

Université de Montréal

**À qui le crédit? Division du travail et du capital dans les
collaborations de recherche**

par

Philippe Mongeon

École de bibliothéconomie et des sciences de l'information
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences de l'information

Décembre 2017

© Philippe Mongeon, 2017

Résumé

Le statut d'auteur attribue à la fois le crédit et la responsabilité aux chercheurs pour leur contribution à l'avancement des connaissances, et permet l'existence d'une économie du capital symbolique où les chercheurs bâtissent leur réputation et progressent dans leur carrière en signant des articles. Ce système repose sur l'existence d'un lien historiquement clair entre le statut d'auteur et la contribution. Or, ce lien est obscurci par le nombre de plus en plus élevé d'auteurs sur les articles scientifiques, par le fait que le statut d'auteur peut être attribué pour des contributions hétérogènes, parfois minimales, ou même nulles, et par la subjectivité du processus d'attribution du statut d'auteur. Ainsi, lorsqu'un article est signé par plusieurs individus, il est difficile de déterminer les contributions de chacun.

Dans certaines disciplines, les chercheurs participent de plus en plus au régime de protection de la propriété intellectuelle en brevetant leurs découvertes. Contrairement au statut d'auteur, dont les pratiques d'attribution sont de plus en plus hétérogènes et subjectives, le statut d'inventeur d'un brevet est, de par son aspect légal, attribué selon des critères plus stricts. Il est en effet réservé aux individus ayant participé à la conception et à la matérialisation de l'invention. Ainsi, dans le cas des travaux de recherche menant à la fois à un article et à un brevet, la liste des inventeurs du brevet permet, en théorie, d'identifier parmi les auteurs ceux qui ont fait une contribution inventive et, donc, de mieux comprendre la relation entre le statut d'auteur et la nature de la contribution effectuée par chacun des auteurs de l'article.

Notre étude compare les listes d'auteurs et d'inventeurs d'articles et de brevets issus de la même recherche afin de mieux comprendre les pratiques d'attribution des statuts d'auteur et d'inventeur. Nous utilisons la régression linéaire pour analyser la relation entre les caractéristiques de l'article (discipline, impact scientifique potentiel, nombre d'auteurs, collaboration interinstitutionnelle) et la proportion d'auteurs qui sont aussi inventeurs. De plus, nous utilisons la régression logistique pour déterminer dans quelle mesure il est possible de prédire quels auteurs d'un article obtiendront également le statut d'inventeur à partir de leur position dans la liste des auteurs, leur rôle, leur contribution, leur sexe et leur réputation. Les résultats montrent que la proportion moyenne des auteurs qui obtiennent le statut d'inventeur est liée à la discipline, est plus faible pour les recherches impliquant plusieurs institutions, et est

plus élevée pour les articles publiés dans une revue à haut facteur d'impact. Les résultats montrent également que les auteurs occupant la première et la dernière position dans les listes d'auteurs sont plus souvent inventeurs que les auteurs occupant les autres positions, que les auteurs hautement cités sont plus souvent inventeurs que les auteurs peu cités, et que les hommes sont plus souvent inventeurs que les femmes.

Mots-clés : article scientifique, brevet, auteur scientifique, inventeur, collaboration, division du travail, bibliométrie, scientométrie, capital scientifique

Abstract

The function of scientific authorship is to assign credit and responsibility to researchers for their contribution to a piece of knowledge, thus enabling the existence of an economy of reputation economy in which researchers advance their careers by, among others, having their names associated to scientific work. This system relies on the existence of a historically clear link between authorship and contribution. However, this link is obscured by the increasingly high number of authors appearing in the bylines of scientific papers, by the fact that authorship can be given for diverse (sometimes very small) contributions, and by the subjectivity of authorship decisions. Thus, when an article is authored by several researchers, it becomes difficult to determine their respective contributions.

In certain research fields, researchers can be involved in the commercialization of research, and often patent their discoveries. Patents are thus another form of contribution which is valued in the evaluation of researchers. Unlike the heterogeneous and subjective authorship practices, inventorship is, because of its legal aspect, awarded using stricter criteria. It is reserved for individuals who designed and materialized the invention. In some cases, research may lead to both an article and a patent. The list of inventors can then, in theory, help identify the authors who have made an inventive contribution and thus provide a better understanding of the relationship between authorship and the nature of the individual researchers' contribution to the work.

Our study compares the lists of authors and inventors of articles and patents reporting the same research. Specifically, we use linear regression to analyse the relationship between the characteristics of articles (discipline, potential impact, number of authors, interinstitutional collaboration) and the proportion of authors who also appear as inventors. In addition, we use logistic regression to determine to what extent the inventor status of authors can be predicted by their position in the byline of the paper, their role in the research, the nature of their contribution, their gender and their reputation. The results show that the average proportion of authors who are named inventors differs by discipline, is lower for research involving multiple institutions, and is higher when the article is published in a top Impact Factor journal. The results also show that the authors in the first and last positions of the byline are more often inventors than authors

in other positions, that highly cited authors are more often inventors than authors with fewer citations, and that men are more often inventors than women.

Keywords : scientific article, patent, authorship, inventorship, collaboration, bibliometrics, scientometrics, scientific capital

Table des matières

Introduction.....	1
Objectifs de l'étude.....	4
Chapitre 1. Revue de la littérature	7
Introduction.....	7
1.1. La structure et le fonctionnement de la communauté scientifique	7
1.1.1. Bourdieu et les concepts de champ, d'agent, de capital et d'habitus.....	8
1.1.2. Merton, les normes et le système de reconnaissance de la science	13
1.1.3. Liens entre les travaux de Merton, Bourdieu, et notre recherche	21
1.1.4. En résumé.....	24
1.2. L'auteur scientifique	25
1.2.1. Les fonctions du statut d'auteur en science	25
1.2.2. Les normes et pratiques d'attribution du statut d'auteur en contexte collaboratif..	27
1.2.3. Politiques éditoriales des revues savantes.....	36
1.2.4. En résumé.....	39
1.3. Le brevet d'invention et le statut d'inventeur	40
1.3.1. Le brevet d'invention	40
1.3.2. L'inventeur.....	43
1.3.3. En résumé.....	49
1.4. L'appariement des articles et des brevets	50
1.4.1. Techniques de détection des liens entre les articles et les brevets.....	50
1.4.2. En résumé.....	59
1.5. Résultats des études sur l'attribution des statuts d'auteur et d'inventeur réalisées à partir de paires brevet-article.....	59
Conclusion	63
Chapitre 2. Méthodologie	67
Introduction.....	67
2.1. Approche méthodologique générale	68
2.2. Sources et traitement des données	70

2.2.1. Sources de données	70
2.2.2. Traitement des données.....	73
2.2.3. Échantillon	92
2.2.4. En résumé.....	93
2.3. Analyse des données	94
2.3.1. Hypothèses de recherche.....	94
2.3.2. Prédiction du ratio inventeurs/auteurs (HS ₁ à HS ₄)	96
2.3.3. Prédiction du statut d'inventeur (HS ₅ à HS ₁₁)	99
2.3.4. Liens entre les hypothèses de recherche, les variables, les sources de données, et les types d'analyse.....	103
2.3.5. En résumé.....	103
2.4. Qualité de la recherche.....	104
2.4.1. Validité de conclusion statistique	104
2.4.2. Validité interne.....	105
2.4.3. Validité de construit.....	106
2.4.4. Validité externe	107
2.4.5. Fidélité	107
2.4.6. Objectivité.....	108
Conclusion	108
Chapitre 3. Résultats	110
Introduction.....	110
3.1. Prédicteurs du ratio inventeurs/auteurs	110
3.1.1. Le ratio inventeurs/auteurs en fonction du nombre d'auteurs, de la collaboration interinstitutionnelle, du facteur d'impact de la revue de publication et de la discipline	111
3.1.2. En résumé.....	125
3.2. Prédicteurs du statut d'inventeur	125
3.2.1. L'obtention du statut d'inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.....	126
3.2.2. L'obtention du statut d'inventeur en fonction du rôle, du sexe et de la réputation en recherche biomédicale	136

3.2.3. L'obtention du statut d'inventeur en fonction de la contribution, du sexe et de la réputation en recherche biomédicale	140
3.2.4. En résumé.....	147
Conclusion	148
Chapitre 4. Discussion	151
Introduction.....	151
4.1. Le ratio inventeurs/auteurs dans les équipes de recherche	151
4.1.1. Enjeux, structure, culture et normes de la discipline	151
4.1.2. Impact potentiel de la recherche	154
4.1.3. Les caractéristiques de l'équipe et la concentration du statut d'inventeur	156
4.2. Qui est inventeur?	159
4.2.1. Entre contribution et rapports de force	159
4.2.2. Un effet Saint-Matthieu dans l'attribution du statut d'inventeur?	162
4.2.3. Inégalités femmes-hommes dans l'attribution du statut d'inventeur	164
4.3. Limites de la recherche	165
Conclusion	167
Conclusion générale.....	169
Implications théoriques.....	171
Implications méthodologiques	172
Implications pratiques.....	173
Recherches futures	174
Bibliographie.....	i
Bibliographie – documents légaux.....	xviii
Législation.....	xviii
Jurisprudence	xviii
Annexe 1. Extrait des recommandations de l'International Committee of Medical Journal Editors.....	xix
Annexe 2. Exemple de description des contributions individuelles des coauteurs d'un article publié dans le <i>Journal of the American Medical Association</i>	xxii

Annexe 3. Exemple de brevet (première page seulement)	xxiii
Annexe 4. Nom, description et source des champs de la base de données SQL utilisée pour la recherche	xxiv
Annexe 5. Schéma de la base de données relationnelle utilisée pour la recherche	xxvi
Annexe 6. Tableaux et figures pour la vérification du respect des postulats de la régression linéaire.....	xxvii
Annexe 7. Tableaux de corrélation pour la vérification des postulats de la régression logistique	xxxi

Liste des tableaux

Tableau I. Informations utilisées par Dornbusch et al. (2013) pour l'appariement des brevets et des articles.....	53
Tableau II. Catégories de PBA de Coward et Franklin (1989, traduction libre, p. 60-61).....	55
Tableau III. Propriétés des techniques de liaison des articles et des brevets (adapté de Bassecoulard et Zitt, 2005, p.689, notre traduction).....	58
Tableau IV. Facteurs influençant l'attribution des statuts d'auteur et d'inventeur.....	66
Tableau V. Exemple d'extraction des SN dans le titre d'un article.....	76
Tableau VI. Nombre et proportion de PBA valides en fonction du score de similarité.	77
Tableau VII. Nombre de PBA, d'articles et de brevets par discipline.....	81
Tableau VIII. Exemple des modalités de la variable « position » pour chacun des auteurs d'un article ayant huit auteurs.	83
Tableau IX. Exemple d'article comportant une séquence d'auteurs périphériques listés en ordre alphabétique.	86
Tableau X. Nombre et proportion des auteurs par rôle et par discipline pour les articles en ordre alphabétique partiel.	86
Tableau XI. Nombre et proportion de libellés de contribution par catégorie.	88
Tableau XII. Nombre d'auteurs distincts et d'inventeurs distincts par discipline.	89
Tableau XIII. Évaluation de l'algorithme Wiki-Gendersort à l'aide de quatre bases de données associant un sexe à des prénoms (adapté de Bérubé et al. [2017]).	91
Tableau XIV. Nombre et proportion des auteurs par sexe et par discipline.	91
Tableau XV. Taille des échantillons pour chaque question de recherche.	92
Tableau XVI. Exemple de valeurs de probabilités, rapport de cotes et logit (p).	100
Tableau XVII. Hypothèses vérifiées, variables prédites et variables prédictives pour chacun des modèles globaux de régression logistique.	101
Tableau XVIII. Liens entre les hypothèses de recherche, les sources de données et le type d'analyse.	103
Tableau XIX. Statistiques descriptives sur le nombre d'auteurs des PBA par discipline en fonction de la collaboration interinstitutionnelle et du facteur d'impact de la revue de publication.....	112

Tableau XX. Statistiques descriptives sur le nombre d'inventeurs des PBA par discipline en fonction de la collaboration interinstitutionnelle et du facteur d'impact de la revue de publication.....	113
Tableau XXI. Statistiques descriptives sur le ratio inventeurs/auteurs des PBA par discipline en fonction de la collaboration interinstitutionnelle et du facteur d'impact de la revue de publication.....	114
Tableau XXII. Qualité de l'ajustement et valeur prédictive du modèle de régression linéaire multiple hiérarchique visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI.....	120
Tableau XXIII. Significativité statistique du modèle de régression linéaire visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI.....	121
Tableau XXIV. Coefficients du modèle de régression linéaire visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI.....	122
Tableau XXV. Coefficients du modèle de régression linéaire visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI (recherche biomédicale exclue).....	122
Tableau XXVI. Coefficient de la régression linéaire visant à prédire le ratio inventeurs/auteurs pour chaque discipline à l'aide du nombre d'auteurs de la collaboration interinstitutionnelle et du FI.....	123
Tableau XXVII. Qualité de l'ajustement et capacité prédictive des modèles de régression logistique visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs	131
Tableau XXVIII. Coefficients du modèle de régression logistique global visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs, de la discipline, du sexe et de la réputation.....	132
Tableau XXIX. Coefficients du modèle de régression logistique excluant la recherche biomédicale et visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs, de la discipline, du sexe et de la réputation.....	133

Tableau XXX. Coefficients du modèle de régression logistique pour chaque discipline visant à prédire le statut d’inventeur en fonction de la position dans la liste des auteurs, du sexe, et de la réputation.	134
Tableau XXXI. Qualité de l’ajustement et capacité prédictive du modèle de régression logistique hiérarchique visant à prédire l’obtention du statut d’inventeur en fonction du rôle, du sexe et de la réputation.....	138
Tableau XXXII. Coefficient du modèle de régression logistique visant à prédire le statut d’inventeur en fonction du rôle, du sexe et de la réputation.	139
Tableau XXXIII. Qualité de l’ajustement et capacité prédictive du modèle de régression logistique hiérarchique visant à prédire l’obtention du statut d’inventeur en fonction de la contribution, du sexe et de la réputation, et contrôlant pour la position dans la liste des auteurs.	145
Tableau XXXIV. Coefficient du modèle de régression logistique hiérarchique visant à prédire l’obtention du statut d’inventeur en fonction de la contribution, du sexe et de la réputation, et contrôlant pour la position dans la liste des auteurs.	146
Tableau XXXV. Tableau récapitulatif des résultats des hypothèses.	150

Liste des figures

Figure 1. Relation entre les variables du modèle visant à prédire le ratio inventeurs/auteurs des PBA (HS ₁ à HS ₄).	98
Figure 2. Modèle de régression logistique visant à prédire l'obtention du statut d'inventeur par les auteurs des PBA à partir de la position dans la liste des auteurs, de la discipline, du sexe et de la réputation (HS ₅ , HS ₆ , HS ₁₀ et HS ₁₁).....	101
Figure 3. Modèle de régression logistique visant à prédire l'obtention du statut d'inventeur par les auteurs des PBA à partir du rôle, du sexe et de la réputation (HS ₇ , HS ₈ , HS ₁₀ et HS ₁₁)... ..	102
Figure 4. Modèle de régression logistique visant à prédire l'obtention du statut d'inventeur les auteurs des PBA à partir de la contribution, du sexe et de la réputation (HS ₉ , HS ₁₀ et HS ₁₁). ..	102
Figure 5. Ratio inventeurs/auteurs moyen en fonction du nombre d'auteurs et par discipline.	115
Figure 6. Ratio inventeurs/auteurs moyen en génie en fonction du nombre d'auteurs pour l'ensemble des articles, les articles en collaboration interinstitutionnelle et les articles publiés dans une revue à haut FI.	116
Figure 7. Ratio inventeurs/auteurs moyen en recherche biomédicale en fonction du nombre d'auteurs pour l'ensemble des articles, les articles en collaboration interinstitutionnelle et les articles publiés dans une revue à haut FI.	117
Figure 8. Ratio inventeurs/auteurs moyen en sciences naturelles, en fonction du nombre d'auteurs pour l'ensemble des articles, les articles en collaboration interinstitutionnelle et les articles publiés dans une revue à haut FI.	117
Figure 9. Proportion d'auteurs en génie ayant le statut d'inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.	128
Figure 10. Proportion d'auteurs en recherche biomédicale ayant le statut d'inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.	128
Figure 11. Proportion d'auteurs en sciences naturelles ayant le statut d'inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.	129
Figure 12. Proportion d'auteurs ayant le statut d'inventeur en fonction du rôle, du sexe et de la réputation.	137

Figure 13. Proportion d’auteurs ayant effectué un type de contribution en fonction de la position dans la liste des auteurs.	142
Figure 14. Proportion d’auteurs ayant le statut d’inventeur selon la contribution effectuée et la position dans la liste des auteurs.	143
Figure 15. Proportion d’auteurs ayant le statut d’inventeur en fonction du sexe, de la réputation et de la contribution.	144

Liste des sigles

BTT.....	Bureau de transfert de technologie
CWTS.....	Centre for Science and Technology Studies
EASE.....	European Association of Science Editors
FI.....	Facteur d'impact
ICMJE.....	International Committee of Medical Journal Editors
NBER.....	National Bureau of Economic Research
NSF.....	National Science Foundation
OEB.....	Office européen des brevets
OMPI.....	Organisation mondiale de la propriété intellectuelle
ORI.....	Office of Research Integrity
OST.....	Observatoire des sciences et des technologies
PatFT.....	USPTO Patent Full-Text and Image Database
PBA.....	Paire brevet-article
SCI.....	Science Citation Index
SCIE.....	Science Citation Index Expanded
SN.....	Syntagme nominal
USPTO.....	United States Patent and Trademark Office
WoS.....	Web of Science

Remerciements

La réalisation d'une thèse de doctorat est un parcours sinueux et parsemé de défis et d'embûches de toutes sortes. J'ai eu la chance d'être entouré de personnes et d'organisations qui m'ont non seulement permis de relever ces défis et de surmonter les difficultés, mais qui ont également contribué à faire des quatre dernières années une aventure inoubliable. Je tiens donc à remercier tous ceux et celles que j'ai côtoyés au cours de mon parcours doctoral et qui ont contribué de près ou de loin à son succès.

