Flood CG, Girouard L, Lajoie F, MacMillan JB, Mainprize TC; Society of Obstetricians and Gynaecologists of Canada. Choice of surgery for stress incontinence. *J Obstet Gyneacol Can* 2005; 27(10): 964-80.
18. Lapitan MC, Cody DJ, Grant AM. Open retropubic colposuspension for urinary incontinence in women. *Cochrane Database Syst Rev* 2005 ; (3) : CD002912.
19. Alcalay M, Monga A, Stanton SL. Burch colposuspension: 10 – 20 year follow-up. *Br J Obstet Gyneacol* 1995; 102 (9): 740-5.
20. Nilsson CG, Kuuva N, Falconer C, Rezapour M, Ulmsted U. Long-term Results of the Tension-Free Vaginal Tape (TVT) Procedure for Surgical Treatment of Female Stress Urinary Incontinence. *Int Urogynecol J* 2001; Suppl. 2: S5-S8.
21. DeTayrac R, Deffieux X, Droupy S, Chauveaud-Lambling A, Calvanèse-Benamour L, Fernandez H. A prospective randomized trial comparing tension-free vaginal tape and transobturator suburethral tape for surgical treatment of stress urinary incontinence. *Am J Obstet Gynecol* 2004; 190: 602-8.
22. Nilsson CG, Falconer C, Rezapour M. Seven-year follow-up of the tension-free vaginal tape procedure for treatment of stress urinary incontinence. *Obstet Gynecol* 2004; 104(6): 1259-62.
23. Giberti C, Gallo F, Cortese P, Schenone M. Transobturator tape for treatment of female stress urinary incontinence: objective and subjective results after a mean follow-up of two years. *Urology* 2007; 69(4):703-7.
24. Felce D, Perry J. Quality-of-life: Its Definition and Measurement. *Res Dev Disabil* 1995; 16(1): 51-74.
25. World Health Organisation (1947) Constitution. *WHO Chronicle 1*. World Health Organization, Geneva.
26. Moons P, Budts W, De Geest S. Critique on the conceptualisation of quality of life: A review and evaluation of different conceptual approaches. *Int J Nurs Stud* 2006; 43(7):891-901.
27. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; 2(14): 1-74.
28. Browne JP. Health-related quality-of-life studies in urology: conceptual and methodological considerations. *World J Urol* 1999; 17: 193-8.
29. Spector PE. "Validation." "Reliability and Norms". *Summated rating scale construction: an introduction*. Lewis-Beck MS, editor, Newbury Park, Sage Publications, chapters 6-7, 46-69, 1992.
30. Contandriopoulos AP, Champagne F, Potvin L, Denis JL, Boyle P. « Choix d'une stratégie de recherche ». « Planification opérationnelle de la recherche ». *Savoir préparer une recherche. La définir, la structurer, la financer*. 1^{ère} éd.. Montréal, Les presses de l'Université de Montréal, chapitres 2-3, 33-90, 1990.