Tout d'abord, je remercie du fond du cœur mes directeurs de recherche, Vincent Larivière et Christine Dufour, pour leur accompagnement tout au long de ce parcours. Un merci spécial pour la dernière année qui a été particulièrement intense. Je n'y serais pas arrivé sans votre soutien et le sacrifice de quelques-unes de vos fins de semaine. Je tiens également à remercier les membres du jury, Guillaume Cabanac, Marie Demoulin et Clément Arsenault pour la relecture de ma thèse et leurs précieux commentaires.

Pour leur soutien technique, je tiens à remercier Mario Rouette et Jean-Pierre Robitaille de l'Observatoire des sciences et des technologies, Nees Jan van Eck du Centre for Science and Technology Studies de l'Université de Leyde, My-Linh Luu et Nicolas Bérubé de la chaire de recherche du Canada sur les transformations de la communication savante, ainsi que Marielle Saint-Germain remercier de l'École de bibliothéconomie et des sciences de l'information.

Pour le soutien financier qui m'a permis de me concentrer sur ma mes activités de recherche, je remercie le Conseil de recherches en sciences humaines du Canada, la Chaire de recherche du Canada sur les transformations de la communication savante, l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal, le Centre interuniversitaire de recherche sur la science et la technologie, la Faculté des études supérieures de l'Université de Montréal, et l'International Society for Scientometrics and Informetrics.

Pour le soutien moral, la rétroaction et pour avoir fait du C-2070 un espace de travail et d'échanges d'une richesse inestimable, je remercie toute l'équipe de la Chaire de recherche sur les transformations de la communication savante. Un merci spécial à Adèle Paul-Hus, Antoine Archambault, Diane Marie Plante, Fei Shu, Gita Ghiasi, Maxime Sainte-Marie, Nadine

Desrochers, Nicolas Bérubé, Rodrigo Costas, Sarah Cameron-Pesant, Stefanie Haustein et Tim Bowman.

Finalement, je remercie Jessie pour son support et sa patience, ainsi que mes parents pour leur appui inconditionnel dans l'ensemble du parcours m'ayant conduit jusqu'à la réalisation de cette thèse.

Introduction

Le rôle de la science en tant qu'institution et, par extension, celui des chercheurs, est l'avancement des connaissances (Merton, 1942). L'efficacité de l'institution scientifique est maximisée lorsque ceux qui remplissent le mieux ce rôle sont reconnus et récompensés (Cole et Cole, 1973; Merton, 1957). Les chercheurs n'ayant pas de droit de propriété sur leurs contributions, qui généralement deviennent publiques lors de leur publication dans une revue savante, leur principale récompense est la reconnaissance de leurs pairs (le capital symbolique, au sens de Bourdieu, 1986b). La science s'est donc dotée d'un système de reconnaissance (p. ex. prix honorifiques, postes prestigieux, promotion, octroi de ressources) visant à récompenser les chercheurs s'étant le plus illustrés. Le statut d'auteur joue un rôle crucial dans ce système puisqu'il permet d'attribuer le crédit aux individus pour leurs contributions (Birnholtz, 2006). C'est lui qui rend possible l'identification des meilleurs chercheurs et qui permet à la science de fonctionner de façon méritocratique. Le système de recherche scientifique est toutefois en constante évolution, et le problème de recherche abordé dans cette thèse provient de deux tendances s'étant graduellement manifestées au cours du dernier demi-siècle : 1) l'augmentation du nombre moyen d'auteurs signant les articles scientifiques, et 2) l'augmentation de la participation des chercheurs universitaires au régime de protection de la propriété intellectuelle et à la commercialisation de la recherche (Wong et Singh, 2010).

La collaboration accrue en science (Beaver et Rosen, 1979) et la taille croissante des équipes de recherche sont associées à une augmentation continue du nombre moyen d'auteurs par article (Larivière, Gingras, Sugimoto et Tsou, 2015; Wuchty, Jones et Uzzi, 2007). Le simple fait qu'un article ait plusieurs auteurs rend pratiquement impossible la détermination précise de la contribution respective de chacun des auteurs et donc de la part du crédit qui leur revient (Rennie, Yank et Emanuel, 1997). Plus le nombre d'auteurs est élevé, plus l'exercice devient hasardeux étant donné la diversité des pratiques de division du travail entre les membres de l'équipe (Larivière et al., 2016; Mongeon, Smith, Joyal et Larivière, 2017) et d'attribution du

statut d'auteur¹ (Pontille, 2006). Ces difficultés sont par ailleurs exacerbées par l'usage croissant d'indicateurs bibliométriques ne tenant pas compte de la diversité des pratiques de collaboration et d'attribution du statut d'auteur pour l'évaluation de la recherche, et ce au détriment des évaluations qualitatives par les pairs (Gingras, 2014).

L'autre tendance ayant une influence sur le système de reconnaissance est la participation accrue des chercheurs à la commercialisation de la recherche et plus spécifiquement, leur recours accru au régime de protection de la propriété intellectuelle (Malissard, 2010; Malissard, Gingras et Gemme, 2003; Owen-Smith et Powell, 2003; Thursby, Fuller et Thursby, 2009). Les brevets étant devenus une forme de contribution valorisée dans l'évaluation des chercheurs, ils constituent un autre moyen pour ces derniers de faire avancer leur carrière (Haeussler et Colyvas, 2011; Owen-Smith, 2003). Le statut d'inventeur est au brevet ce que le statut d'auteur est à l'article; cependant, les articles et les brevets opèrent dans des champs distincts (respectivement le champ scientifique et le champ économique) et sont donc régis par des règles et des normes différentes (Packer et Webster, 1996). Notamment, les contributions menant généralement au statut d'auteur (p. ex. les contributions techniques, la collecte de données, la rédaction du manuscrit) ne suffisent pas pour obtenir le statut d'inventeur sur un brevet (Haeussler et Sauermann, 2013). De plus, contrairement à l'article, le brevet n'apporte pas qu'un gain de capital symbolique pour les inventeurs, mais également (et surtout) un gain potentiel de capital économique, ce qui pourrait inciter les chercheurs à limiter le nombre d'inventeurs au strict minimum. Conséquemment, lorsqu'un travail de recherche mène à la fois à la publication d'un article et au dépôt d'une demande de brevet, le nombre d'auteurs de l'article est généralement supérieur au nombre d'inventeurs du brevet (Ducor, 2000; Haeussler et Sauermann, 2013; Lissoni, Montobbio et Zirulia, 2013). La valorisation du statut d'inventeur pourrait donc désavantager les individus dont la contribution n'est reconnue que dans le champ scientifique. Malgré les bouleversements potentiels que pourrait causer cette nouvelle forme de contribution dans le système de reconnaissance scientifique, la relation entre les statuts d'auteur et d'inventeur d'articles et de brevets issus de la même recherche demeure peu connue.

¹ Par attribution du statut d'auteur, nous entendons à la fois la nomination des individus en tant qu'auteurs d'un article et l'ordonnement des auteurs dans la liste.

Ces deux tendances peuvent à elles seules nuire au bon fonctionnement du système de reconnaissance scientifique tel que l'a représenté Merton (1957). En effet, pour Merton (1957) le bon fonctionnement de l'institution scientifique repose sur sa capacité à reconnaître et récompenser l'excellence. Or, le nombre de plus en plus élevé d'auteurs sur les articles scientifiques fait en sorte que le statut d'auteur à lui seul ne permet plus de déterminer le mérite individuel d'un chercheur. De plus, le brevet et le statut d'inventeur répondent à des impératifs qui ne sont pas purement dans l'intérêt de l'institution scientifique. Dans la présente étude, nous complétons ce cadre mertonien à l'aide de concepts issus de la théorie de Pierre Bourdieu (les principaux concepts en question seront définis dans le prochain chapitre). Pour Bourdieu, la science est un champ social défini par des rapports de force entre les chercheurs dont les actions sont pragmatiques et découlent des stratégies qu'ils adoptent dans le but de dominer le champ (Bourdieu, 1976). Dans ce cadre, l'attribution du statut d'auteur peut être vue comme un processus influencé par les rapports de force et les stratégies des agents impliqués. Cela peut, comme les deux tendances mentionnées ci-dessus, rendre plus ambigu le lien entre le statut d'auteur et le « mérite » du chercheur. En somme, la combinaison de ces deux cadres théoriques permet de conceptualiser les pratiques d'attribution des statuts d'auteur et d'inventeur comme étant, d'une part, guidées par des normes et visant à maximiser l'efficacité de l'institution scientifique et, d'autre part, une action pragmatique guidée avant tout par des considérations stratégiques individuelles.

Qu'en est-il alors, dans le cadre de travaux de recherche menant à la fois à une publication et à un brevet, du lien entre la contribution d'un chercheur, sa position dans la liste des auteurs et son inclusion dans la liste des inventeurs? Comment, et à quel point, la liste des auteurs d'un article peut-elle nous renseigner sur les contributions respectives de chacun d'eux à la recherche, et quels renseignements supplémentaires le brevet issu de cette même recherche peut-il fournir à cet égard? Peut-on observer empiriquement, à partir des listes d'auteurs et d'inventeurs, l'influence des rapports de force sur les pratiques d'attribution du capital symbolique et du capital économique que confèrent, respectivement, les articles et les brevets? Telles sont les questions qui émanent de la problématique décrite et qui sous-tendent notre recherche.

Du point de vue empirique, de nombreuses études antérieures se sont intéressées aux critères et pratiques d'attribution du statut d'auteur (p. ex. Frandsen et Nicolaisen, 2010; Henriksen, 2016; Jabbehdari et Walsh, 2017; Pontille, 2004, 2016) et à la division du travail dans les équipes de recherche (p. ex. Haeussler et Sauermann, 2015; Larivière et al., 2016; Perneger et al., 2017). D'autres études se sont intéressées aux liens entre la recherche scientifique et l'innovation technologique (p. ex. Etzkowitz, 1998; Leydesdorff et Meyer, 2003; Mansfield, 1991), ou encore au brevetage des résultats de la recherche académique (p. ex. Owen-Smith et Powell, 2003; Packer et Webster, 1996; Weiner, 1987). Un nombre plus restreint de travaux ont mis en relation les pratiques d'attribution des statuts d'auteur et d'inventeur, et très peu ont analysé des paires d'articles et de brevets issus de la même recherche à cette fin. À notre connaissance, cette recherche antérieure se limite à trois études, soit 1) l'étude de Ducor (2000) qui a comparé les listes d'auteurs et d'inventeurs de 40 paires de brevets et d'articles issus de la même recherche (PBA) dans le domaine de la biologie moléculaire, 2) l'étude de Haeussler et Sauermann (2013), qui ont réalisé une enquête auprès de deux mille chercheurs en sciences de la vie en Allemagne, et 3) l'étude de Lissoni, Montobbio et Zirulia (2013), qui ont analysé 680 PBA en chimie, en électronique, en pharmacologie et en biologie. Ces études suggèrent que le nombre d'inventeurs est généralement inférieur au nombre d'auteurs, que les contributions intellectuelles mènent davantage au statut d'inventeur que les contributions techniques, et que les femmes et les jeunes chercheurs sont plus souvent exclus de la liste des inventeurs. Elles sont toutefois limitées par le fait qu'elles ont de relativement petits échantillons de PBA (Ducor, 2000; Lissoni et al., 2013) dans des disciplines spécifiques (Ducor, 2000; Haeussler et Sauermann, 2013), dans des pays spécifiques (Haeussler et Sauermann, 2013; Lissoni et al. 2013), ou en utilisant des méthodes pouvant causer un biais dans l'échantillon (Haeussler et Sauermann, 2013).

Objectifs de l'étude

Notre étude examine un ensemble de PBA afin de mieux comprendre les pratiques de division du travail et d'attribution des statuts d'auteurs et d'inventeurs. Plus particulièrement nous analysons les listes d'auteurs et d'inventeurs afin de déterminer, d'une part, la relation entre divers facteurs et le ratio inventeurs/auteurs des PBA et, d'autre part, la relation entre

divers facteurs et l'obtention du statut d'inventeur. Cet objectif général se subdivise en quatre objectifs spécifiques, auxquels sont associées dix questions de recherche :

Objectif 1. Déterminer la relation entre les caractéristiques de la recherche et le nombre d'auteurs et d'inventeurs.

- QR1. Quelle est la relation entre la discipline de recherche et le ratio inventeurs/auteurs des PBA?
- QR2. Quelle est la relation entre l'impact scientifique potentiel de la recherche et le ratio inventeurs/auteurs des PBA?

Objectif 2. Déterminer la relation entre les caractéristiques de l'équipe de recherche et le nombre d'auteurs et d'inventeurs.

- QR3. Quelle est la relation entre la taille de l'équipe de recherche et le ratio inventeurs/auteurs des PBA?
- QR4. Quelle est la relation entre la collaboration interinstitutionnelle et le ratio inventeurs/auteurs des PBA?

Objectif 3. Déterminer la relation entre les contributions respectives des auteurs et l'obtention du statut d'inventeur.

- QR5. Quelle est la relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur?
- QR6. La relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur est-elle liée à la discipline?
- QR7. Quelle est la relation entre le rôle dans l'équipe de recherche et l'obtention du statut d'inventeur?
- QR8. Quelle est la relation entre le type de contribution à la recherche et l'obtention du statut d'inventeur?

Objectif 4. Déterminer la relation entre des caractéristiques individuelles des auteurs et l'obtention du statut d'inventeur.

- QR9. Quelle est la relation entre la réputation d'un auteur et l'obtention du statut d'inventeur?

QR10. Quelle est la relation entre le sexe d'un auteur et l'obtention du statut d'inventeur?

Notre recherche adopte une approche quantitative et utilise des méthodes bibliométriques pour l'analyse de PBA formées à partir des métadonnées des brevets du United States Patent and Trademark Office (USPTO) et de celles des articles indexés par le Web of Science (WoS). Elle contribue à l'avancement des connaissances en étudiant notamment l'existence de différences disciplinaires en termes de collaboration scientifique et de division du travail, ainsi qu'en mettant l'accent sur les contributions intellectuelles créatives telles qu'identifiées par l'attribution du statut d'inventeur. Notre recherche permet aussi de mieux comprendre la relation entre divers facteurs liés à la recherche et aux caractéristiques des chercheurs impliqués, et l'obtention du statut d'inventeur. En étudiant l'existence potentielle de liens entre des facteurs non liés à la contribution à la recherche (p. ex. le sexe et l'âge académique) et l'obtention du statut d'inventeur, notre recherche contribue à explorer le caractère subjectif des pratiques d'attribution des statuts d'auteur. Cette thèse a donc des implications pratiques pour les processus et politiques scientifiques, notamment en ce qui concerne l'évaluation de la recherche. Il s'agit de la première recherche à grande échelle des pratiques d'attribution des statuts d'auteur et d'inventeur portant sur plusieurs disciplines, les études antérieures ayant analysé de petits échantillons ou s'étant limitées à une seule discipline. De plus, aucune recherche antérieure n'a utilisé des PBA pour aborder la question de la division du travail au sein des équipes de recherche.

Dans les chapitres suivants, nous présentons la revue de la littérature, une description détaillée de la méthodologie employée, une présentation des résultats obtenus, une discussion de ces résultats, ainsi qu'une conclusion.

Chapitre 1. Revue de la littérature

Introduction

Notre cadre de référence est construit en grande partie à partir des théories de Pierre Bourdieu et Robert K. Merton. La première partie de la revue de littérature visera donc à décrire la structure et le fonctionnement de la communauté scientifique en utilisant notamment les concepts de champs, d'habitus et de capital de Pierre Bourdieu. Les concepts de Bourdieu seront ensuite complétés par une présentation des normes institutionnelles de la science ainsi que de son système de reconnaissance, théorisés notamment par Robert K. Merton. Pour enrichir ce cadre théorique, nous présenterons dans la deuxième partie une revue de la littérature portant sur le concept d'auteur en science, sur ses fonctions, ainsi que sur les normes et pratiques d'attribution du statut d'auteur. De plus, nous nous tournerons dans la troisième partie vers le brevet et le statut d'inventeur et présenterons les critères de brevetabilité des inventions ainsi que les règles et pratiques d'attribution du statut d'inventeur. Dans la quatrième partie, nous examinerons les différentes méthodes d'appariement d'articles et de brevets ayant été utilisées dans des recherches antérieures. Certaines de ces études s'étant intéressées à des problématiques proches de la nôtre, nous en résumerons également les résultats. Nous concluons ce chapitre en résumant les principaux éléments de notre cadre de référence ainsi constitué.

1.1. La structure et le fonctionnement de la communauté scientifique

Dans la présente section, nous décrivons le rôle des pratiques d'attribution des statuts d'auteur et d'inventeur dans la communauté scientifique, et plus spécifiquement dans le fonctionnement de l'économie de la réputation (du capital symbolique) qui la caractérise. Pour ce faire, nous présentons un cadre théorique fondé principalement sur les travaux des sociologues Pierre Bourdieu et Robert K. Merton. Nous définissons les concepts pertinents et les mettons en relation entre eux ainsi qu'avec notre objet d'étude. Ces concepts permettent d'appréhender la communauté scientifique comme un espace social dont la structure dynamique

détermine les pratiques des membres de la communauté, notamment en matière d'attribution des statuts d'auteur et d'inventeur, tout en étant également déterminée par ces pratiques.

1.1.1. Bourdieu et les concepts de champ, d'agent, de capital et d'habitus

1.1.1.1. Champ et agent

Selon Bourdieu (1975), l'espace social est divisé en plusieurs champs distincts et relativement autonomes ayant des intérêts et enjeux qui leur sont propres. Les champs artistiques et politiques en constituent des exemples. Le champ scientifique, pour sa part, est un champ de production de biens symboliques (Bourdieu, 1976) dont l'enjeu principal est l'accroissement des connaissances scientifiques. Les champs sont composés d'agents qui possèdent différentes formes de capital (le concept de capital et ses différentes formes sont présentés plus loin) et peuvent être définis comme

les individus qui, reconnus par leurs pairs, sont en mesure de se soustraire à des intérêts externes, d'apprécier la valeur d'enjeux internes, d'agir selon des normes, se distinguant de la sorte des profanes en vue de faire ce que, comme professionnels, ils sont les seuls capables de faire. (Pinto, s.d.)