31. DeVellis RF. "Validity". *Scale Development: Theory and Applications*. 1st ed.. Thousand Oaks, Sage Publications, Seawell MH, Wiley K, Hoffman CA, Robinson J, editors. Volume 26, chapter 4, 49-59, 2003.
32. Norton C. The effects of urinary incontinence in women. *Int Rehab Med* 1982; 4: 9.
33. Fedorkow DM. *Prevalence of Urinary Incontinence, Pelvic Organ Prolapse and Anal Incontinence in Women*. Female Pelvic Medicine and Reconstructive Pelvic Surgery. 1st ed. Drutz HP, Herschorn S, Diamant NE. London, Springer, Chapter 2, 11-22, 2003.
34. Bonniaud V, Raibaut P, Guyatt G, Amarenco G, Parratte B. Symptom and quality of life assessment in urinary disorders. *Ann Readapt Med Phys* 2005; 48(6): 392-403.
35. Donovan JL, Badia X, J. Corcos J, Gotoh M, Kelleher CJ and Naughton M, Symptom and quality of life assessment. In: P. Abrams, L. Cardozo, S. Khoury and A. Wein, Editors, *Incontinence: Second International Consultation Proceedings*. Plymouth: Health Publication (2002), pp. 267-316.
36. Amarenco G, Marquis P, Leriche A, Richard F, Zerbib F, Jacquetin B. Une échelle spécifique d'évaluation de la perturbation de la qualité de vie au cours des troubles mictionnels : l'échelle Ditrovie. *Ann Readapt Med Phys* 1997; 40: 21-6.
37. Amarenco G, Arnould B, Carita P, Haab F, Labat JJ, Richard F. European psychometric validation of the CONTILIFE ((R)): A quality of life questionnaire for urinary incontinence. *Eur Urol* 2003; 43(4): 391-404.
38. Wagner TH, Patrick DL, Bavendam TG, Martin ML, Buesching DP. Quality of life of persons with urinary incontinence; development of a new measure. *Urology* 1996; 47: 67-71 (discussion 71-1).
39. Costa P, Perrouin-Verbe B, Colvez A, Didier JP, Marquis P, Marrel A, et al. Quality of life in spinal cord injury patients with urinary difficulties: development of validation of qualiveen. *Euro Urol* 2001; 39: 107-13.
40. Wyman JF, Harkins SW, Choi SC, Taylor JR, Fantl JA. Psychosocial impact of urinary incontinence in women. *Obstet Gynecol* 1987; 70: 378-381.
41. Aseltine RH, Carlson KJ, Fowler FJ, Barry MJ. Comparing Prospective and Retrospective Measures of treatment Outcomes. *Med Care* 1995; 33 (suppl): AS67-76.
42. Mancuso CA, Charlson ME. Does recollection error threaten the validity of cross-sectional studies of effectiveness? *Med Care* 1995; 33 (suppl): AS77-88.
43. Litwin MS, McGuigan K. Accuracy of recall in health-related quality-of-life assessment among men treated for prostate cancer. *J Clin Oncol* 2000; 17: 2882-88.
44. Legler J, Potosky AL, Gilliland FD, Eley JW, Stanford JL. Validation Study of Retrospective Recall of Disease-Targeted Function: Results from the Prostate Cancer Outcomes Study. *Med Care* 2000; 38: 847-57.
45. Potosky AL, Harlan LC, Stanford JL, Gilliland FD, Hamilton AS, Albertsen PC et al. Prostate cancer practice patterns and quality of life : The Prostate Cancer Outcome Study. *J Natl Cancer Inst* 1999; 91: 1719-24.

46. Howard GS, Ralph KM, Gulanic NA, Maxwell SE, Nance SW, Gerber SK. Internal validity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Appl Psychol Meas* 1979; 3 : 1-23.
47. Sprangers MA. Response-shift bias : A challenge to the assessment of patients' quality of life in cancer clinical trials. *Cancer Treat Rev* 1996; 22 (suppl A): 55-62.
48. Herrmann D. Reporting current, past and changed health status: what we know about distortion. *Med Care* 1995; 33 (suppl.): AS89-94.
49. Karakiewicz P, Shariat SF, Naderi A, Kadmon D, Slawin KM. Reliability of remembered international index of erectile function domain in men with localized prostate cancer. *Urology* 2005; 65: 131-5.
50. Viktrup L, Lose G. Do fertile women remember the onset of stress incontinence? *Acta Obstet Gynecol Scand* 2001; 80: 952-5.
51. Wilcox AJ, Horney LF. Accuracy of spontaneous abortion recall. *Am J Epidemiol* 1984; 120:727-33.
52. Bryant D, Norman G, Stratford P, Marx RG, Walter SD, Guyatt G. Patients undergoing knee surgery provided accurate ratings of preoperative quality of life and function 2 weeks after surgery. *J Clin Epidemiol* 2006; 59: 984-93.
53. Blackwelder WC. "Equivalence Trials". *Encyclopedia of Biostatistics*, New York, John Wiley and Sons. Armitage P, Colton T, editors. Volume 2, 1367-1372, 2001.
54. Albo M, Wruck L, Baker J, Brubaker L, Chai T, Dandreo KJ, Diokno A, Kraus S, Kusek JW, Lemack G, Lowder J, Steers W. The Relationships Among Measures of Incontinence Severity in Women Undergoing Surgery for Stress Urinary Incontinence. *J Urol* 2007; 177:1810-14.
55. Burnand B, Kernan WN, Feinstein AR. Indexes and boundaries "quantitative significance" in statistical decisions. *J Clin Epidemiol* 1990; 43: 1273-84.
56. Portnoy LG, Watkins MP. "Statistical measures of reliability". *Foundation of clinical research : Application to practice*. Norwall, Appleton & Lange, chapter 26, 505-528, 1993.
57. <http://web.missouri.edu/~ucasadorg/icc text.htm> since August 14, 2000.
58. Portnoy LG, Watkins MP. "Reliability". *Foundation of clinical research : Application to practice*. Norwall, Appleton & Lange, chapter 5, 53-67, 1993.
59. Tinsley HEA, Tinsley DJ. Uses of Factor Analysis in Counseling Psychology Research. *Journal of Counseling Psychology* 1987; 34: 414-24.
60. Nunnally JC. *Psychometric Theory* 1978, 2nd ed. New York: McGraw-Hill.
61. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Medecine* 2000; 30: 1-15.
62. Corcos J, Behluoli H, Beaulieu S. Identifying cut-off scores with neural networks for interpretation of the incontinence impact questionnaire. *Neurourol Urodynam* 2002; 21: 198-203.
63. Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality of life research. *Sco Sci Med* 1999; 48: 1531-48.

64. Langford CF, Elmissiry MM, Ghoniem GM. Do women have realistic expectations of treatment for stress urinary incontinence? *Neurourol Urodynam* 2008; 27: 480-4.

Annexe I : Questionnaire sur les données socio-démographiques

S'il vous plaît, madame, veuillez répondre aux questions suivantes concernant vos caractéristiques socio-démographiques personnelles

- | | |
|--|--|
| <p>1. Quel âge aviez-vous lors de votre dernier anniversaire ? ____ ans</p> <p>2. Que considérez-vous être votre groupe ethnique ou culture d'origine ?</p> <p>Blanc, caucasien.....1</p> <p>Noir, Afro-américain.....2</p> <p>Latino, Hispanique.....3</p> <p>Asiatique, Oriental, Insulaire du Pacifique.....4</p> <p>Amérindien, Autochtone.....5</p> <p>Autre, précisez _____.....6</p> <p>3. Quelle est votre préférence linguistique ?</p> <p>Français.....1</p> <p>Anglais.....2</p> <p>Autre, précisez _____.....3</p> <p>4. Quel énoncé décrit le mieux votre situation courante ?</p> <p>Vie commune avec épouse ou partenaire.....1</p> <p>Relation significative mais sans vie commune.....2</p> <p>Pas de relation significative.....3</p> <p>5. Quel est votre état civil actuel ?</p> <p>Jamais marié.....1</p> <p>Marié.....2</p> <p>Séparé.....3</p> <p>Divorcé.....4</p> <p>Veuf.....5</p> | <p>6. Avez-vous un emploi rémunéré ?</p> <p>Oui, à temps plein..... 1</p> <p>Oui, à temps partiel.....2</p> <p>Non, mais je suis à la recherche d'emploi.....3</p> <p>Non, je suis à la retraite.....4</p> <p>7. Avez-vous déjà fumé des cigarettes ?</p> <p>1 Oui 2 Non</p> <p>Si oui, combien de cigarettes par jour fumez ou fumiez-vous ? ____ cigarettes</p> <p>Si oui, pendant combien d'années avez-vous fumé ? ____ années</p> <p>8. Quel est votre niveau de formation ?</p> <p>Niveau primaire ou moins..... 1</p> <p>Études secondaires ou techniques..... 2</p> <p>Diplômé d'études secondaires ou techniques.....3</p> <p>Études collégiales.....4</p> <p>Diplômé d'études collégiales.....5</p> <p>Diplômé d'études universitaires / professionnelle.....6</p> <p>9. Quel est votre revenu combiné approximatif de votre ménage ?</p> <p>0-10,000 \$..... 1</p> <p>10,001 – 20,000 \$.....2</p> <p>20,001 – 30,000 \$.....3</p> <p>30,001 – 50,000 \$.....4</p> <p>50,001 – 75,000 \$.....5</p> <p>Plus de 75,000 \$.....6</p> |
|--|--|