Un premier élément important de cette définition de l'agent est la notion de reconnaissance par les pairs. En effet, certaines conditions d'entrée doivent être satisfaites pour qu'un individu puisse être reconnu comme agent par les autres agents du champ (Bourdieu, 1975). Quant à l'appropriation des normes et des enjeux du champ, elle renvoie à la notion d'habitus qui sera définie un peu plus loin.

La structure d'un champ est déterminée par l'ensemble de ses agents ainsi que par la quantité et les types de capital qu'ils possèdent (Bourdieu, 1987). Cette structure est dynamique, puisqu'un champ est un espace de lutte où les agents investissent leurs capitaux afin d'y améliorer leur positionnement et, éventuellement, afin de le dominer (Bourdieu, 1975). Selon Bourdieu (1993), il existe aussi des métachamps qui intègrent plusieurs autres champs (c.-à-d. que les agents des métachamps et les luttes qui y sont à l'œuvre influencent directement les autres champs). Le champ étatique en est un exemple puisque l'État joue notamment un rôle important dans l'attribution des ressources (p. ex. subventions aux entreprises et financement de la recherche) dans les autres champs (Jourdain et Naulin, 2011). Le champ du pouvoir (dont

l'État fait également partie) est un autre exemple de métachamp. Il est composé des agents les plus riches en capital symbolique (concept présenté plus loin) dans les différents champs (Jourdain et Naulin, 2011). Les dirigeants de grandes entreprises et les vedettes du cinéma, par exemple, peuvent y accéder et utiliser leur fortune ou leur renommée pour se positionner dans d'autres champs. Un champ peut aussi comporter plusieurs sous-champs qui y sont pleinement intégrés. Par exemple, bien que la communauté scientifique soit formée d'un ensemble d'agents partageant globalement le même enjeu (la production de connaissances), on y retrouve une multitude de communautés d'agents regroupés autour d'enjeux spécifiques (p. ex. la production des connaissances qui sont d'intérêt pour le groupe), de pratiques et de normes qui leur sont propres. Le champ scientifique peut ainsi être divisé en sous-champs disciplinaires. L'existence de sous-champs et de métachamps témoigne de l'autonomie seulement partielle des champs, puisqu'elle fait en sorte que des luttes externes aux champs soient susceptibles d'en modifier la structure interne (Bonnwitz, 2002).

1.1.1.2. Capital

Bourdieu a développé le concept de capital à partir du capital culturel de Durkheim et du capital économique de Marx, en y ajoutant le capital social comme troisième type de capital principal (Durand, s.d.). Pour Bourdieu, il existe aussi une forme de capital spécifique à chaque champ (p. ex. le capital scientifique) qui s'acquiert seulement à l'intérieur du champ et qui n'a en principe aucune utilité à l'extérieur.

Les agents investissent leur capital dans les divers champs où ils sont actifs afin d'obtenir davantage de capitaux de divers types. Ainsi, les gains en capital sont fonction de la capacité de l'agent à faire les placements les plus rentables (Bonnwitz, 2002) et donc de ses connaissances du champ et de sa capacité à y naviguer. La répartition du capital entre les agents constitue donc à fois la structure du champ ainsi qu'un ensemble de contraintes déterminant les possibilités d'action des agents (Bourdieu, 1986b).

Le capital peut exister sous forme objectivée (p. ex. les ressources financières) ou sous forme incorporée par les agents (p. ex. les connaissances et les savoir-faire). Les sections suivantes présentent brièvement les trois principaux types de capital (économique, social et culturel). Nous présentons aussi le capital scientifique puisqu'il s'agit d'un type de capital

propre au champ scientifique sur lequel porte notre thèse. Nous présentons finalement la notion de capital symbolique, qui englobe toutes les formes de capital dans la mesure où elles sont reconnues par les agents du champ (Bourdieu, 1987).

1.1.1.2.1. Capital économique

Le capital économique désigne l'ensemble des ressources économiques d'un agent, aussi bien son patrimoine matériel que ses revenus et ses moyens de production (terres, usines, travail) (Bourdieu, 1986b). Ce capital permet d'acquérir d'autres types de capitaux, moyennant un certain effort de transformation (Bourdieu, 1986b). Dans le champ scientifique, par exemple, le chercheur investit différentes formes de capital économique (travail, fonds de recherche, équipement) afin de contribuer à l'avancement des connaissances et dans le but d'obtenir en retour un gain en capital scientifique.

1.1.1.2.2. Capital social

Le capital social est un ensemble de ressources disponibles ou potentielles liées à l'appartenance à un groupe, à un réseau durable d'agents qui connaissent et qui reconnaissent les liens qui les unissent (Bourdieu, 1980). Trois éléments déterminent la quantité de capital social que possède un agent, soit l'étendue de son réseau, le type et la quantité de capital détenu par les membres de ce réseau, ainsi que la capacité qu'a l'agent de mobiliser ce capital (Bourdieu, 1980). Le capital social, comme le capital économique, facilite l'acquisition d'autres types de capitaux. Par exemple, en participant à un projet de recherche collaboratif, un agent met à profit son capital social en mobilisant le capital économique de ses collaborateurs afin d'obtenir pour lui-même un gain en capital scientifique en investissant une quantité moindre de son propre capital économique.

Selon Bourdieu (1980), l'existence d'un réseau d'agents est le fruit d'efforts d'instauration et d'entretien, les échanges entre les agents leur permettant de se reconnaître entre eux et de reconnaître leur appartenance au groupe. Ces échanges créent le groupe et en déterminent les limites, faisant des membres du groupe les gardiens des limites de celui-ci. L'inclusion d'un agent « indésirable » constitue un danger qui justifie l'existence d'institutions favorisant les échanges légitimes et limitant les échanges illégitimes (Bourdieu, 1980). Ces institutions assurent ainsi la reproduction du capital social au moyen d'événements (p. ex. les

congrès scientifiques), de lieux (p. ex. les laboratoires de recherche) et de pratiques (p. ex. l'évaluation par les pairs).

1.1.1.2.3. Capital culturel

Selon Bourdieu (1979), le capital culturel peut être incorporé, objectivé ou institutionnalisé. Le capital culturel incorporé est constitué des « dispositions durables de l'organisme » (p. 3). Il s'agit des connaissances d'un agent, de sa culture. Ce dernier ne peut donc pas en hériter et doit l'acquérir au moyen d'un investissement en temps, d'un apprentissage. Le capital culturel objectivé peut prendre diverses formes comme, par exemple, des œuvres d'art, des livres ou des instruments. Bien que sous cette forme le capital culturel puisse être transmis d'un agent à l'autre, son appropriation par l'agent nécessite la possession du capital culturel incorporé nécessaire à son utilisation. Par exemple, un livre (capital culturel objectivé) est utile seulement pour l'agent qui a la capacité de le lire et de le comprendre (capital culturel incorporé). Alors que le capital culturel incorporé n'est pas nécessairement reconnu par les autres agents du champ, le capital culturel institutionnalisé, dont les diplômes scolaires sont un exemple, permet à l'agent d'obtenir une reconnaissance constante et juridiquement garantie des connaissances et compétences acquises (Bourdieu, 1979).

1.1.1.2.4. Capital symbolique

La connaissance et la reconnaissance des différentes formes de capital par les agents ayant incorporé l'habitus du champ (le concept d'habitus est défini ci-dessous) confèrent à celui qui les possède un capital symbolique (Bourdieu, 1987). La répartition du capital symbolique détermine la structure des rapports de force dans la lutte que se livrent les agents pour la domination d'un champ, et le pouvoir dont jouissent les agents dépend inévitablement de la reconnaissance de ce pouvoir par les autres agents. Le capital symbolique est ainsi une « sorte de métacapital puisqu'il est l'objectif suprême de l'action des agents » (Durand, s.d.).

1.1.1.2.5. Capital scientifique

Le capital scientifique est une forme de capital qui est spécifique au champ scientifique et qui en détermine la structure (Bourdieu, 1975). Il s'agit d'une « espèce particulière de capital symbolique qui consiste dans la reconnaissance (ou le crédit) accordée par l'ensemble des

pairs-concurrents au sein du champ scientifique » (Bourdieu, 1997, p. 20). Il « s'acquiert principalement par les contributions reconnues au progrès de la science » (Bourdieu, 1997, p. 23). Le capital confère à l'agent qui le possède de l'autorité scientifique, capacité socialement reconnue « de parler et d'agir légitimement en matière de science » (Bourdieu, 1975, p. 89), et « peut être accumulé, transmis et même reconverti en d'autres espèces sous certaines conditions » (Bourdieu, 1975, p. 93). Comme les agents se font compétition pour la domination du champ scientifique, le capital scientifique a de particulier qu'il est conféré à un agent par ses compétiteurs, qui sont donc « moins enclins à [le] lui accorder sans discussions ni examen » (Bourdieu, 1975, p. 91). De plus, l'autorité scientifique d'un agent est fonction de la valeur et de l'originalité de sa contribution à l'avancement des connaissances aux yeux de la collectivité (Bourdieu, 1997).

L'article scientifique (ou toute autre forme de publication savante), en tant qu'objet matérialisant une contribution à l'avancement des connaissances faite par un chercheur, est un exemple de ce que Bourdieu appelle les biens symboliques (Bourdieu, 1971). Ces derniers sont des objets intellectuels ou culturels qui ont une certaine valeur en capital symbolique dans un champ particulier (Lebaron, 2008).

1.1.1.3. Habitus

L'habitus d'un individu est sa façon de voir le monde, de s'y positionner et d'y positionner les autres agents. Il correspond ainsi à la structure du champ dans sa forme incorporée par les agents sous forme de schèmes de perception, de compréhension et d'action (Bourdieu, 1989). Les agents acquièrent l'habitus spécifique à un champ par l'apprentissage et l'expérience prolongée dans ce dernier. C'est par ce processus d'appropriation de l'habitus du champ que les agents en comprennent la structure, les normes et les enjeux, et qu'ils sont en mesure de choisir les meilleures stratégies d'investissement de leurs capitaux afin d'en acquérir davantage. Le champ et l'habitus s'influencent donc mutuellement. En effet, l'habitus est structuré par le champ dont il est la forme intériorisée par les agents, mais le champ est aussi structuré par l'habitus, ce dernier guidant les actions qui, elles, modifient la structure du champ.

Le développement par un agent de l'habitus spécifique à un champ est influencé par sa trajectoire sociale et donc par d'autres habitus préalablement acquis. En effet, il existe pour

Bourdieu un habitus primaire que les agents acquièrent dès leur enfance. Il est composé, d'une part, de l'habitus familial, acquis dès l'enfance et déterminé par la quantité et le type de capitaux des parents (c.-à-d. de leur classe sociale) et, d'autre part, par l'habitus scolaire, développé au cours du parcours scolaire effectué durant l'enfance. Cet habitus primaire conditionnera par la suite l'acquisition de l'habitus secondaire développé au fil des expériences d'un agent dans différents champs (Bonnewitz, 2002).

Comme l'habitus est basé sur la structure objective du champ commune à tous les agents, il en découle une certaine régularité dans la conduite de ces derniers. Cette régularité laisse néanmoins place à un certain flou, puisque c'est l'esprit pratique des agents qui guide leurs actions, et non le respect de règles ou lois. En effet, comme le dit Bourdieu (1986a), « Il faut se garder de chercher dans les productions de l'habitus plus de logique qu'il n'y en a : la logique de la pratique, c'est d'être logique jusqu'au point où être logique cesserait d'être pratique » (p. 41).

1.1.2. Merton, les normes et le système de reconnaissance de la science

La présente section décrit les intérêts, les enjeux et les normes de la science théorisés par Merton et d'autres par la suite, ainsi que le système de reconnaissance de la science.

1.1.2.1. Les normes institutionnelles de la science

L'éthos de la science moderne, que Merton (1942) décrit comme « un ensemble de valeurs et de normes auxquelles est tenu le chercheur » (p. 268-269, notre traduction), comporte quatre normes (ou idéaux) institutionnelles : l'universalisme, le communalisme, le désintéressement et le scepticisme organisé.

1.1.2.1.1. L'universalisme

L'universalisme fait référence au caractère impersonnel de la science : l'ethnocentrisme et le patriotisme, par exemple, n'ont pas leur place en science, qui est universelle. Ainsi, les contributions à l'avancement des connaissances doivent être évaluées (et conséquemment acceptées ou rejetées) de façon objective, sans tenir compte des caractéristiques individuelles et sociales des personnes qui les ont produites. De la même façon, l'accès à une carrière

scientifique doit être assuré pour tous les talents, indépendamment des origines socio-économiques ou culturelles (Merton, 1942).

1.1.2.1.2. Le communalisme

Selon la norme du communalisme², l'ensemble des connaissances scientifiques appartient à la collectivité. En effet, les découvertes se prêtent peu à la notion de propriété individuelle, qui se limite en fait à la reconnaissance et à l'estime que le chercheur reçoit en échange de sa contribution (Merton, 1942). Cette norme du communalisme de la science se manifeste également d'une part dans l'importance de la diffusion des découvertes, nécessaire à leur inclusion dans l'inventaire des connaissances et, d'autre part, dans le fait que les chercheurs reconnaissent leur dette envers leurs prédécesseurs, notamment en les citant dans leurs propres travaux (Merton, 1942).

1.1.2.1.3. Le désintéressement

Selon la norme du désintéressement, qui ne doit pas être confondue avec l'altruisme, l'éthos scientifique est caractérisé par une passion pour la connaissance, la curiosité et le désir d'améliorer la condition humaine. Selon Merton (1942), si ces caractéristiques se retrouvent généralement chez les chercheurs, ce n'est pas tant parce qu'ils en sont dotés à la base, mais plutôt parce que les mécanismes de contrôle de la science récompensent le désintéressement, et sanctionnent son contraire. Par exemple, la rareté des cas de fraude scientifique serait attribuable à ces mécanismes de contrôle plutôt qu'à une surreprésentation d'individus honnêtes parmi les rangs des chercheurs (Merton, 1942).

1.1.2.1.4. Le scepticisme organisé

La norme du scepticisme organisé est visible dans l'importance accordée à la logique et aux vérifications empiriques en science et a des implications méthodologiques et institutionnelles. Du point de vue méthodologique, elle fait référence à la nécessité que les énoncés scientifiques soient fondés sur la logique et les faits plutôt que sur des croyances. Du point de vue institutionnel, le scepticisme organisé demande, d'une part, que les contributions

² Merton utilise le terme « communisme » dans le texte original, mais le terme « communalisme » est parfois utilisé pour éviter la confusion avec la doctrine politique.

des chercheurs soient rigoureusement examinées par leurs pairs et, d'autre part, que toute croyance soit remise en question et que les chercheurs tentent d'en valider (ou invalider) la pertinence à l'aide de méthodes scientifiques. La science peut ainsi parfois susciter la désapprobation d'autres institutions (p. ex. l'État ou l'Église) dans la mesure où elle remet en question les croyances (Merton, 1942).

1.1.2.2. Le système de reconnaissance de la science

Il découle de la norme du communalisme (selon laquelle les connaissances scientifiques appartiennent à la collectivité) que la principale source de motivation des chercheurs est la reconnaissance de leurs pairs (Zuckerman, 1977). Il s'agit en effet, selon Cole et Cole (1973), de la seule forme de propriété à laquelle les chercheurs ont droit en échange de leur contribution :

Because recognition is so important to scientists, there must be a reward system that identifies and honors scientific excellence wherever it is found. If a scientist desires to acquire "property", he can only do so through recognition by the system, since there are no other legitimate ways to obtain property in science. (p. 46)

Le rôle du chercheur étant de contribuer à l'avancement des connaissances, la communauté scientifique s'est graduellement dotée d'un système de reconnaissance permettant de récompenser les chercheurs qui remplissent le mieux ce rôle (Merton, 1957). Ce système est efficace lorsque ces derniers obtiennent la reconnaissance et l'estime de la communauté et lorsque les chercheurs les plus prometteurs obtiennent les ressources nécessaires pour réaliser leur potentiel (Merton, 1957). Créant ainsi un climat propice à la culture de l'excellence, le système est bénéfique à la fois pour les chercheurs récompensés et pour l'ensemble de la communauté scientifique (Merton, 1973).

1.1.2.2.1. Les formes de reconnaissance

Le système de reconnaissance comporte plusieurs formes de reconnaissance, qui constituent une hiérarchie. Au sommet de cette hiérarchie, l'éponymie est la pratique selon laquelle on associe le nom du chercheur à sa découverte (p. ex. la loi de Zipf, le système de Copernic, la constante de Planck), à une discipline (ex. Comte, père de la sociologie) ou encore à une période (p. ex. l'époque freudienne, l'ère darwinienne) (Cole et Cole, 1973; Garfield,

1983, Merton, 1957). Il s'agit, selon Merton (1957), de la forme la plus permanente et la plus prestigieuse de reconnaissance institutionnalisée en science. Les récompenses éponymiques peuvent aussi être hiérarchisées selon le niveau de prestige qui y est associé. Du plus prestigieux au moins prestigieux, on retrouve le fait d'avoir son nom associé à une période, à une discipline, puis à une découverte (Merton, 1957). Un autre type de récompense hautement prestigieuse est l'octroi de prix honorifiques comme le prix Nobel, la nomination en tant que membre de l'Académie des Sciences, de la Royal Society ou d'autres institutions du même sexe, ou encore l'obtention de titres de noblesse (Merton, 1957). Les chercheurs ayant reçu ces plus grands honneurs constituent l'élite et se situent donc au sommet de la structure stratifiée de la science (Cole et Cole, 1973). Même s'ils ne jouissent pas du même prestige que les gagnants du prix Nobel, les administrateurs de la science siègent aussi au sommet de cette structure grâce à leur grande influence dans la communauté scientifique (Cole et Cole, 1973).

Les places où siègent les individus auxquels ont été décernés les plus grands honneurs sont limitées, ce qui fait en sorte que de nombreux chercheurs dont les accomplissements sont tout aussi importants que ceux des lauréats n'obtiennent pas les mêmes honneurs (Merton, 1968). Ainsi, il existe plusieurs prix honorifiques plus ou moins prestigieux pour souligner les accomplissements de chercheurs s'étant démarqués, mais ne faisant pas partie de l'élite (Cole et Cole, 1973), mais ces honneurs sont également limités en nombre et ne sont donc octroyés qu'à une minorité des chercheurs (Zuckerman, 1970). Par exemple, Cole et Cole (1967) avaient montré que dans le domaine de la physique, moins du tiers des chercheurs (même parmi ceux occupant des postes au sein des départements les plus prestigieux) avaient obtenu au moins un prix dans leur carrière. Ainsi, il existe d'autres mécanismes de reconnaissance, par exemple l'octroi de postes importants au sein de départements prestigieux, ou encore la nomination en tant que rédacteur en chef d'une revue (Zuckerman, 1970). Finalement, on trouve au bas de la hiérarchie des formes de reconnaissance scientifique la visibilité : le fait d'être cité, de voir ses contributions reconnues par les collègues du domaine (Cole et Cole, 1973).

1.1.2.2.2. La stratification sociale en science

La stratification sociale en science constitue un effet systématique de la différenciation et de l'évaluation des chercheurs (Cole et Cole, 1973), qui sont ainsi positionnés dans une

structure qui détermine l'accès aux ressources et aux opportunités (Merton, 1968). Pour Cole et Cole (1973), les deux principaux facteurs qui entrent en jeu dans le positionnement des individus dans la structure hiérarchique sont leur potentiel et leurs contributions. Lorsqu'ils font leur entrée dans la structure, les nouveaux chercheurs n'ont généralement pas encore contribué à la science. Ceux à qui l'on reconnaît le plus de potentiel seront recrutés par les plus prestigieuses institutions, dotées de plus de ressources, ce qui leur donnera un avantage concurrentiel sur les autres chercheurs débutants. C'est ainsi que prend forme la structure stratifiée de la science, au croisement des processus d'autosélection et de sélection sociale (Merton, 1979). Ensuite, les contributions vont conférer aux chercheurs du pouvoir et de l'autorité, et ce, pour une durée indéterminée et indépendamment de leurs accomplissements subséquents (Merton et Zuckerman, 1973). C'est par ce processus que les chercheurs (mais aussi les institutions, les pays et les domaines de recherche) se distinguent entre eux, selon une distribution asymétrique et stratifiée des ressources, de la productivité, de la visibilité et du prestige (Cole et Cole, 1973).

1.1.2.2.3. L'accumulation des avantages et des désavantages

Une des caractéristiques de la stratification sociale en science est qu'elle mène à une accumulation des avantages et des désavantages, et par conséquent à l'accroissement de l'écart entre les privilégiés et les non-privilégiés (Zuckerman, 1998). Le fait d'identifier tôt les chercheurs les plus prometteurs et de leur octroyer un certain avantage sur les autres a des effets à court et à long terme sur leur performance et leur carrière. Lorsqu'il est parfaitement fonctionnel, ce système méritocratique crée un accroissement exponentiel de l'écart entre les privilégiés et les non-privilégiés, les chercheurs les plus performants obtenant plus de ressources qui leur permettront d'être encore plus performants et donc d'obtenir encore plus de ressources, et ainsi de suite (Zuckerman, 1998). Le système deviendrait dysfonctionnel si l'allocation des ressources se basait sur des critères non pertinents (p. ex. le sexe, la religion ou l'ethnie). Le processus d'accumulation des avantages et des désavantages contribue donc à la stratification sociale en science, celle-ci n'étant pas problématique pour le système, mais plutôt essentielle à son optimisation et à l'avancement des connaissances (Zuckerman, 1998).

1.1.2.2.4. L'effet Saint-Matthieu

La valeur des contributions et la reconnaissance que les chercheurs obtiennent en retour sont basées sur le jugement inévitablement subjectif des pairs, ce qui peut causer certains dysfonctionnements dans le système dont, notamment, ce que Merton (1968) appelle l'effet Saint-Matthieu et qu'il définit comme :

The accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark. (p. 53)

Selon Cole et Cole (1973), l'effet Saint-Matthieu serait en fait un produit de la stratification sociale en science constitué de « l'influence de tous les aspects de la stratification sur la réception des idées scientifiques » (p. 214, notre traduction). Il amplifierait donc le processus d'accumulation d'avantages et de désavantages dont il se distingue par son caractère dysfonctionnel. En effet, comme nous l'avons vu au début de la section 1.1.2.2, le système de la reconnaissance permet de maximiser l'efficacité de l'institution scientifique lorsque les chercheurs qui remplissent le mieux leur rôle sont reconnus et lorsque les chercheurs les plus prometteurs obtiennent les ressources nécessaires pour réaliser leur potentiel (Merton, 1957). Or, en surestimant le mérite de certains individus, parfois au détriment d'autres individus, l'effet Saint-Matthieu peut rendre le système méritocratique inefficace. Il peut aussi constituer une violation de la norme d'universalisme de la science voulant que les contributions soient jugées pour elles-mêmes et non en fonction des caractéristiques du chercheur (Merton, 1942).

L'effet Saint-Matthieu peut se produire à divers niveaux, notamment dans l'attribution du crédit pour les découvertes, l'évaluation par les pairs des contributions soumises pour publication, la visibilité de la recherche et des chercheurs, l'attribution de prix honorifiques ou encore l'attribution des ressources. En ce qui concerne l'attribution du crédit pour les découvertes, l'effet Saint-Matthieu peut se manifester dans des cas de découvertes simultanées par des chercheurs indépendants, où il explique le fait que la communauté scientifique attribue généralement le crédit pour la découverte au chercheur le plus connu au détriment de l'autre chercheur. Un phénomène semblable peut être observé au sein d'une équipe de recherche. Dans ce cas, l'effet Saint-Matthieu mène l'observateur externe à surestimer la contribution des chercheurs les plus connus de l'équipe et à sous-estimer celle des chercheurs les moins connus

(Zuckerman, 1998). L'effet Saint-Matthieu peut aussi faire en sorte qu'au sein même de l'équipe de recherche, la contribution de certains membres (p. ex. le directeur du laboratoire) soit surestimée par les autres membres, phénomène particulièrement intéressant dans le cadre de notre recherche. Comme nous le verrons plus loin, l'attribution du statut d'auteur est guidée en partie par la nature et l'ampleur des contributions des différents chercheurs. Ainsi, l'effet Saint-Matthieu peut avoir un impact sur la liste des auteurs (et des inventeurs), soit en ce qui concerne l'attribution ou non du statut à certains membres de l'équipe, soit en ce qui concerne l'ordre dans lequel les noms apparaissent.

Dans le processus d'évaluation par les pairs, l'effet Saint-Matthieu fait en sorte que l'évaluation faite par un chercheur bien connu a un poids plus élevé, susceptible d'être déterminant dans la décision de publier ou non un article. En effet, il est plus difficile pour un rédacteur en chef de revue savante de remettre en cause le jugement d'un évaluateur jouissant d'une excellente réputation dans le domaine. De la même manière, un article dont l'auteur est un chercheur renommé soumis à l'évaluation par les pairs à simple aveugle ou ouverte (c.-à-d. que l'identité de l'auteur est connue par l'évaluateur et vice-versa) aura tendance à être jugé moins sévèrement, soit par crainte de l'affrontement, à cause d'un sentiment d'inaptitude à juger de la qualité du travail d'un chercheur réputé, ou encore parce qu'on présume de la qualité de l'article sans l'avoir examiné minutieusement, et ce, quel que soit le statut de l'évaluateur (Merton et Zuckerman, 1973).

L'effet Saint-Matthieu augmente aussi la visibilité des publications et des communications des chercheurs plus connus par rapport à celles des chercheurs moins connus. Étant donné le rythme rapide auquel de nouveaux articles scientifiques sont publiés, ces derniers s'ajoutant à un inventaire de publications de plus en plus volumineux, il devient plus difficile pour les chercheurs de se tenir à jour. Dans ce cadre, la réputation des auteurs peut être utilisée comme moyen d'identifier les publications les plus importantes dans un domaine, ce qui peut causer là aussi une amplification de l'effet Saint-Matthieu (Merton, 1968). De plus, selon Zuckerman (1970), les prix honorifiques ont tendance à être donnés à ceux qui en ont reçu auparavant, les prix les moins prestigieux étant en quelque sorte un prérequis pour l'obtention de plus grands honneurs.

En somme, l'effet Saint-Matthieu met en lumière le caractère subjectif de l'évaluation de la recherche et de l'attribution du crédit pouvant causer un certain dysfonctionnement du système de reconnaissance de la science.

1.1.2.2.5. L'effet Matilda

Les femmes en sciences peuvent être victimes de l'effet Matilda (Rossiter, 1993) qui désavantage les femmes en les privant d'une part (ou de la totalité) du crédit qui leur revient pour leurs contributions à l'avancement des connaissances, et qui constitue donc, comme l'effet Saint-Matthieu, une violation de la norme d'universalisme. L'effet Matilda s'est manifesté tout au long de l'histoire de la science, alors que de nombreuses femmes ayant participé aux plus grandes découvertes scientifiques des derniers siècles ont vu les grands honneurs (notamment le Prix Nobel) octroyés à leurs collaborateurs masculins, et ce, même lorsque les faits démontrent que le rôle de ces derniers avait été égal ou moindre (Rossiter, 1993). Bien qu'aujourd'hui l'écart entre les hommes et les femmes en science ait été grandement réduit, l'effet Matilda persiste et les hommes sont toujours plus nombreux que les femmes à remporter des prix honorifiques, indépendamment de leur représentation parmi les nominés (Lincoln, Pincus, Koster et Leboy, 2012). Moss-Racusin, Dovidio, Brescoll, Graham et Handelsman (2012) ont observé que les candidats masculins pour un poste de responsable de laboratoire sont perçus comme étant plus compétents, plus employables et méritant un salaire plus élevé que les femmes. Selon Haeussler et Sauermann (2013), le sexe n'est pas un prédicteur significatif de l'obtention du statut d'inventeur. Par contre, l'étude de Lissoni et al. (2013) a montré que les femmes étaient plus souvent exclues de la liste des inventeurs d'un brevet que les hommes. L'effet Matilda se manifeste non seulement dans les inégalités entre les hommes et les femmes dans l'attribution de prix, mais également à tous les niveaux du système de reconnaissance de la science (Larivière, Ni, Gingras, Cronin et Sugimoto, 2013). Il est particulièrement important dans le cadre de notre recherche puisqu'il peut avoir une influence sur la valorisation du travail des femmes au sein d'équipe de recherche, et de ce fait sur l'attribution des statuts d'auteure et d'inventrice.

Malgré les progrès importants accomplis au cours des dernières décennies, les disparités entre hommes et femmes sont encore présentes dans la plupart des sphères de la science. Il y a

une érosion des effectifs féminins à chaque étape de la carrière académique (Shen, 2013), les femmes obtiennent moins de fonds de recherche (Ley et Hamilton, 2008), représentent une plus faible proportion des auteurs (Larivière et al., 2013; West, Jacquet, King, Correll et Bergstrom, 2013), sont moins citées (Larivière et al., 2013), et brevettent moins (Ding, Murray et Stuart, 2006) que les hommes. Les femmes sont aussi moins centrales que les hommes dans les réseaux de collaboration (Ghiasi, Larivière et Sugimoto, 2015). Dans les équipes de recherche, les femmes ont tendance à effectuer des tâches moins importantes que leurs collaborateurs masculins (Macaluso, Larivière, Sugimoto et Sugimoto, 2016).

La participation des femmes au régime de la propriété intellectuelle est un secteur dans lequel les disparités hommes-femmes sont particulièrement grandes, les femmes brevetant beaucoup moins que les hommes (Ding et al., 2006; Murray et Graham, 2007; Sugimoto, Ni, West et Larivière, 2015). Quelques causes possibles de cette disparité ont été proposées dans la littérature, dont l'absence de liens de collaboration avec l'industrie (Meng, 2016), la plus grande difficulté des femmes à concilier le travail et la famille (Murray et Graham, 2007), la persistance des stéréotypes (Murray et Graham, 2007), leurs réserves sur la commercialisation de la science (Murray et Graham, 2007) et leurs attitudes envers le risque et la concurrence (Stephan et El-Ganainy, 2007).

1.1.3. Liens entre les travaux de Merton, Bourdieu, et notre recherche

Nous avons présenté dans les sections 1.1.1 et 1.1.2 les principaux éléments des théories de Bourdieu et Merton. Bien que ces deux perspectives s'inscrivent dans des écoles de pensées distinctes dont les visions du monde social diffèrent, elles ont de nombreux points communs et complémentaires en ce qui concerne l'objet de notre étude. Dans cette section, nous positionnons notre étude dans ce cadre théorique tout en exposant quelques-uns des points de divergence et de convergence entre les perspectives de Bourdieu et de Merton.

Notre thèse s'intéresse aux pratiques de collaboration et d'attribution des statuts d'auteur et d'inventeur dans le cadre de projets de recherche collaboratifs ayant mené à la fois à un article scientifique et à un brevet. Ces actions prennent place dans un espace social théorisé de façon différente par Merton et Bourdieu. Pour Merton, la science est une institution régie par des normes qui sont à la fois des prescriptions de natures méthodologiques et morales (Merton,

1942). Les actions des chercheurs sont ainsi d'abord et avant tout guidées par leur fonction dans le système et un ensemble de normes (présentées à la section 1.1.2.1). C'est en remplissant leur fonction (c.-à-d. en contribuant à l'avancement des connaissances) qu'ils obtiendront la reconnaissance de leurs pairs (Merton, 1957), principale source de motivation des chercheurs (Zuckerman, 1977).

Bourdieu critique cette approche et affirme que « la "fonction" au sens du "fonctionnalisme" de l'école américaine n'est autre chose que l'intérêt des dominants » (Bourdieu, 1975, p. 97). Selon lui, le chercheur n'est pas guidé par un sens moral du devoir découlant des intérêts de la communauté scientifique, mais plutôt par des stratégies visant à acquérir du capital symbolique et éventuellement dominer les autres agents du champ :

Le marché des biens scientifiques a ses lois, qui n'ont rien à voir avec la morale. [...]. Il faut savoir reconnaître pour telles les stratégies qui, dans les univers où l'on a intérêt au désintéressement, tendent à dissimuler les stratégies. (Bourdieu, 1975, p. 100)

Cette divergence fait en sorte que les comportements des chercheurs que nous observons dans notre recherche, par exemple, peuvent être interprétés différemment selon que l'on adopte une perspective mertonienne (le chercheur guidé par les intérêts de l'institution scientifique et son sens moral) ou bourdieusienne (le chercheur guidé par ses intérêts et son sens pratique).

L'article scientifique, autre élément important de notre étude, joue quant à lui un rôle similaire dans les théories de Merton et de Bourdieu. Pour le premier, c'est en contribuant à l'inventaire des connaissances, en publiant ses découvertes que le chercheur obtient la reconnaissance de ses pairs (Merton, 1957) et améliore sa position dans la structure stratifiée de la science et de l'accès aux ressources et opportunités (Cole et Cole, 1973). Pour Bourdieu l'article joue un rôle similaire en ce qu'il s'agit d'un bien symbolique par lequel un chercheur acquiert du capital symbolique et améliore sa position dans le champ (Bourdieu, 1975). Le brevet n'est pas directement abordé par Merton et Bourdieu. Comme nous le verrons dans la section 1.3 portant sur le brevet d'invention et le statut d'inventeur, il s'agit d'un document opérant en principe dans un autre champ (le champ économique) et qui implique un potentiel de gain en capital économique (mais aussi en capital symbolique).

Notre thèse s'intéresse particulièrement à l'attribution des statuts d'auteur et d'inventeur. Selon Bourdieu, tous les choix scientifiques (p. ex. l'objet de la recherche, le lieu de publication) sont « orienté[s] vers la maximisation du profit proprement scientifique, c'est-à-dire de la reconnaissance susceptible d'être obtenue des pairs-concurrents » (Bourdieu, 1976, p. 91). Ainsi, les décisions quant à l'attribution du statut d'auteur peuvent être considérées comme des choix scientifiques résultant de négociations entre les agents du collectif de travail. C'est une lutte à deux niveaux, les agents luttant, d'une part, pour obtenir une place dans les listes des auteurs et, d'autre part, pour obtenir la meilleure position dans la liste (p. ex. être le premier auteur). Dans la perspective mertonienne, l'attribution de la reconnaissance est présentée comme un mécanisme normé dont la fonction est d'accroître la performance du système. Dans la perspective bourdieusienne, nous pouvons concevoir l'attribution des statuts d'auteur et d'inventeur comme étant le produit d'une négociation à forces inégales où chacun des agents impliqués cherche à satisfaire ses propres intérêts.

Notre étude s'intéresse également aux facteurs sociaux influençant l'attribution des statuts d'auteur et d'inventeur, un de ces facteurs étant la position d'un chercheur dans la structure sociale. Merton et Bourdieu sont similaires sur ce point, en ce que les deux auteurs s'entendent sur l'idée que les chercheurs réputés ont tendance à avoir un accès plus facile aux ressources et à l'accroissement de leur réputation. En effet, Bourdieu affirme que « le capital va au capital » (Bourdieu, 1976, p. 94), ce qui fait écho au principe de l'accumulation des avantages et des désavantages (Zuckerman, 1998) et à l'effet Saint-Matthieu (Merton, 1968) qui sont des éléments centraux du cadre mertonien.

Les pratiques de collaboration et d'attribution de la reconnaissance aux chercheurs sont influencées par un ensemble de normes (Merton, 1942) et par la stratification sociale donnant un accès préférentiel aux ressources aux chercheurs les plus performants (Cole et Cole, 1973). Ces pratiques sont aussi le fruit de négociations à forces inégales entre un ensemble d'agents, influencées par les relations de pouvoir en place, l'esprit pratique des chercheurs et leur habitus (Bourdieu, 1986a).