10. Avez-vous été atteint ou êtes atteint d'une des maladies suivantes ?

	Non	Oui
a) Diabète.....	1	2
b) Infarctus, douleur de poitrine.....	1	2
c) Accident cérébro- vasculaire	1	2
d) Amputation	1	2
e) Trouble de circulation, troubles respiratoires	1	2
f) Asthme, emphysème, troubles respiratoires	1	2
g) Ulcère d'estomac	1	2
h) Maladie rénale.....	1	2
i) Dépression majeure	1	2
j) Épilepsie	1	2
k) Alcoolisme ou problème d'alcool	1	2
l) Toxicomanie.....	1	2

Annexe II IIQ-30

Veillez inscrire la date d'aujourd'hui (jour/mois/année), s.v.p. : __/__/____

Les pertes d'urine ou le prolapsus de la vessie vous nuisent-ils dans les activités suivantes?

1. Vos tâches ménagères (la cuisine, le ménage, le lavage)?
 pas du tout un peu modérément beaucoup
2. L'entretien et les réparations autour et dans la maison?
 pas du tout un peu modérément beaucoup
3. Le magasinage?
 pas du tout un peu modérément beaucoup
4. Vos loisirs et vos passe-temps?
 pas du tout un peu modérément beaucoup
5. Vos activités physiques telle la marche, la natation, ou une autre forme d'exercice?
 pas du tout un peu modérément beaucoup
6. Vos sorties (cinéma, concerts, etc.)?
 pas du tout un peu modérément beaucoup
7. Vos déplacements de moins de **20 minutes** en voiture ou en autobus?
 pas du tout un peu modérément beaucoup
8. Vos déplacements de plus de **20 minutes** en voiture ou en autobus?
 pas du tout un peu modérément beaucoup
9. Vous rendre quelque part où vous n'êtes pas certaines de trouver des toilettes accessibles?
 pas du tout un peu modérément beaucoup
10. Partir en vacances?
 pas du tout un peu modérément beaucoup
11. Aller à l'église ou dans un lieu de culte?
 pas du tout un peu modérément beaucoup
12. Faire du bénévolat?
 pas du tout un peu modérément beaucoup

Les pertes d'urine ou le prolapsus de la vessie vous nuisent-ils dans les activités suivantes?

13. Faire un travail rémunéré à l'extérieur de la maison?
 pas du tout un peu modérément beaucoup
14. Recevoir des amis chez vous?
 pas du tout un peu modérément beaucoup
15. Participer à des activités sociales?
 pas du tout un peu modérément beaucoup
16. Vos relations avec vos amis?
 pas du tout un peu modérément beaucoup
17. Vos relations avec votre famille?
 pas du tout un peu modérément beaucoup
18. Vos relations sexuelles?
 pas du tout un peu modérément beaucoup
19. Votre façon de vous habiller?
 pas du tout un peu modérément beaucoup
20. Votre santé mentale?
 pas du tout un peu modérément beaucoup
21. Votre santé physique?
 pas du tout un peu modérément beaucoup
22. Votre sommeil?
 pas du tout un peu modérément beaucoup
23. La peur des odeurs vous limite-t-elle dans vos activités?
 pas du tout un peu modérément beaucoup
24. La peur de vivre des situations embarrassantes vous limite-t-elle dans vos activités?
 pas du tout un peu modérément beaucoup

De plus, votre problème de santé suscite-t-il chez vous les réactions suivantes?

25. La nervosité ou l'anxiété?

- pas du tout un peu modérément beaucoup

26. La peur?

- pas du tout un peu modérément beaucoup

27. La frustration?

- pas du tout un peu modérément beaucoup

28. La colère?

- pas du tout un peu modérément beaucoup

29. La dépression?

- pas du tout un peu modérément beaucoup

30. La gêne?

- pas du tout un peu modérément beaucoup

Veillez, s'il vous plaît, vous assurer d'avoir répondu à toutes les questions.

MERCI BEAUCOUP POUR VOTRE COLLABORATION!

Annexe III SF-12

Veillez inscrire la date d'aujourd'hui (jour/mois/année), s.v.p. : ___/___/___

1. En général, trouvez-vous que votre santé est:

(Encerclez une seule réponse)

- Excellente 1
 Très bonne 2
 Bonne 3
 Passable 4
 Mauvaise 5

2. Les questions suivantes portent sur les activités que vous pourriez avoir à faire au cours d'une journée normale. Votre état de santé actuel vous limite-t-il dans ces activités? Si oui, dans quelle mesure?