L'approche quantitative utilisée, dont nous discuterons en détail au chapitre 2, permet d'analyser un grand nombre de PBA et d'observer des tendances dans les pratiques des chercheurs, celles-ci étant dans une certaine mesure une manifestation des normes

institutionnelles (Merton, 1942). Bien que les intérêts individuels soient les principales motivations des chercheurs dans la perspective bourdieusienne, ces actions sont néanmoins guidées par leur habitus, or :

[l]'habitus, comme système de dispositions à la pratique, est un fondement objectif de conduites régulières, donc de la régularité des conduites, et si l'on peut prévoir les pratiques [...], c'est que l'habitus est ce qui fait que les agents qui en sont dotés se comporteront d'une certaine manière dans certaines circonstances. (Bourdieu, 1986a, p. 40)

Cette régularité du comportement fait en sorte qu'il est possible de faire sens des tendances observées afin de mieux comprendre la relation entre la structure du champ, les intérêts individuels, et les pratiques de collaboration et d'attribution des statuts d'auteur et d'inventeur.

1.1.4. En résumé

La présente section nous a permis de définir le champ scientifique, ses objectifs, ses normes, et les formes de capitaux qui en définissent la structure. La combinaison des théories de Bourdieu et de Merton nous permet d'élaborer un cadre théorique où les chercheurs sont guidés à la fois par des normes institutionnelles qui découlent des objectifs de la communauté scientifique et qui sont partagées par tous les chercheurs, ainsi que par leurs stratégies individuelles et pragmatiques à des fins de domination du champ. L'habitus du champ, plus ou moins incorporé par les chercheurs, est ce qui guide ces stratégies.

Nous avons également vu que les chercheurs ont comme objectif commun l'avancement des connaissances et que le système est doté de mécanismes permettant de récompenser les chercheurs dont les contributions sont les plus importantes. La reconnaissance que ces derniers obtiennent grâce à leurs contributions est le moteur de leur ascension dans la structure stratifiée de la science. Or nous avons présenté différents facteurs pouvant avoir un impact sur cette reconnaissance, notamment l'effet Saint-Matthieu et l'effet Matilda, ce qui permet de conceptualiser l'attribution du crédit dans le champ scientifique comme un processus qui, bien que guidé par des normes, est également subjectif.

Finalement, nous avons mis en relation les principaux éléments des théories de Bourdieu et de Merton et exposé les liens entre ces cadres théoriques et notre recherche. Nous avons vu que malgré les importantes différences entre ces deux théories, elles ont plusieurs points en

commun, notamment le principe de stratification sociale donnant aux chercheurs dominants un avantage sur les autres.

Dans ces cadres, l'accumulation de la reconnaissance de leurs pairs (Merton, 1957) ou de capital symbolique (Bourdieu, 1987) est essentielle pour la progression des chercheurs dans leur carrière. Ils obtiennent cette reconnaissance en contribuant à l'inventaire des connaissances scientifiques et en apposant leur nom à leurs découvertes afin d'en obtenir le crédit (Merton, 1957). Nous présentons ainsi dans la prochaine section la notion d'auteur scientifique ainsi que les normes et les pratiques d'attribution du statut d'auteur dans les équipes de recherche.

1.2. L'auteur scientifique

Le concept d'auteur est au cœur du système de reconnaissance du champ scientifique, puisqu'il s'agit du lien unissant toute contribution à l'individu (ou aux individus) qui en est responsable. Dans le présent document, le terme auteur est utilisé pour désigner les individus dont le nom apparaît dans la liste des auteurs d'une publication scientifique. Le statut d'auteur n'est pas toujours attribué à l'ensemble des individus ayant participé à une recherche; nous appellerons donc « contributeurs » l'ensemble des individus (qu'ils soient auteurs ou non) impliqués dans une recherche. La présente section se divise en trois sous-sections dans lesquelles nous abordons respectivement 1) les fonctions du statut d'auteur en science, 2) les normes et pratiques qui en guident l'attribution, et 3) les politiques et lignes directrices portant sur l'attribution du statut d'auteur.

1.2.1. Les fonctions du statut d'auteur en science

Selon Birnholtz (2006), les fonctions du statut d'auteur en science sont d'attribuer le crédit, la propriété, et la responsabilité pour les découvertes, ainsi que de permettre l'existence d'une économie de la réputation. Nous décrivons chacune de ces fonctions plus en détail ci-dessous.

1.2.1.1. Le crédit

Comme nous l'avons vu, la reconnaissance des pairs est la principale récompense qu'obtiennent les chercheurs, ce capital symbolique leur permettant ensuite d'obtenir d'autres

récompenses ayant une valeur symbolique, mais aussi financière. Il est donc important pour les chercheurs d'obtenir ce crédit pour leurs contributions et, comme nous le verrons à la section 1.2.2.3, les codes d'éthique d'associations disciplinaires (et, dans une moindre mesure, les politiques éditoriales des revues) mettent l'accent sur l'importance de donner à tous les contributeurs le crédit qui leur est dû. Pour reprendre quelques éléments de la théorie de Bourdieu présentés plus tôt, le crédit est en fait le capital symbolique obtenu par les chercheurs dans leur lutte pour la domination du champ scientifique (Bourdieu, 1976). C'est grâce au statut d'auteur que peut fonctionner le système de reconnaissance de la science octroyant une forme institutionnalisée de crédit (sous forme de prix honorifiques, de postes, par exemple) aux chercheurs qui se sont démarqués comme auteurs d'importantes contributions à l'avancement des connaissances.

1.2.1.2. La propriété et la responsabilité

La notion de propriété liée au statut d'auteur s'applique moins à la science qu'aux autres domaines (p. ex. le domaine littéraire). En effet, les droits de propriété en science se résument à la reconnaissance de la contribution du chercheur aux résultats rapportés (Merton, 1957). Selon Biagioli (2003), seule l'expression originale des découvertes scientifiques, généralement sous forme textuelle, est la propriété des auteurs, ce qui leur permet d'être protégés contre le plagiat.

La notion de responsabilité va de pair avec celle de propriété et a une importance particulière en science (Birnholtz, 2006; Rennie et Flanagin, 1994). Dans un système de reconnaissance basé sur la contribution des chercheurs à l'avancement des connaissances, le bon fonctionnement du système dépend à la fois de sa capacité à récompenser le mérite et le respect des normes et de sa capacité à punir, ou du moins à ne pas récompenser, le démerite et le non-respect des normes. La fraude scientifique (p. ex. la fabrication ou la falsification des résultats de recherche, ou encore le plagiat) est un exemple de non-respect des normes de la science ayant un impact négatif sur la carrière des chercheurs impliqués (Azoulay, Bonatti et Krieger, 2015; Lu, Jin, Uzzi et Jones, 2013; Mongeon et Larivière, 2016), leurs institutions et même l'ensemble de leur discipline (Azoulay, Furman, Krieger et Murray, 2015). Bien qu'ils demeurent relativement rares par rapport à l'ensemble de la recherche scientifique, les cas de fraudes scientifiques ont fortement augmenté au cours des vingt dernières années (Steen, Casadevall et

Fang, 2013). Cela a contribué à l'intérêt des éditeurs pour l'élaboration de lignes directrices sur l'attribution du statut d'auteur qui, comme nous le verrons plus loin, mettent particulièrement l'accent sur la responsabilité des chercheurs quant à la validité du contenu des articles (Resnik et Master, 2011).

1.2.1.3. L'économie de la réputation

Les notions de réputation et d'économie de la réputation (Whitley, 2000) sont intimement liées à celle du capital symbolique. En effet, le statut d'auteur, entre autres, fournit du capital scientifique au chercheur (Birnholtz, 2006). Dans la mesure où ce capital est connu et reconnu par les autres chercheurs, il est converti en capital symbolique et confère ainsi au chercheur un certain pouvoir dans le champ (Bourdieu, 1987). L'économie de la réputation se manifeste aussi dans la stratification sociale du champ scientifique ainsi qu'à tous les niveaux de la hiérarchie des récompenses, de la simple citation à l'obtention du prix Nobel, forme institutionnalisée du capital symbolique et de la réputation du chercheur (Cole et Cole, 1973).

1.2.2. Les normes et pratiques d'attribution du statut d'auteur en contexte collaboratif

La recherche scientifique est une activité de plus en plus collaborative. Dans l'ensemble des disciplines, le nombre moyen d'auteurs par article a augmenté de façon continue au cours du dernier siècle (Larivière et al., 2015; Wuchty et al., 2007). Cette augmentation de la collaboration relève de nombreux facteurs, notamment du fait que la recherche est de plus en plus complexe et de plus en plus interdisciplinaire, nécessitant des expertises variées (Katz et Martin, 1997). De plus, les chercheurs sont portés à collaborer pour partager les coûts de plus en plus élevés de la recherche, notamment des équipements (Katz et Martin, 1997).

L'attribution du statut d'auteur peut ainsi devenir très complexe lorsqu'une découverte scientifique est le fruit du travail de nombreux individus dont les statuts, rôles et contributions sont variables (Rennie et Yank, 1998; Smith et Williams-Jones, 2012). Il faut non seulement déterminer qui mérite sa place dans la liste des auteurs et qui sera (peut-être) remercié, mais aussi déterminer l'ordre des auteurs, qui se veut généralement un reflet de la part plus ou moins importante du crédit qui revient à chacun (Pontille, 2016; Zuckerman, 1968). Dans le même

ordre idée, lorsqu'un article est signé par plusieurs individus, il est difficile, voire impossible, pour un observateur externe de déterminer qui a fait quoi et la part du crédit qui revient à chacun des auteurs (Rennie et al., 1997). Essentiellement, la signification de la liste des auteurs d'un article donné dépend de deux éléments : 1) la division du travail entre les individus impliqués dans la recherche (qui a fait quoi?) et 2) les décisions qui ont été prises au niveau de l'attribution du statut d'auteur de l'article (quels noms apparaîtront dans la liste des auteurs, et dans quel ordre?).

1.2.2.1. Les formes de collaboration et de division du travail

Laudel (2002) a étudié les formes de collaboration et de division du travail de recherche à l'aide de 101 entrevues semi-structurées auprès de chercheurs. Elle distingue six types de collaboration. Le premier est la collaboration impliquant une division du travail créatif (*collaboration involving a division of labour*). L'aspect créatif du travail est important, puisque l'auteure distingue les contributions intellectuelles des contributions techniques. Dans ce type de collaboration, les collaborateurs ont un objectif de recherche commun et effectuent une contribution intellectuelle importante dans le but d'atteindre cet objectif. Le deuxième type de collaboration est la collaboration de service (*service collaboration*), où un ou plusieurs chercheurs engagés dans une collaboration impliquant une division du travail font appel à des collaborateurs qui exécuteront des tâches routinières ou des tâches techniques pouvant nécessiter une expertise particulière. Dans le troisième type de collaboration, le collaborateur n'effectue pas de tâches liées à la réalisation du projet, mais fournit du matériel ou de l'équipement à l'équipe de recherche (*provision of access to research equipment*). Le quatrième type de collaboration est la transmission de savoir-faire (*transmission of know-how*), que Laudel (2002) décrit comme l'activité non créative de transmission d'une information mémorisée pouvant être utile pour la réalisation d'une recherche. Les deux derniers types de collaboration ne sont pas liés au processus de recherche lui-même. Il s'agit de 1) la stimulation et l'encouragement réciproque (*mutual stimulation*) où les interactions informelles entre les chercheurs leur permettent de développer leurs idées, par exemple, et de 2) la fourniture de commentaires ou de rétroaction sur un travail effectuée (*trusted assessorship*).

Comme nous le verrons plus en détail dans la prochaine section (1.2.2.2) ces formes de collaboration ne mènent pas toutes à l'obtention du statut d'auteur. En fait, selon l'analyse de Laudel (2002), seules les collaborations impliquant une division du travail créatif et, plus rarement, les collaborations de service mènent à l'obtention du statut d'auteur.

Alors que la typologie des formes de collaboration de Laudel (2002) est axée sur la nature des contributions des collaborateurs, Subramanyam (1983) propose quatre types de collaboration axés sur les statuts hiérarchiques des individus impliqués, soit 1) la collaboration professeur-étudiant, 2) la collaboration entre collègues, 3) la collaboration superviseur-assistant, et 4) la collaboration chercheur-consultant. Les collaborations ne diffèrent pas seulement selon leurs formes ou selon le type de relation entre les chercheurs impliqués, mais également selon les raisons pour lesquelles des individus décident de collaborer. Selon Melin (2000), les chercheurs peuvent collaborer pour accéder à des compétences particulières, à des données ou à de l'équipement; ils peuvent aussi collaborer pour des raisons sociales (p. ex. pour rendre service à un ami) ou encore dans le cadre d'une relation superviseur-étudiant. Ces différents motifs de collaborer peuvent mener à différentes formes de collaboration.

Le nombre d'individus requis pour mener à bien un travail de recherche peut aussi avoir une influence sur le mode d'organisation du travail (et donc les formes de collaboration) qui sera privilégié. Par exemple, Walsh et Lee (2015) ont montré que plus le nombre d'individus impliqués dans une recherche est élevé, plus le groupe a tendance à être organisé selon une structure bureaucratique. Ce type de structure est caractérisé par une plus importante division du travail, la spécialisation et la standardisation des tâches, la présence de relations hiérarchiques, et la décentralisation de la prise de décision (Walsh et Lee, 2015).

Les formes de collaboration privilégiées par les chercheurs ne sont pas seulement déterminées par leurs préférences personnelles, mais aussi par la nature des travaux de recherche effectués. Celle-ci est définie par l'objectif institutionnel de production de connaissances dont découlent des impératifs techniques et moraux (Merton, 1942). Cet objectif commun ou, selon Bourdieu (1976), les intérêts des agents dominants définissent les enjeux scientifiques (les objets de recherche) qui sont d'intérêts pour les agents du champ qui recherchent les investissements les plus payants en capital symbolique (Bourdieu, 1976). L'objet de recherche définit les ressources humaines et matérielles ainsi que les savoirs et savoir-faire requis, ce qui peut en

quelque sorte imposer aux chercheurs certains types de collaboration et de division du travail. Cette idée est bien illustrée par la notion de culture épistémique (Knorr Cetina, 1999) qui désigne la particularité des pratiques de création des connaissances dans une discipline.

1.2.2.2. L'attribution du statut d'auteur

Toute collaboration est caractérisée par une hiérarchie et une division des rôles où chaque individu occupe une position particulière (Gläser, 1963). Celle-ci détermine, notamment, la nature et l'importance de la contribution respective de chacun, ainsi que les rapports de forces qui donneront à certains individus un plus grand poids dans les processus décisionnels, incluant les décisions portant sur l'attribution du statut d'auteur. Ces décisions ont essentiellement deux volets : 1) déterminer qui sera auteur de l'article et 2) déterminer dans quel ordre les noms des auteurs seront listés. Nous décrivons ces deux volets dans les sections suivantes, et nous présentons ensuite quelques typologies d'auteurs identifiées lors de notre revue de la littérature.

1.2.2.2.1. Qui est auteur?

Alors qu'en littérature la notion d'auteur est indissociable de l'acte d'écriture, ce n'est pas toujours le cas en science, où d'autres types de contribution peuvent mener au statut d'auteur. De plus, l'écriture à elle seule n'est pas nécessairement suffisante pour l'acquisition de ce statut, sa valeur étant « dépendante de son articulation à d'autres tâches qui, associées les unes aux autres, rendent possible l'accès à la signature » (Pontille, 2006, p. 84). Il est aussi possible de signer sans écrire lorsque d'autres tâches comme la conception de l'étude, la collecte des données ou la revue de la littérature sont perçues comme des activités qui justifient, à elles seules, l'obtention du statut d'auteur (Pontille, 2006). Ainsi, l'une des caractéristiques importantes de l'attribution du statut d'auteur scientifique est la pluralité des types de contributions pouvant mener à l'obtention de ce statut.

Quels sont donc les types de contribution qui sont (ou qui devraient) être récompensés par l'octroi du statut d'auteur? Selon, Laudel (2002), les collaborations impliquant une division du travail créatif mènent généralement à l'obtention du statut d'auteur pour les individus impliqués. Ce lien entre la contribution substantielle et créative et le statut d'auteur soulève peu de débats dans la littérature. Cependant, la question de l'octroi du statut d'auteur à ceux dont la

contribution à la recherche est de nature technique, routinière ou de moindre ampleur soulève davantage de questions.

Hagstrom (1964) aborde la question de l'attribution du statut d'auteur pour des contributions techniques en distinguant la collaboration traditionnelle de la collaboration moderne³. Traditionnellement, le technicien était généralement un travailleur peu qualifié exécutant des tâches simples. Il n'était pas engagé dans une quête de solutions à des problèmes scientifiques et résolvait des problèmes qui lui étaient soumis par d'autres en échange de capital économique (et non de capital scientifique) (Hagstrom, 1964). C'est donc le chercheur qui détenait l'autorité de définir la nature, l'étendue et le sens du travail du technicien à son service. Dans la collaboration moderne, le technicien est plutôt un professionnel qualifié qui exécute des tâches complexes que le chercheur qui l'emploie ne saurait pas nécessairement effectuer par lui-même, ce qui mène à l'attribution croissante du statut d'auteur aux techniciens (Hagstrom, 1964). Cette professionnalisation est exemplifiée dans l'anecdote, racontée par Knorr Cetina (1999), d'un chercheur qui, malgré une maîtrise de la théorie, s'est avéré incapable d'exécuter les manipulations requises dans son domaine de recherche. Cela dit, cette tendance n'est pas généralisée et les contributions techniques mènent, encore aujourd'hui, moins souvent au statut d'auteur que les autres types de contributions (Haeussler et Sauermann, 2015; Hong, 2008, Larivière et al., 2016; Laudel, 2002).

Le choix d'attribuer le statut d'auteur à un individu ayant fait une contribution technique peut aussi être influencé par le statut de l'individu qui accomplit la tâche. Un technicien salarié n'obtiendra généralement pas le statut d'auteur (Pontille, 2016), mais un collègue chercheur ayant effectué une contribution similaire pourra obtenir le statut d'auteur comme forme de paiement pour un service rendu (Pontille, 2016; Shibayama, Walsh et Baba, 2012).