(Encerclez un seul chiffre par ligne)

- | | Mon état
de santé
me limite
<u>beaucoup</u> | Mon état
de santé
me limite
<u>un peu</u> | Mon état
de santé
ne me limite
<u>pas du tout</u> |
|--|--|--|--|
| a. Dans les activités modérées comme
déplacer une table, passer l'aspirateur,
jouer aux quilles ou au golf | 1 | 2 | 3 |
| b. Pour monter quelques étages à pied | 1 | 2 | 3 |

3. Au cours des 4 DERNIÈRES SEMAINES, avez-vous eu l'une ou l'autre des difficultés suivantes, au travail ou à vos autres activités quotidiennes habituelles, à cause de votre ÉTAT DE SANTÉ PHYSIQUE.

(Encerclez un seul chiffre par ligne)

- | | <u>Oui</u> | <u>Non</u> |
|---|-------------------|-------------------|
| a. Avez-vous accompli moins de choses que vous
l'auriez voulu | 1..... | 2 |
| b. Avez-vous été limitée dans la nature de vos
tâches ou de vos autres activités? | 1..... | 2 |

4. **Au cours des 4 DERNIÈRES SEMAINES, avez-vous eu l'une ou l'autre des difficultés suivantes, au travail ou à vos autres activités quotidiennes habituelles, à cause de l'ÉTAT DE votre moral**
(Encerclez un seul chiffre par ligne)

	<u>Oui</u>	<u>Non</u>
a. Avez-vous accompli moins de choses que vous l'auriez voulu	1	2
b. Avez-vous fait votre travail ou vos autres activités avec moins de soin qu'à l'habitude?	1	2

5. **Ces questions portent sur les 4 DERNIÈRES SEMAINES. Pour chacune des questions suivantes, donnez la réponse qui s'approche le plus de la façon dont vous vous êtes sentie.**
(Encerclez un seul chiffre par ligne)

	<u>Tout</u> le <u>temps</u>	<u>La</u> plupart <u>du temps</u>	<u>Souvent</u>	<u>Quelque</u> <u>fois</u>	<u>Rarement</u>	<u>Jamais</u>
a. Vous êtes-vous sentie calme et sereine?	1	2	3	4	5	6
b. Avez-vous eu beaucoup d'énergie?	1	2	3	4	5	6
c. Vous êtes-vous sentie triste et abattue?	1	2	3	4	5	6

6. **Au cours des 4 DERNIÈRES SEMAINES, dans quelle mesure votre état physique ou moral a-t-il nui à vos activités sociales habituelles (famille, amis, voisins ou autres groupes)?**
(Encerclez une seule réponse)

Pas du tout	1
Un peu	2
Moyennement	3
Beaucoup	4
Énormément	5

**7. Au cours des 4 DERNIÈRES SEMAINES, dans quelle mesure la douleur a-t-elle
nui à vos activités habituelles (au travail comme à la maison)?**

(Encerchez une seule réponse)

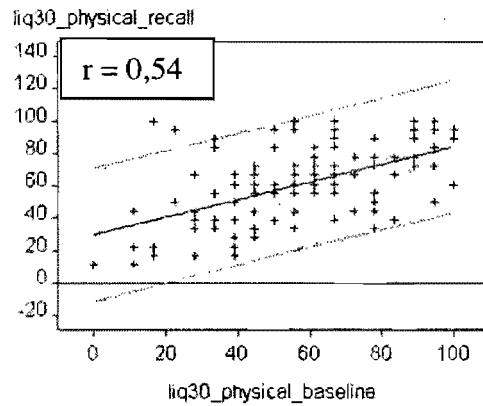
- Pas du tout 1
- Un peu 2
- Moyennement 3
- Beaucoup 4
- Énormément 5

Veillez, s'il vous plaît, vous assurer d'avoir répondu à toutes les questions.

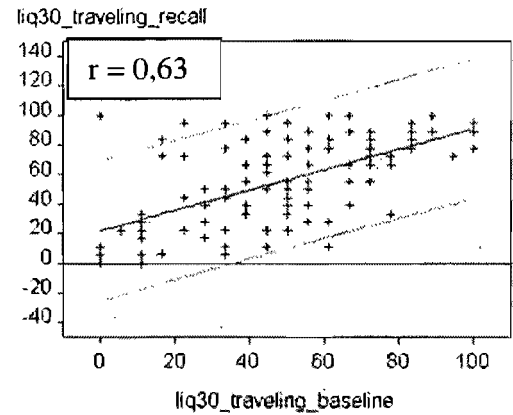
MERCI BEAUCOUP POUR VOTRE COLLABORATION!

Annexe IV Scatterplots of recalled QOL scores compared to baseline preoperative QOL scores for the IIQ subscales: physical (a), traveling (b), emotional (c), social (d).

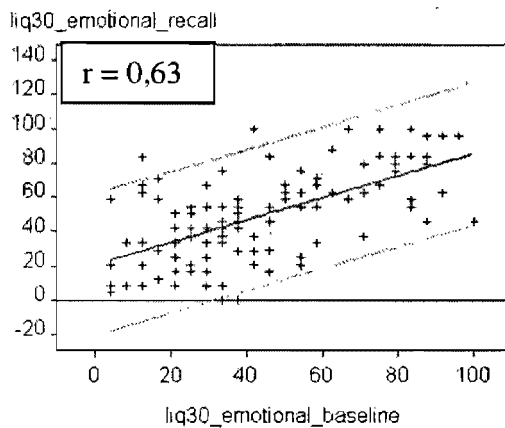
(a)



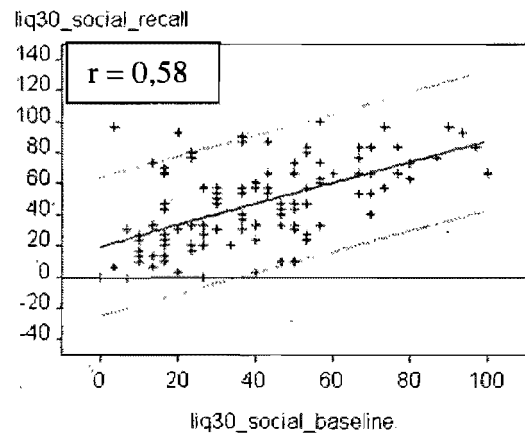
(b)



(c)



(d)



r : Pearson's correlation coefficient

Annexe V Patient's recalled scores of preoperative quality of life compared to prospective baseline scores of preoperative quality of life for the SF-12 Physical (PCS) (2.1) and Mental (MCS) (2.2) scales.

Figure 2.1

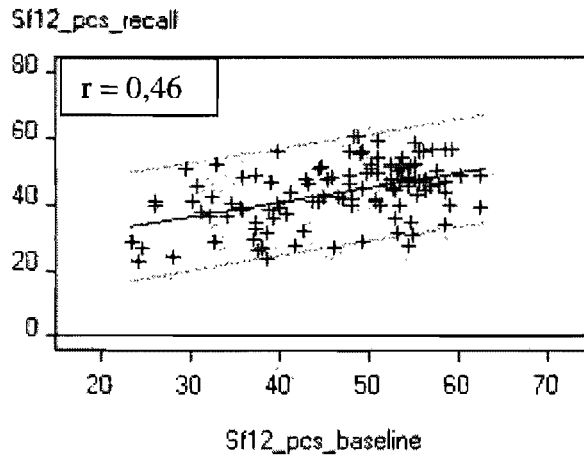
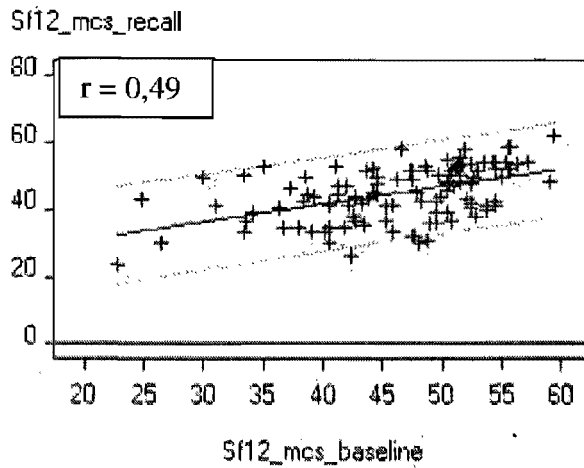


Figure 2.2



r : Pearson's correlation coefficient

Annexe VI Univariate linear regression where the difference between Baseline score and recalled score is the dependant variable and patient characteristics are the independent variables for SF-12.