L'importance d'une découverte peut aussi avoir un effet sur l'attribution du statut d'auteur. Par exemple, Jabbehdari et Walsh (2017) ont observé que les auteurs ayant effectué des contributions mineures ou techniques à la recherche ont plus souvent tendance à être exclus de la liste des auteurs lorsqu'il s'agit d'articles hautement cités, ce qui indique que les chercheurs

³ Bien que l'article de Hagstrom soit paru en 1964, son idée de collaboration moderne peut être considérée comme valide encore aujourd'hui en ce qu'il fait référence au fait que les techniciens sont de plus en plus qualifiés.

sont peut-être moins enclins à partager le crédit pour les découvertes qu'ils considèrent majeures et dont le gain potentiel en capital symbolique est plus important. Dans la même étude, Jabbehdari et Walsh (2017) ont également observé que la probabilité que des individus ayant effectué des contributions mineures ou techniques soient exclus de la liste des auteurs est plus faible dans les collaborations interinstitutionnelles comparativement aux recherches impliquant une seule institution.

La relation entre la nature et l'ampleur d'une contribution et l'obtention du statut d'auteur est également guidée par des normes implicites liées à la culture disciplinaire. Une même contribution peut mener au statut d'auteur dans une discipline donnée, mais non dans une autre, et des pratiques courantes dans une discipline peuvent même être considérées non éthiques dans d'autres (Bozeman et Youtie, 2016). Par exemple, en droit, le superviseur de thèse, bien qu'ayant contribué à la conception et au développement du projet de son étudiant, ne signera pas l'article (Pontille, 2004). En sociologie, l'écriture est la contribution la plus essentielle, la notion d'auteur étant semblable à celle qui prévaut en littérature (Pontille, 2004). Une distinction existe toutefois entre la sociologie française et la sociologie américaine, où d'autres activités comme la collecte et l'analyse des données peuvent aussi mener à l'obtention du statut d'auteur (Pontille, 2004). Cela montre que les cultures disciplinaires ne sont souvent pas indépendantes des contextes locaux. En physique, on retrouve le cas particulier des mégacollaborations dont les articles sont signés par des centaines d'auteurs, phénomène que Cronin (2001) appelle « *hyperauthorship* ». Dans ce contexte, c'est l'appartenance au groupe qui justifie l'obtention du statut d'auteur, et non une contribution spécifique à un article spécifique (Biagioli, 2003; Birnholtz, 2006; Knorr Cetina, 1999). Ce modèle d'attribution du statut d'auteur évacue la notion de responsabilité individuelle au profit d'une responsabilité collective, et où la crédibilité d'un article est corrélée au nombre de signatures (Biagioli, 2003).

Dans certains cas, il peut exister une relation très faible, voire inexistante entre la contribution effectuée et l'obtention du statut d'auteur. Il est fréquent, par exemple, que des auteurs honorifiques (guest ou honorary authors) signent des articles alors que leur contribution était très superficielle ou encore nulle, alors que d'autres, les auteurs fantômes (ghost authors), ne signent pas alors qu'ils ont réalisé une partie importante du travail (Bates, Anić, Marušić et Marušić, 2004; Flanagin et al., 1998; Godlee, 1996; Huth, 1986b; Katz et Martin, 2007; Mowatt

et al., 2002; Sismondo, 2009; Wager, 2007a; Wislar, Flanagin, Fontanarosa et Deangelis, 2011). De plus, il est fréquent que les individus ainsi exclus de la liste des auteurs soient remerciés dans l'article (Cronin et Overfelt, 1994; Laudel, 2002). Les différences disciplinaires observées dans le nombre d'auteurs moyen par article sont d'ailleurs grandement réduites lorsqu'on tient compte à la fois des auteurs et des individus remerciés (Paul-Hus et al., 2017), ce qui suggère que ces différences ne découlent pas seulement du nombre de chercheurs ayant contribué à la recherche, mais peut-être aussi de différences disciplinaires dans les pratiques d'attribution du statut d'auteur, possiblement en ce qui concerne la valorisation des contributions des techniciens ou des assistants de recherche. Il importe ici de souligner la distinction entre l'auteur fantôme et le remercié. Le premier désigne l'individu qui n'a pas obtenu le statut d'auteur alors que sa contribution aurait normalement justifié une telle reconnaissance. Les remerciements sont plutôt utilisés, en principe, pour souligner une contribution dont l'ampleur ne justifie pas l'attribution du statut d'auteur (Kassirer et Angell, 1991). L'existence des auteurs honorifiques, des auteurs fantômes et des contributeurs remerciés met en évidence le fait que la liste des auteurs peut dans certains cas mener à une sous-estimation (Moulopoulos, Sideris et Georgilis, 1983; Paul-Hus, Mongeon, Sainte-Marie et Larivière, 2017; Rennie et Flanagin, 1994) ou à une surestimation (King, 2000) du nombre de chercheurs ayant contribué à une recherche.

1.2.2.2.2. L'ordonnancement des auteurs

L'attribution du statut d'auteur ne consiste pas seulement à nommer les auteurs, mais également à déterminer dans quel ordre les noms apparaîtront, cet ordre se voulant le reflet de l'importance de la contribution respective des auteurs (Zuckerman, 1968). Dans les paragraphes suivants, nous faisons un survol des différents types d'ordonnancement qui sont utilisés par les chercheurs.

Les noms des auteurs peuvent être listés en ordre alphabétique afin de distribuer le crédit entre les membres de l'équipe de façon « égalitaire » (Zuckerman, 1968, p. 278, notre traduction). L'ordre alphabétique est utilisé couramment dans certaines disciplines comme les mathématiques, l'économie, la gestion et les finances (Waltman, 2012), mais aussi, plus ou moins fréquemment, dans les autres disciplines (Mongeon et al., 2017; Waltman, 2012). Dans le cas des mégacollaborations en physique, par exemple, comme on ne tient pas compte de la

contribution respective des individus dans l'attribution du statut d'auteur, les auteurs sont généralement listés en ordre alphabétique (Knorr Cetina, 1999). Il arrive également que les auteurs soient d'abord regroupés (p. ex., par pays, ou par institutions) et soient ensuite listés en ordre alphabétique à l'intérieur de chacun des groupes⁴. Le fait que l'usage de l'ordre alphabétique soit assez fréquent dans certaines disciplines sans être généralisé apporte une certaine ambiguïté puisqu'il se peut que le premier et/ou le dernier auteur soient perçus comme ayant effectué une contribution plus importante que les autres auteurs (Zuckerman, 1968). Il est aussi possible que l'ordre alphabétique soit le fruit du hasard et que les auteurs soient perçus comme ayant effectué une contribution équivalente alors que ce n'est pas le cas (Zuckerman, 1968).

Les auteurs peuvent aussi être ordonnés en fonction de l'importance de leur contribution, le premier auteur étant typiquement celui qui a effectué la plus importante contribution (Zuckerman, 1968). Il s'agit d'une forme d'ordonnement souvent utilisée dans les sciences humaines et sociales où l'écriture est la contribution la plus importante, et parfois même nécessaire pour obtenir le statut d'auteur (Pontille, 2004). Une des limites de l'ordonnement selon l'importance de la contribution est qu'il ne peut y avoir qu'un seul auteur à chaque position de la liste. Pour contourner cette limite, de plus en plus d'articles mentionnent que certains auteurs ont effectué une contribution égale (Hu, 2009). Selon Zuckerman (1967), il arrive aussi que les chercheurs plus reconnus cèdent la première position à leurs collaborateurs moins reconnus et dont la contribution à l'article a été plus faible.

Une variante de l'ordre basé sur la contribution décrit au paragraphe précédent est souvent utilisée dans certaines disciplines des sciences naturelles ou biomédicales où le travail de recherche s'effectue typiquement dans des laboratoires. La différence est que les deux extrémités de la liste des auteurs (la première et la dernière position) sont les plus importantes (Pontille, 2004, 2006). Le premier auteur, qui est généralement un doctorant ou un chercheur postdoctoral responsable de l'expérimentation. Le dernier auteur est généralement le directeur du laboratoire; celui qui détient l'autorité. Les autres contributeurs sont listés entre ces deux

⁴ Voir l'article suivant pour un exemple : <http://science.sciencemag.org/content/338/6114/156>

extrêmes en fonction de l'importance de leur contribution. Les contributions les moins importantes se retrouvant au centre de la liste (Pontille, 2004, 2006).

Zuckerman (1968) décrit un dernier type d'ordonnancement que nous appellerons « ordre alphabétique partiel » et selon lequel un sous-ensemble des auteurs est en ordre alphabétique afin de mettre en évidence la contribution particulière d'un auteur (normalement le premier ou le dernier). Cette idée a été reprise par Mongeon et al. (2017) qui ont mesuré la prévalence des cas où seul un sous-ensemble des auteurs situés au centre de la liste est en ordre alphabétique. L'ordre alphabétique est ici utilisé pour créer une distinction entre les auteurs principaux (au début de la liste), les auteurs ayant joué un rôle de supervision (à la fin de la liste) et les autres (au milieu de la liste, en ordre alphabétique), dont la contribution est généralement moindre (Mongeon et al., 2017).

Bien que les modes d'ordonnancement des auteurs présentés ci-dessus soient les plus fréquemment utilisés, il arrive aussi que les chercheurs utilisent d'autres critères, parfois arbitraires ou ludiques, pour déterminer l'ordre des auteurs (Cabanac, 2015).

1.2.2.2.3. Typologies d'auteurs

Certains auteurs ont proposé de diviser les auteurs en trois groupes. Par exemple, Perneger et al. (2017) ont analysé les contributions individuelles des auteurs de 1 139 articles du domaine biomédical et distingué trois types de profils : 1) le penseur (*thinker*), qui conçoit l'étude, obtient le financement, et révisé l'article, 2) le soldat, qui fournit du matériel, du soutien administratif ou technique, ou participe à la collecte des données, et 3) le scribe, qui analyse et interprète les résultats et rédige l'article. De leur côté, Baerlocher, Newton, Gautam, Tomlinson et Detsky (2007) distinguent les trois types d'auteurs suivants :

1. Les auteurs principaux (*primary*) : ils participent à la planification et à la réalisation de l'étude ainsi qu'à la rédaction de l'article. Ils sont en mesure d'expliquer l'ensemble des résultats de l'étude, et se portent garants de l'exactitude des informations rapportées dans l'article. Ils apparaissent généralement au début de la liste des auteurs.
2. Les auteurs seniors/superviseurs (*senior/supervisory*) : ils participent à la planification et à la supervision de l'étude ainsi qu'à la révision substantielle de l'article. Comme les auteurs principaux, ils sont en mesure d'expliquer l'ensemble des résultats de l'étude et

se portent aussi garants de l'exactitude des informations rapportées dans l'article. Ils apparaissent généralement à la fin de la liste des auteurs.

3. Les auteurs contributeurs (*contributing*)⁵ : Ils ont fait une contribution substantielle à la recherche, mais ils ne remplissent pas les critères pour l'obtention du statut d'auteur principal ou superviseur décrit ci-dessus. Ils n'ont pas à être en mesure d'expliquer l'ensemble des résultats de l'étude et ne sont pas responsables de l'exactitude des informations rapportées dans l'article. Ils apparaissent généralement au milieu de la liste des auteurs.

Un parallèle peut être fait entre les typologies de Perneger et al. (2017) et de Baerlocher et al. (2007), les scribes s'apparentant aux auteurs principaux, les penseurs aux auteurs superviseurs, et les soldats aux auteurs contributeurs. De plus, les soldats et les auteurs contributeurs s'apparentent aux techniciens auxquels font référence Shapin (1989) et Hagstrom (1964), entre autres. Il est important toutefois de noter que le technicien selon la définition de Hagstrom (1964) est typiquement un salarié n'ayant pas d'intérêt dans les enjeux scientifiques, alors que les soldats et les auteurs contributeurs peuvent être d'autres chercheurs impliqués dans ce que Laudel (2002) appelle une collaboration de service.

1.2.3. Politiques éditoriales des revues savantes

Les sociétés savantes et les rédacteurs en chef de revues ont développé des lignes directrices en matière d'attribution du statut d'auteur. Les efforts déployés en ce sens sont une réponse à trois tendances, soit 1) le manque de standardisation des pratiques (Osborne et Holland, 2009), 2) l'augmentation du nombre moyen d'auteurs par article (Larivière et al., 2015; Wuchty et al., 2007), et 3) l'augmentation des cas de fraude scientifique (Pontille, 2016; Steen, 2011) rappelant l'importance de l'attribution de la responsabilité aux auteurs (Rennie et Flanagan, 1994). Les deux premières tendances sont particulièrement marquées en sciences naturelles, en génie et en sciences biomédicales (Wuchty et al., 2007), et la dernière surtout dans le domaine biomédical (Fang, Steen et Casadevall, 2012). Il est donc peu surprenant que les politiques sur l'attribution du statut d'auteur soient très fréquentes dans ces disciplines (Bošnjak

⁵ Dans leur étude, Mongeon et al. (2017) reprennent la typologie de Baerlocher et al. (2007), mais utilisent le terme « auteur périphérique » plutôt que « auteur contributeur ».

et Marušić, 2012; Wager, 2007b). En effet, l'analyse des politiques éditoriales de 252 revues du domaine biomédical effectuée par Wager (2007b) a démontré que 59 % d'entre elles contenaient des directives concernant l'attribution du statut d'auteur. C'était également le cas de 53 % de l'échantillon aléatoire de 110 revues en sciences naturelles, génie et recherche biomédicale analysé par Bošnjak et Marušić (2012). À l'inverse, très peu de revues en arts et humanités (6 % des revues analysées par Bošnjak et Marušić, 2012) ont des politiques sur l'attribution du statut d'auteur, ce qui est peu surprenant puisqu'il s'agit de disciplines où le modèle classique de l'auteur unique est encore la norme (Cronin, 2005). Avec environ un tiers des revues ayant des politiques sur l'attribution du statut d'auteur, les sciences sociales (Bošnjak et Marušić, 2012) et la bioéthique (Resnik et Master, 2011), champ multidisciplinaire au croisement de la science et de la philosophie, se situent entre ces deux extrêmes.

Il est rare que les revues créent leurs propres règles en matière d'attribution du statut d'auteur. La plupart font référence aux codes d'éthique d'associations professionnelles ou aux politiques des maisons d'édition, qui renvoient souvent aux recommandations de l'International Committee of Medical Journal Editors (ICMJE). Ces dernières sont donc une référence majeure en la matière. En effet, même si leur adoption n'est pas généralisée dans le domaine biomédical (Bhopal et al., 1997; Bošnjak et Marušić, 2012; Wager, 2007b, 2009), les recommandations de l'ICMJE exercent une influence allant au-delà de cette discipline. Par exemple, le Committee on Publication Ethics (COPE) ainsi que l'European Association of Science Editors (EASE) y font référence, tout comme certaines maisons d'édition commerciales telles qu'Elsevier, Springer et Wiley-Blackwell.

Selon les recommandations de l'ICMJE (voir Annexe 1, p. ix), l'obtention du statut d'auteur nécessiterait à la fois une contribution substantielle à la recherche et au processus d'écriture (minimalement, la révision du manuscrit et l'approbation de sa version finale). Les auteurs doivent aussi s'engager à faciliter l'enquête et la résolution de toute question liée à l'intégrité du travail, cette dernière étant une préoccupation à la fois éthique et légale pour les éditeurs (Huth, 1986b; Kassirer et Angell, 1991; Leash, 1997). D'ailleurs, un livre blanc du Council of Science Editors (2012) indique que « l'ultime raison d'identifier les auteurs et les autres contributeurs est d'assigner la responsabilité pour le travail rapporté » (p. 21, notre traduction). Cette importance particulière que les éditeurs accordent à la notion de responsabilité

se reflète également dans les études d'Osborne et Holland (2009), Bebeau et Monson (2011) ainsi que Resnik et Master (2011), cette dernière étude montrant que plus de 70 % des politiques de revues mentionnent la responsabilité des auteurs, alors que c'est le cas de seulement 13 % des codes d'éthique des sociétés savantes, qui mettent plutôt l'accent sur la juste attribution du crédit aux chercheurs pour leurs contributions.

Malgré tout, il n'y a pas de consensus sur ce qui constitue une contribution substantielle justifiant l'attribution du statut d'auteur (Claxton, 2005). Même si elles mentionnent parfois le type de contribution nécessaire pour obtenir le statut d'auteur, les lignes directrices restent généralement vagues en ce qui concerne l'ampleur de la contribution requise, stipulant simplement qu'elle doit être substantielle. De plus, les éditeurs eux-mêmes se déclarent peu convaincus de l'efficacité des lignes directrices et de leur capacité à réduire la prévalence des auteurs honorifiques ou fantômes (Wager, Fiack, Graf, Robinson et Rowlands, 2009). En somme, bien que les lignes directrices en matière d'attribution du statut d'auteur puissent aider les chercheurs dans le processus de nomination et d'ordonnement des auteurs, elles semblent avoir un effet relativement limité sur les pratiques.

Afin de pallier les problèmes d'attribution du crédit et de la responsabilité associés à l'augmentation du nombre d'auteurs par article, Rennie, Yank et Emanuel (1997) ont proposé un modèle qui rejette complètement la notion d'auteur et qui la remplace par les notions de contributeur et de garant. Dans ce modèle, les contributeurs sont tous ceux qui ont participé à la recherche, indépendamment de l'ampleur et de la nature de leur contribution, et les garants sont responsables de l'intégrité de l'ensemble du projet. Afin de savoir quelle a été leur contribution respective, chaque contributeur doit fournir une description du travail effectué (voir exemple à l'Annexe 4, p. xxii).

Ce modèle est appliqué, souvent partiellement, par de plus en plus de revues (p. ex. *The Lancet*, *Journal of the American Medical Association*, *British Medical Journal*, *Nature* et *PLoS*), qui exigent que les contributions de chacun des auteurs soient indiquées. De même, le livre blanc du Council of Science Editors (2012), cité plus tôt, indique que les éditeurs devraient exiger que les auteurs et contributeurs identifient leurs contributions respectives afin de fournir cette information au lecteur. La description des contributions se révèle donc un élément parfois mentionné, notamment par les règles de l'ICMJE, mais pas toujours exigé. Or, Rennie, Flanagin

et Yank rapportaient dans un éditorial publié en 2000 dans la revue *JAMA* que très peu d'auteurs sont portés à préciser leur contribution lorsque cela n'est pas exigé et que, conséquemment, la revue allait dorénavant exiger que les contributions des auteurs soient déclarées (Rennie, Flanagin et Yank, 2000).