SF-12 (PCS)	β	Standard Error (SE)	P-value	R ²
• Age	0.03	0.06	0.5397	0.0034
• Living situation	-0.04	1.30	0.9767	0.0000
• Marital Status	-0.48	1.21	0.6928	0.0014
• Employment status	-2.29	1.24	0.0679	0.0297
• Smoking status	0.69	1.23	0.5769	0.0028
• Education category	-0.95	0.69	0.1736	0.0166
• Income category	-0.97	0.82	0.2443	0.0144
• Comorbidity	1.69	1.20	0.1620	0.0175
• Duration to recall	-0.06	0.13	0.6386	0.0020

SF-12 (MCS)	β	Standard Error (SE)	P-value	R ²
• Age	0.02	0.04	0.5669	0,0030
• Living situation	-0.87	1.03	0,4012	0.0064
• Marital Status	-0.35	0.96	0.7117	0.0012
• Employment status	-1.63	0.99	0.1018	0.0239
• Smoking status	0.21	0.98	0.8270	0.0004
• Education category	0.12	0.56	0.8230	0.0005
• Income category	-0.61	0.69	0.3753	0.0084
• Comorbidity	1.28	0.95	0.1823	0.0160
• Duration to recall	0.00	0.11	0.9772	0.0000

Annexe VII Multivariate analysis of age, employment status, education category and income category on recalled score using IIQ-30.

Models	Regression coefficient (β)				R²
	Age	Employment status	Education Category	Income Category	
Model 1	1.03	-7.67	-1.51	-11.84	0.1059
Model 2	1.06	-3.99	-6.46		0.0733
Model 3	1.02	-8.52		-12.63	0.1056
Model 4	1.19*		-2.74	-11.87	0.1035
Model 5		-1.61	-0.82	-12.69	0.0800
Model 6	0.07	-8.04			0.0645
Model 7	1.15*		-6.98		0.0726
Model 8	1.22*			-13.47	0.1023

* p < 0.05

Annexe VIII Sociodemographic characteristics and baseline QOL scores between incontinents and continents subjects

Patient characteristics	Incontinents (n=113)	Continents (n=51)
Age at baseline (years)		
Mean	55.6	38.1
Range	(30-77)	(18-59)
Race		
White	110 (97.4%)	48 (94.1%)
Marital status		
Married	53 (46.9%)	8 (15.7%)
Living situation		
Living with a spouse or a partner	85 (75.2%)	31 (60.8%)
Smoking status		
Current smoker	68 (60.2%)	32 (62.8%)
Education		
Did not complete high school	34 (30.1%)	6 (11.8%)
Completed high school	26 (23.0%)	8 (15.7%)
Completed college or university	53 (46.9%)	37 (72.6%)
Employment situation		
Full or part time job	73 (64.6%)	47 (92.2%)
Income*		
Less than \$20,000	18 (15.9%)	10 (19.6%)
Between \$20,000 – 50,000	33 (29.2%)	13 (25.5%)
More than \$50,000	45 (39.8%)	28 (54.9%)
Co-morbidities**	51 (45.1%)	11 (21.6%)
IIQ baseline scores (mean)		
Total	187.4	2.2
Physical	57.1	0.5
Traveling	48.7	0.8
Social	40.2	0.2
Emotional	41.5	0.7
SF-12 baseline scores (mean)		
PCS	46.6	56.0
MCS	46.4	45.2

* 17 (15.0%) patients did not respond to this question

** Subjects with one or another past or present co-morbidity: diabetes, myocardial infarction, stroke, amputation, leg circulation problems, asthma, stomach ulcer, kidney disease, major depression, epilepsy, alcoholism and toxicomania

Curriculum vitae

Annick Larochelle, M.D.

Formation

93/09 à 98/12	Études médicales	Université de Montréal, Faculté de Médecine, Montréal, (Qc)
99/07-04/06	Résidence en obstétrique gynécologie	Université de Montréal, Faculté de médecine, Montréal (Qc)

Diplômes, certificats et permis

1999	Diplômée en médecine (M.D.) Université de Montréal, Faculté de médecine, Montréal, (Qc)
2000	Conseil médical du Canada (L.M.C.C.) – date : 18 décembre 2000 Certificat # 88562
2004	Collège des médecins du Québec – 1 ^{er} juillet 2004 Permis d'exercice # 04-052
2004	Collège des médecins du Québec – 1 ^{er} juillet 2004 (C.S.P.Q. – Certificat # 16796) Examen du certificat de spécialiste en obstétrique-gynécologie réussi les 21 et 22 mai 2005
2004	Diplôme d'études supérieures (D.E.S.) Spécialité: Obstétrique-Gynécologie Université de Montréal, Faculté de médecine, Montréal, (Qc)
2004	Collège royal des médecins et chirurgiens du Canada (F.R.C.S.C. - Certifiée en obstétrique-gynécologie)

Formation complémentaire et post-doctorale

2004 – 2005	6 ^{ième} année de résidence orientée en urogynécologie dans le CHUM, l'hôpital Sainte-Justine et l'hôpital Maisonneuve-Rosemont, supervisée par le Dr. Stéphane Ouellet.
2004 – 2008	Maîtrise en Science biomédicales avec option recherche clinique (Msc). Université de Montréal, département de médecine sociale et préventive.
09/2005 – 01/2006	Programme de chirurgie pelvienne avancée, endoscopie et urogynécologie, Maternité de l'Hôtel-Dieu, C.H.U. de Clermont-Ferrand, France, sous la direction du Professeur Bernard Jacquetin.
02/2007 – 06/2007	Programme de fellowship en urogynécologie au Civic Hospital d'Ottawa, Ontario, Canada, sous la supervision du Dr Joyce Schachter.

Projets de recherche

- Titre : Accuracy of the BVI 3000 in the assessment of postvoid residual volume (PVR) among patients coming to a urogynecologic clinic
Noms: G. Al-Shaikh, A. Larochelle, J. Schachter, C. Campbell, K. Baker.
Site: Civic Hospital, Ottawa, (ON).
- Titre : Les complications per-opératoires et postopératoires de la bandelette transobturatrice pour le traitement de l'incontinence urinaire de stress
Noms: A. Larochelle, L. Leboeuf, S. Ouellet
Site : Hopital St-Luc, Montréal, (QC)
- Titre: Traitement des grossesses avec allo-immunisation plaquettaire
Noms: A. Larochelle, A. Skoll, M. David
Site : CHU Sainte-Justine, Montréal, (QC)

Rayonnement

- Publications : 2

Publications de recherche dans les revues non dotées d'un comité de lecture.

« Le pessaire : une solution efficace au prolapsus génital ». Le Clinicien; juin 2007; volume 22; numéro 6; p. 63-66.

Abrégés publiés

“Safety of the Transobturator (TOT) procedure (Monarc™ AMS): A review of 114 patients from one institution.” Présentation en « Read-by-title » au congrès de l' « International Continence Society » (ICS) à Montréal, du 29 août au 2 septembre 2005.

- Conférencière invitée : 5

- Affiches et exhibits scientifiques : 3

- Présentation scientifique nationale : 1

Responsabilités professionnelles

2006 à ce jour	Chargée d'enseignement clinique au département d'obstétrique gynécologie de l'universitaire de Montréal
2007 à ce jour	Membre du comité des titres de l'hôpital Sainte-Justine de Montréal
2007 à ce jour	Responsable des réunions scientifiques du département d'obstétrique gynécologie de l'hôpital Sainte-Justine

Titres honorifiques, prix, distinctions

Juin 2004	Lauréate 2003-2004 pour l'excellence des procédures endoscopiques par le comité pédagogique post-doctoral, département obstétrique-gynécologie, Faculté de médecine, Université de Montréal, via le « Special Excellence in Endoscopies Procedures » de l'AAGL (American Association of Gynecologic Laparoscopists).
-----------	--