L'information que ces déclarations de contributions fournissent est toutefois limitée et n'est pas parfaitement fiable. En effet, seulement le type de contribution est indiqué, ce qui ne permet pas de connaître l'ampleur du travail effectué. De plus, selon Ilakovac, Fister, Marusic et Marusic (2006), le fait que la déclaration des contributions est généralement faite par un seul des auteurs en faisant appel à sa mémoire peut induire des imprécisions dans l'information rapportée. De plus, le tiers des répondants à une enquête effectuée auprès de plus de 6 000 chercheurs sont d'avis que les déclarations de contributions ne sont pas plus informatives que l'ordre des auteurs quant aux contributions effectuées par chacun (Sauermann et Haeussler, 2017).

1.2.4. En résumé

Dans cette section, nous avons présenté les fonctions du statut d'auteur en science, soit d'attribuer le capital scientifique et la responsabilité aux chercheurs pour leur travail, et de permettre l'existence d'une économie de la réputation. Nous avons ensuite discuté de l'augmentation du nombre moyen d'auteurs par article due, notamment, à la collaboration accrue en science, ainsi que de la tension sur le système de reconnaissance scientifique que cela engendre. Cette tension est due en partie au fait que plus la liste d'auteurs d'un article est longue, plus il est difficile de déterminer qui a fait quoi et quelle part du crédit revient à chacun des auteurs.

Après avoir discuté des différentes formes de collaboration, nous avons fait un survol des études antérieures sur les pratiques d'attribution du statut d'auteur, celles-ci mettant en évidence l'existence de différences disciplinaires. S'il est effectivement possible d'observer une certaine homogénéité des pratiques d'attribution du statut d'auteur au sein des disciplines, les travaux de Pontille (2016), notamment, montrent que le lien entre la contribution d'un chercheur et sa position dans la liste des auteurs peut grandement varier d'un article à l'autre (même au sein d'une même discipline). Nous avons également vu que diverses formes d'ordonnement

des auteurs sont utilisées par les chercheurs, et présenté quelques typologies d'auteurs proposées par certains auteurs.

Dans la dernière partie, nous avons fait un survol de la littérature sur les politiques éditoriales des revues savantes en matière d'attribution du statut d'auteur. Ces politiques sont plus fréquentes dans les disciplines où le nombre d'auteurs des articles est généralement plus élevé, notamment dans le domaine biomédical, où les lignes directrices de l'ICMJE font autorité. La capacité de ces dernières et des politiques éditoriales des revues à « contrôler » les pratiques d'attribution du statut d'auteur semble toutefois limitée. Finalement, nous avons discuté du nombre croissant de revues, surtout dans le domaine biomédical, qui exigent que les articles qu'elles publient incluent une déclaration des contributions individuelles de chacun des auteurs.

1.3. Le brevet d'invention et le statut d'inventeur

Notre recherche visant à comparer les listes d'auteurs et d'inventeurs d'articles et de brevets afin de mieux comprendre les pratiques de collaboration et d'attribution des statuts d'auteur et d'inventeur, nous présentons dans la présente section les notions de brevet d'invention et d'inventeur.

1.3.1. Le brevet d'invention

Le brevet d'invention (voir l'Annexe 3, p. xxiii), régi au Canada par la Loi sur les brevets⁶, fait partie du régime de la propriété intellectuelle, qui a pour fonction de stimuler la recherche et l'innovation en récompensant le créateur pour les productions qu'il réalise et partage avec la collectivité (Dubé, 2010a). Cet objectif est aussi clairement présent dans la Constitution américaine selon laquelle « Le Congrès aura le pouvoir [...] de promouvoir le progrès de la science et des arts utiles, en garantissant aux auteurs et aux inventeurs, pour une durée limitée, le droit exclusif à leurs écrits et découvertes respectifs.⁷ » (notre traduction)

La protection qu'apporte le brevet est double. D'une part, le brevet vise la protection de la propriété intellectuelle comme objet d'échange commercial et, d'autre part, il vise à protéger

⁶ *Loi sur les brevets*, LRC 1985, c. P-4, art 1.

⁷ US Const art I § 8.

l'inventeur qui fournit une contribution intellectuelle (Kan, 2009a). Cela est bien illustré par un jugement de la Cour suprême du Canada :

The strongest economic argument is that, without the protection of copyright and patent, research and creative thought would not be financed by the groups (business firms, manufacturers and publishers) who traditionally financed such enterprises.

A patent represents a quid pro quo. The Quid to the patentee is the monopoly; the Quo is that he presents to the public the knowledge which they have not got.⁸

L'invention est définie dans la loi canadienne sur les brevets comme « Toute réalisation, tout procédé, toute machine, fabrication ou composition de matières, ainsi que tout perfectionnement de l'un d'eux, présentant le caractère de la nouveauté et de l'utilité.⁹ » Pour être brevetable, une invention doit être :

1. nouvelle (la première au monde);
2. utile (fonctionnelle et exploitable);
3. inventive (constituer un apport inventif et ne pas être évidente pour une personne possédant les connaissances de base et travaillant dans le domaine de votre invention). (Office de la propriété intellectuelle du Canada, 2017)

Le brevet d'invention est un document décrivant une invention respectant ces trois critères. Il contient un titre, une description générale de l'invention ainsi qu'une ou plusieurs revendications. Ces dernières sont les éléments nouveaux, utiles et inventifs que les demandeurs cherchent à protéger en déposant une demande de brevet. C'est cette partie du brevet qui a une valeur légale et qui délimitera le monopole d'exploitation dont pourra jouir le titulaire du brevet. Chaque revendication fera l'objet d'une évaluation afin de déterminer si les trois critères sont respectés et si la propriété intellectuelle peut être attribuée aux demandeurs.

Il est généralement reconnu que la recherche scientifique financée par les fonds publics doit contribuer à la croissance économique, si bien que la commercialisation des résultats de recherche est parfois vue comme la troisième mission des universités (Etzkowitz et Leydesdorff,

⁸ *Pope Appliance Corp v Spanish River Pulp and Paper Mills*, [1929] AC 269.

⁹ *Loi sur les brevets*, LRC 1985, c. P-4, art 1.

2000; Meyer, 2003). Cependant, pour les chercheurs du milieu universitaire qui entretiennent des liens étroits avec l'industrie, le choix de diffuser leur découverte au moyen d'articles scientifiques ou de brevets résulte d'une négociation entre des intérêts et des normes scientifiques qui divergent des intérêts et normes du monde des affaires (Gans, Murray et Stern, 2017). En effet, malgré les liens qui existent entre elles, la recherche universitaire et la recherche privée constituent des champs distincts régis par des normes, des valeurs et des systèmes de reconnaissance distincts, où sont à l'œuvre différents agents (Dasgupta et David, 1994; Packer et Webster, 1996). Même s'il arrive fréquemment qu'une recherche mène à la fois à la publication d'un article et à l'obtention d'un brevet, toute découverte scientifique n'est pas nécessairement brevetable et toute invention n'est pas nécessairement d'intérêt pour la communauté scientifique. Cela peut être dû d'une part au critère d'utilité de l'invention (ce qui exclut la plupart des résultats de la recherche fondamentale) et au fait que certains éléments (p. ex. les organismes vivants) ne sont pas brevetables selon les lois sur la propriété intellectuelle et, d'autre part, au fait que le concept de nouveauté (ou d'originalité) en science n'est pas équivalent au concept de nouveauté en propriété intellectuelle (Packer et Webster, 1996). Dans leur étude, Packer et Webster (1996) font une distinction entre les notions de découverte et d'invention : alors qu'une découverte a lieu lorsqu'on dévoile l'existence ou certaines propriétés d'un phénomène naturel ou social, une invention est plutôt la création humaine et non évidente (originale). La nouveauté et l'originalité d'une invention sont évaluées de façon absolue et universelle à l'aide de comparaison au domaine public (toute forme de publication ou de divulgation) et à l'état des connaissances de base dans le domaine. Par contre, la nouveauté et l'originalité d'un article scientifique sont évaluées de façon relative et locale en considérant l'état des connaissances dans la communauté spécifique auquel l'article est destiné (Packer et Webster, 1996).

1.3.2. L'inventeur

1.3.2.1. Qu'est-ce qu'un inventeur?

L'inventeur est une personne physique¹⁰ (le statut d'inventeur ne pouvant être attribué à une personne morale) à qui est attribuée la paternité de l'invention, celle qui a conçu la solution au problème et qui l'a matérialisée. Il s'agit d'un droit moral qui consacre le lien entre le créateur et son œuvre (Dubé, 2010a). La reconnaissance des contributions intellectuelles par le statut d'inventeur est essentielle pour l'organisation sociale et elle a une valeur tant pour les créateurs que pour les utilisateurs, indépendante de celle de l'attribution de la propriété (Kan, 2009b). Il est important de distinguer la notion d'inventeur de celle de titulaire du brevet. Alors que l'inventeur jouit d'un droit moral qui est inaliénable, le titulaire du brevet jouit d'un droit économique transférable lui donnant un monopole d'exploitation de l'invention (Dubé, 2010b, p. 55). Autrement dit, c'est lui qui en est le propriétaire. À moins d'une entente contraire, les inventeurs sont titulaires du brevet. Cependant, les entreprises sont généralement titulaires des brevets lorsque l'invention a lieu dans le cadre de leurs activités. Bien qu'ils ne bénéficient pas des droits d'exploitation des brevets en ce cas, les inventeurs obtiennent généralement une compensation financière prenant souvent la forme d'une part des profits générés par l'invention. Le droit à une compensation, sa nature et son ampleur sont régis par des lois qui varient grandement d'un pays à l'autre. Par exemple, en France et en Allemagne, les employés qui sont inventeurs ont droit à une compensation dont la valeur est calculée au moment de l'invention, alors qu'au Royaume-Uni la compensation n'est pas automatique et est calculée en fonction des bénéfices réels obtenus par le ou les titulaires (Peberdy et Strowel, 2010). Cela dit, tant dans les entreprises que les universités, le partage des revenus de la commercialisation des inventions est abordé dans les politiques, les conventions collectives ou le contrat de travail. Dans les universités canadiennes, la part des inventeurs est généralement entre 25 % et 50 % des revenus nets (Couture et Malissard, 2010). Un survol des politiques sur la propriété intellectuelle des

¹⁰ Selon le grand dictionnaire terminologique, un inventeur est une « *personne physique* [emphase ajoutée] mentionnée comme telle dans un brevet ou une demande de brevet ». (http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=2084450)

universités américaines démontre qu'elles sont similaires à celles du Canada, la part des inventeurs allant du quart à la moitié des revenus nets générés par les brevets¹¹.

Puisque la propriété intellectuelle fait l'objet de lois, on pourrait s'attendre à ce que le concept d'inventeur et les critères d'attribution du statut y soient mieux définis que le concept d'auteur l'est en science. Or ce n'est pas le cas : le concept d'inventeur et les critères qui déterminent l'attribution de ce statut ne sont pas définis dans les lois sur les brevets. En effet, plusieurs législations, dont les législations canadienne¹², italienne (Lissoni et al., 2013), française et européenne (Grosset-Fournier et Dacheux, 2012), ne proposent pas de définition du terme « inventeur », alors que d'autres donnent des définitions plutôt minimales, comme le Royaume-Uni où l'inventeur est défini comme « the actual deviser of an invention¹³ » ou encore les États-Unis où l'inventeur est « the individual or, if a joint invention, the individuals collectively who invented or discovered the subject matter of the invention.¹⁴ »

Malgré l'importance que semble posséder l'attribution adéquate du statut d'inventeur, les lois sont relativement silencieuses en ce qui concerne ses critères d'octroi, à l'exception de la loi américaine dont la seule indication est que la demande conjointe ne nécessite pas que les co-inventeurs aient travaillé physiquement ensemble, qu'ils aient fait des contributions d'une ampleur similaire et de même nature, ou qu'ils aient contribué à chaque revendication du brevet¹⁵. Devant cette absence de définition claire de l'inventeur dans la loi, les conflits documentés qui sont liés à l'attribution du statut d'inventeur constituent une source

¹¹ Les politiques sur la propriété intellectuelle des établissements suivants ont été consultés le 1^{er} janvier 2017 :

- Cornell University (https://www.dfa.cornell.edu/sites/default/files/policy/vol1_5.pdf)
- Drexel University (http://drexel.edu/provost/policies/patent_policy/)
- Harvard University (<http://otd.harvard.edu/facultyinventors/resources/policiesandprocedures/statement-ofpolicyinregardtointellectualproperty/>)
- The University of Pennsylvania (<http://www.upenn.edu/research/RevisedPatentPolicy5-19-10.pdf>)
- Stanford University (http://otl.stanford.edu/inventors/inventors_policies.html?headerbar=1#royalty)
- Tufts University (<http://techtransfer.tufts.edu/resources/tufts-policies/intellectual-property-policies/>)
- The University of Utah (<http://regulations.utah.edu/research/7-002.php>)

¹⁰ Recherche effectuée le 14 juillet 2017 sur le site de l'Office de la propriété intellectuelle du Canada (OPIC) et dans la Loi sur les brevets (LRC [1985], ch P-4).

¹³ *Patents Act 1977* (R-U), c 37, art 3.

¹⁴ *Patents*, 35 USC § 100.

¹⁵ *Patents*, 35 USC § 116.

d'information incontournable pour déterminer comment la notion d'inventeur a été interprétée et mise en pratique par les juges dans des contextes précis.

1.3.2.2. Pratiques d'attribution du statut d'inventeur

L'attribution du statut d'inventeur est du ressort des demandeurs du brevet eux-mêmes et ne fera l'objet d'aucune vérification¹⁶, sauf lorsqu'elle est contestée en cour (Lissoni et al., 2013). Ainsi, il faut se tourner vers la jurisprudence (principalement américaine¹⁷) pour comprendre la façon dont la loi est interprétée et appliquée pour déterminer à qui le statut d'inventeur revient de droit. Or cette jurisprudence montre que l'élément déterminant est la conception : processus intellectuel de formulation claire et bien définie de l'invention permettant à une personne versée dans l'art du domaine de la comprendre et de pouvoir la mettre en pratique (Ho, 2004). La décision suivante de la cour américaine illustre bien l'importance de la conception :

The threshold question in determining inventorship is who conceived the invention. Unless a person contributes to the conception of the invention, he is not an inventor. ... Insofar as defining an inventor is concerned, reduction to practice, per se, is irrelevant.¹⁸

Il faut donc non seulement formuler le résultat escompté, mais aussi contribuer à la résolution du problème de façon concrète et inventive¹⁹. Chaque inventeur doit avoir fait une contribution de ce type à l'une au moins des revendications du brevet (Sheiness et Canady, 2006), l'accomplissement de tâches routinières ne suffisant pas²⁰. En effet, dans la mesure où l'objectif du régime de la propriété intellectuelle est de stimuler l'innovation en récompensant les personnes pour leurs contributions créatives, il serait inefficace d'attribuer le statut d'inventeur à tous les participants d'un projet, à ceux qui ont effectué des tâches routinières pour quelqu'un du domaine ou à ceux qui ont simplement contribué financièrement sans avoir participé à la

¹⁶ Par exemple, en France, l'article R611-15 du Code de la propriété intellectuelle précise que « L'Institut national de la propriété industrielle ne contrôle pas l'exactitude de la désignation de l'inventeur prévue à l'article R. 612-10. »

¹⁷ Nous nous basons en grande partie sur la jurisprudence américaine parce que, d'une part, nous utilisons les brevets de la United States Patent and Trademark Office (USPTO) dans notre étude et, d'autre part, c'est sur celle-ci que s'appuie la plupart des discussions entourant l'attribution du statut d'inventeur dans le contexte de la recherche scientifique.

¹⁸ *Hardee, In re*, 223 USPQ 1122 (Comm'r Pat 1984).

¹⁹ *Smernoff, Ex parte*, 215 USPQ 545 (Bd App 1982).

²⁰ *Burroughs Wellcome Co v Barr Labs Inc*, 40 F.3d 1223, 32 USPQ2d 1915 (Fed Cir 1994).

conception (Kan, 2009b). C'est pourquoi les techniciens de laboratoire ne sont généralement pas reconnus comme inventeurs. Dans l'attribution du statut d'inventeur, il s'agit donc non seulement d'évaluer qui a fait quoi, mais également de vérifier que ce qui a été accompli n'était ni déjà connu (ce qui contreviendrait à la notion d'innovation), ni une procédure routinière (Armstrong et Murphy, 2012).

Nous avons vu que selon la loi américaine, il n'est pas nécessaire pour les co-inventeurs d'avoir travaillé ensemble, d'avoir fait des contributions de valeur égale et de même nature et d'avoir contribué à chaque revendication du brevet²¹. Il n'est également pas nécessaire que les co-inventeurs aient pris conscience de l'invention au même moment²². Par contre, il doit exister un lien effectif de collaboration entre eux, comme le précise la décision suivante de la cour américaine :

[f]or persons to be joint inventors under Section 116, there must be some element of joint behavior, such as collaboration or working under common direction, one inventor seeing a relevant report and building upon it or hearing another's suggestion at a meeting.²³

L'inventeur n'a pas à avoir lui-même effectué l'ensemble du processus inventif. Il peut déléguer à d'autres les tâches qui ne demandent pas de faire preuve d'inventivité²⁴. Il peut également considérer et adopter les idées d'autrui, tant qu'il conserve le contrôle du processus intellectuel de sélection et de rejet des idées ainsi que de leur mise en application pour la résolution du problème²⁵.

Puisque les brevets peuvent constituer une source importante de revenus, il est souvent préférable de laisser un professionnel s'occuper de l'identification des inventeurs afin qu'elle soit juste (Sheiness et Canady, 2006). Ainsi, les universités aident les demandeurs dans le processus de demande de brevet, souvent par l'entremise d'un bureau de transfert de technologie (BTT). Cela dit, les agents de brevet se fieront généralement aux responsables de laboratoire pour l'identification des inventeurs (Armstrong et Murphy, 2012). Or si les

²¹ *Patents*, 35 USC § 116

²² *Moler v Purdy*, 131 USPQ 276 (Bd Pat Inter 1960).

²³ *Kimberly-Clark Corp v Procter & Gamble Distrib Co*, 973 F.2d 911, 23 USPQ2d 1921 (Fed Cir 1992).

²⁴ *DeBaun, In re*, 687 F.2d 459, 214 USPQ 933 (CCPA 1982); *Fritsch v Lin*, 21 USPQ2d 1731 (Bd Pat App & Inter 1991).

²⁵ *Morse v Porter*, 155 USPQ 280 (Bd Pat Inter 1965).

responsables de laboratoire ont l'habitude de déterminer qui est (ou n'est pas) auteur d'un article scientifique, ils ne connaissent souvent pas très bien les règles en matière d'attribution du statut d'inventeur (Armstrong et Murphy, 2012). Ainsi, il peut parfois être difficile de déterminer qui est inventeur dans le cas de recherches universitaires, ces dernières se faisant souvent en collaboration avec d'autres institutions, groupes de recherche ou contractants externes (Seymore, 2006).

En principe, l'attribution du statut d'inventeur est une question de faits, et ne peut donc pas faire l'objet d'une entente entre les parties, ce qui serait contraire aux politiques publiques²⁶. De plus, la loi américaine ne permet pas d'attribuer le statut d'inventeur à quelqu'un pour toute autre raison qu'une contribution concrète et inventive à l'une au moins des revendications du brevet. L'omission d'un inventeur peut quant à elle rendre le brevet invalide (Seymore, 2006), et donc avoir d'importantes conséquences financières pour son ou ses titulaires. Il est cependant possible de modifier la liste des inventeurs sans invalider le brevet lorsque l'omission résulte d'une erreur de bonne foi²⁷. Bien qu'il soit dans l'intérêt de toutes les parties d'identifier correctement les inventeurs, différents facteurs sociaux, comme la compétition entre les chercheurs, leur statut et leur réputation, peuvent toutefois réduire l'objectivité du processus (Seymore 2006). Les chercheurs peu expérimentés, surtout les étudiants ou les post-doctorants, ne connaissent pas nécessairement bien les lois sur les brevets et sur la propriété intellectuelle, ce qui fait qu'ils ne remettront pas en question les décisions de leurs supérieurs. Même en cas de désaccord avec ces derniers, ils n'auront souvent pas le courage de les affronter (Seymore, 2006). En somme, malgré les apparences purement factuelles de l'attribution du statut d'inventeur, il existe des facteurs sociaux susceptibles d'influencer le processus et auxquels nous nous intéressons dans la présente recherche.

1.3.2.3. Similitudes et différences entre le statut d'auteur et celui d'inventeur

Les statuts d'inventeur et d'auteur sont similaires en ce qu'il s'agit dans les deux cas de droits moraux à la paternité d'une œuvre. Au-delà de ces droits moraux, les articles et les brevets confèrent également des droits patrimoniaux protégeant leur exploitation par leurs propriétaires

²⁶ *Eli Lilly & Co v Aradigm Corp*, 376 F.3d 1352 (Fed Cir 2004).

²⁷ *Patents*, 35 U.S.C. 256; *Stark v Advanced Magnetics Inc*, 119 F.3d 1551, 1556 (Fed Cir 1997).

(les titulaires dans le cas du brevet, et généralement la revue dans le cas de l'article). Cependant, les inventeurs obtiennent typiquement une part des revenus générés par le brevet, ce qui n'est généralement pas le cas des auteurs d'articles scientifiques. Ainsi, l'ajout d'un inventeur peut diminuer les gains potentiels de capital économique pour les autres inventeurs. De plus, l'ajout d'un inventeur employé par une organisation différente peut aussi mener à l'ajout d'un titulaire, et ainsi diminuer les gains potentiels pour le ou les autres titulaires, qui ne seront plus seuls à pouvoir exploiter l'invention. Il y aurait donc un fort incitatif pour les demandeurs de brevets (tant les titulaires que les inventeurs) à ne nommer comme inventeurs que les personnes ayant effectivement contribué à l'invention et ayant droit à ce statut. Ces incitatifs sont moindres en ce qui concerne l'attribution du statut d'auteur, puisque l'ajout d'un ou de plusieurs auteurs ne conduit pas à un tel partage des droits d'exploitation : l'article sera publié dans une seule revue, qui en devient l'unique propriétaire, et les revenus générés ne sont pas partagés avec les auteurs. De plus, contrairement au gain en capital économique que peut apporter le brevet, le partage du capital symbolique entre les coauteurs n'est pas un jeu à somme nulle²⁸ (Nudelman et Landers, 1972). L'ajout d'un coauteur supplémentaire est donc peu « coûteux » pour les autres auteurs, comparativement à l'ajout d'un inventeur dont la part des revenus potentiellement générés par le brevet sera déduite de celle des autres inventeurs.

Une autre distinction importante entre auteur et inventeur est que les lignes directrices pour l'attribution du statut d'auteur n'ont pas de portée juridique et existent sous diverses formes, même au sein d'une même discipline. Inversement, même si elles ne sont pas très clairement définies dans la Loi sur les brevets, les règles d'attribution du statut d'inventeur qui sont explicitées par la jurisprudence sont les mêmes pour tous et sont visiblement plus strictes que les règles d'attribution du statut d'auteur, l'élément déterminant étant la contribution à l'aspect novateur de l'invention. En effet, l'attribution du statut d'inventeur est, en principe, basée uniquement sur la nature de la contribution et ne subit donc pas l'influence de la quantité de travail effectuée et des facteurs sociaux comme le statut et le prestige (Seymore, 2006). On

²⁸ Selon le Grand dictionnaire terminologique, un jeu à somme nulle est un « jeu dans lequel ce qui est rendu par un groupe de joueurs est gagné par un autre groupe de joueurs » (http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=17058567)

peut donc s'attendre, en toute logique, à une plus grande homogénéité dans les pratiques d'attribution du statut d'inventeur que dans celles du statut d'auteur.

1.3.3. En résumé

Cette section a porté sur le brevet d'invention, le statut d'inventeur et les similitudes et différences entre les statuts d'inventeur et d'auteur. Nous avons présenté les critères de brevetabilité d'une invention, soit la nouveauté, l'utilité et l'inventivité. Les diverses lois régissant l'octroi des brevets n'abordant pas (ou très peu) la question de l'attribution du statut d'inventeur, un survol de la jurisprudence a été nécessaire pour illustrer la façon dont la loi est appliquée par les juges. Il est ressorti quatre principaux éléments de ce survol :

1. Un inventeur doit avoir fait une contribution intellectuelle originale et inventive.
2. Un inventeur n'a pas à avoir contribué à l'ensemble de l'invention, mais à au moins une des revendications.
3. Les contributions techniques visant par exemple à matérialiser l'idée ne justifient pas l'obtention du statut d'inventeur.
4. L'inventeur peut considérer et adopter les idées d'autrui, tant qu'il demeure en contrôle de la sélection et du rejet des idées ainsi que de leur mise en application.

La comparaison de ces critères avec les critères d'attribution du statut d'auteur que nous avons décrits à la section précédente permet de souligner que l'attribution du statut d'inventeur est en principe beaucoup plus stricte que celle du statut d'auteur. En effet, seules les contributions intellectuelles inventives mènent au statut d'inventeur, alors que des types de contributions plus variés, notamment les contributions techniques, peuvent mener au statut d'auteur. Une autre différence importante entre le statut d'inventeur et celui d'auteur est que le premier est un vecteur de capital économique pour le chercheur, alors que ce n'est pas le cas du second. Malgré la plus grande homogénéité et une plus grande rigidité dans l'attribution du statut d'inventeur que dans l'attribution du statut d'auteur, les deux sont néanmoins susceptibles d'être influencées par des facteurs autres que la contribution à la recherche (p. ex. le statut hiérarchique).

1.4. L'appariement des articles et des brevets

Dans les sections précédentes, nous avons décrit le fonctionnement du champ scientifique, les mécanismes de reconnaissance et de stratification de ses agents ainsi que les concepts clés de notre étude, soit l'auteur et l'inventeur. Cette étude visant la comparaison d'articles et de brevets issus de la même recherche, la présente section porte sur les techniques utilisées pour identifier divers types de liens entre les brevets et les articles, et plus particulièrement sur celles permettant de former des PBA.

1.4.1. Techniques de détection des liens entre les articles et les brevets

Les techniques de détection des liens entre les articles et les brevets peuvent être divisées en cinq catégories : 1) les techniques basées sur les citations, 2) les techniques basées sur l'activité dans les deux champs (les auteurs-inventeurs), 3) les techniques basées sur les classifications thématiques, 4) les techniques basées sur le contenu (techniques lexicales) et 5) les techniques hybrides combinant deux types de techniques ou plus (Bassecoulard et Zitt, 2005). Les sections suivantes présentent des exemples de ces différentes techniques tirées de la littérature. La plupart des exemples sont des techniques hybrides.

1.4.1.1. Les techniques basées sur les citations

Les techniques basées sur les citations entre brevets et articles sont le plus souvent utilisées pour étudier le transfert d'information et la façon dont les deux champs s'alimentent mutuellement (p. ex. Glänzel et Meyer, 2003; Meyer, 2000; Narin, Hamilton et Olivastro, 1997; Narin et Olivastro, 1998). Ces interactions sont fortement concentrées dans quelques domaines, la majorité des citations dans les brevets vers les articles scientifiques étant concentrées dans une minorité de domaines et vice-versa (Verbeek et al., 2002).

Breschi et Catalini (2010) ont utilisé les citations pour identifier des auteurs-inventeurs à partir de l'ensemble des brevets de l'Office européen des brevets (OEB) enregistrés entre 1990 et 2003 ainsi que les données du *Science Citation Index* (SCI). Ils ont utilisé les références à la « littérature non-brevet » apparaissant dans les brevets afin de repérer la publication dans le SCI, puis ont comparé les auteurs des articles cités aux inventeurs des brevets les citant afin

d'identifier les auteurs-inventeurs. Après l'élimination des faux positifs²⁹, ils ont obtenu un échantillon d'auteurs-inventeurs équivalant à 21 %, 18 % et 4,5 % de l'ensemble des inventeurs dans les domaines des lasers, des biotechnologies et des semi-conducteurs, respectivement. Ces résultats illustrent l'utilité potentielle des références bibliographiques des brevets et des articles pour leur appariement. Ils sont également le reflet de la participation inégale des chercheurs de différents domaines à l'innovation technologique, et plus particulièrement au régime des brevets.

1.4.1.2. Les techniques basées sur les individus actifs dans les deux champs (les auteurs-inventeurs)

Quelques études (p. ex. Czarnitzki, Glänzel et Hussinger, 2007; Czarnitzki, Hussinger et Schneider, 2011; von Proff, 2011) ont identifié les brevets dont les inventeurs sont des professeurs d'université en cherchant le terme « Prof. » dans le nom des inventeurs. Selon Dornbusch, Schmoch, Schulze et Bethke (2013), cela fonctionne particulièrement bien en Allemagne parce qu'il y est commun d'indiquer son titre académique (p. ex. « Prof. ») sur les documents officiels comme les brevets. Cependant, la méthode est peu utile dans les pays où il n'est pas coutume pour un chercheur d'indiquer son titre. De plus, même lorsqu'indiqué par le chercheur, le titre n'est pas toujours inclus dans les métadonnées des notices. C'est le cas, par exemple, de celles de l'OEB. Cette méthode a aussi comme désavantage d'exclure les inventeurs qui œuvrent dans le milieu universitaire sans avoir le statut de professeur.

Thursby et al. (2009) ont utilisé la liste du personnel d'universités afin d'apparier une liste de 34 202 professeurs et ingénieurs américains avec les brevets américains répertoriés dans la base de données du National Bureau of Economic Research (NBER). Ils ont éliminé de leur liste les chercheurs au nom trop commun et conservé uniquement les cas où l'adresse de l'inventeur était située à 50 miles ou moins de l'institution du chercheur (la distance étant calculée à l'aide des codes postaux). Finalement, dans les cas où le titulaire du brevet ne correspondait pas à l'institution du chercheur, une vérification manuelle a été effectuée afin de

²⁹ Un faux positif est une erreur qui a lieu lorsque qu'on obtient un résultat positif alors qu'en réalité le résultat est négatif. Inversement, un faux négatif a lieu lorsqu'on obtient un résultat négatif alors qu'en réalité il est positif. Par exemple, une PBA formée alors que le brevet et l'article ne sont pas liés est un faux positif alors que la non-identification d'une PBA lorsqu'un brevet et un article sont liés est un faux négatif.

valider ou de rejeter la paire. Cette méthode leur a permis d'obtenir un échantillon de 6 472 paires brevet-chercheur. Les limites d'une telle approche sont, d'une part, le coût élevé de création des listes de personnel et, d'autre part, la possibilité que certains membres « non officiels » de l'organisation ne figurent pas sur les listes (Dornbusch et al., 2013).

Boyack et Klavans (2008) ont également contourné les problèmes d'homonymie causés par les chercheurs et les inventeurs au nom trop commun. En limitant leur échantillon aux brevets et aux articles avec des inventeurs et des auteurs aux noms rares, ils ont identifié 18 251 chercheurs distincts apparaissant sur 56 000 brevets américains dont la demande a été déposée entre 2002 et 2006. Cette méthode est donc très efficace pour construire un bon échantillon de brevets et d'articles signés par la même personne, mais son utilité est limitée lorsque l'objectif est de développer une méthode d'appariement exhaustive des articles et des brevets. Pour faire face aux problèmes d'homonymie, il est aussi possible d'utiliser l'affiliation institutionnelle des chercheurs, à condition que cette information soit disponible et que l'institution figure également dans la liste des titulaires du brevet. Or, selon une étude de 2009, aux États-Unis, les universités sont titulaires de la majorité des brevets issus de la recherche en leur sein, alors qu'en Europe, la majorité des brevets universitaires ne sont pas assignés à l'université, mais plutôt à une ou plusieurs tierces parties (Thursby et al., 2009).

Haeussler et Sauermann (2013) ont quant à eux réalisé un sondage auprès de chercheurs allemands et britanniques en biologie et en médecine, leur demandant s'ils avaient dans le passé participé à une recherche ayant mené à la fois à une publication et à un brevet. Ce fut le cas de 45 % des répondants, soit un total de 2 191 chercheurs, ce qui équivaut à 12,7 % de la population de 17 263 chercheurs du domaine identifié par les auteurs à l'aide de PubMed. L'un des avantages de cette méthode est que les paires ainsi formées n'ont pas à être validées d'une autre façon.

Afin de comparer l'impact scientifique des chercheurs qui brevettent et de ceux qui ne le font pas, Meyer (2006) a utilisé le nom et le pays des auteurs des articles du WoS et des inventeurs des brevets de la United States Patent and Trademark Office (USPTO) dans le domaine des nanotechnologies pour identifier des chercheurs-inventeurs. Se limitant aux chercheurs allemands, belges et britanniques, il a trouvé que dans ces pays une proportion respective de 0,5 %, 1,5 % et 0,6 % des chercheurs dans le domaine des nanotechnologies

étaient également inventeurs, et inversement que 27,3 %, 33,8 % et 40,5 % des inventeurs étaient également chercheurs.

Dans le but d'identifier les brevets issus de la recherche universitaire en Allemagne, Dornbusch et al. (2013) ont utilisé une méthode basée sur l'identité des auteurs et des inventeurs pour appairer les brevets de l'OEB aux articles de Scopus. Les informations utilisées pour l'appariement sont présentées dans le Tableau I (p. 53).

Tableau I. Informations utilisées par Dornbusch et al. (2013) pour l'appariement des brevets et des articles.

Critères	Brevets	Articles
Pays	Pays de l'inventeur	Pays de l'institution
Institution	Type de demandeur	Institution de l'auteur
Nom	Nom et prénom de l'inventeur	Nom et prénom de l'auteur
Localisation	Code postal de l'adresse de l'inventeur	Code postal de l'adresse de l'institution
Fenêtre temporelle	Année de la demande	Année de publication
Sujet	Classification technologique	Classification disciplinaire de la revue (Scopus)

Les auteurs ont dans un premier temps limité leur jeu de données aux brevets dont au moins un des titulaires est une université et aux articles dont au moins un des auteurs est affilié à une université allemande. Ils ont ensuite formé des paires à partir des brevets et des articles qui partagent au moins un inventeur et un auteur et dont l'article a été publié dans les deux ans suivant le dépôt de la demande de brevet. Ils n'ont pas tenu compte de l'institution des brevets puisque le demandeur est rarement l'institution elle-même et parce que l'objectif de leur étude était de développer une méthode basée sur l'identité des auteurs et des inventeurs. À partir d'un ensemble de 678 brevets validés, ils ont montré que l'utilisation des deux premiers chiffres du code postal seulement est optimale lorsque l'on accorde une importance égale à la précision et au rappel.

1.4.1.3. Les techniques basées sur les classifications thématiques

Comme nous le verrons dans les exemples ci-dessous, les classifications thématiques des brevets et des articles sont souvent utilisées conjointement à celles basées sur les auteurs-inventeurs afin, encore une fois, de faciliter l'élimination des faux positifs.

Coward et Franklin (1989) ont établi un réseau de cocitations à partir de la littérature scientifique en épitaxie et en microlithographie. Les différentes grappes identifiées à l'intérieur du réseau leur ont permis d'identifier et de définir des spécialités de recherche ainsi que les relations et la proximité entre elles. Ils ont ensuite fait l'appariement des brevets aux spécialités à l'aide d'une méthode en deux étapes. Premièrement, ils ont identifié automatiquement les PBA potentielles à l'aide des noms des auteurs et des inventeurs, du nom des institutions ainsi que des listes de références des brevets et des articles. Les noms des auteurs et des inventeurs ont permis de former beaucoup plus de PBA potentielles que les institutions et les références. La seconde étape a été de vérifier manuellement les paires identifiées afin d'éliminer les erreurs. Cette étape a comporté trois opérations :

1. Comparer les titulaires du brevet aux institutions auxquelles sont affiliés les auteurs de l'article. Si les titulaires du brevet sont situés dans un autre pays que les institutions citées dans l'article, la paire est rejetée.
2. Comparer le titre du brevet à la spécialité de l'article. En cas de contraste évident, la paire est rejetée.
3. Comparer les co-inventeurs et les coauteurs.

Il en est ressorti 255 PBA, que les auteurs ont divisées en trois catégories, en fonction de la probabilité que le brevet et l'article soient effectivement issus de la même recherche. Le Tableau II (p. 55) décrit les critères utilisés pour la classification.

Tableau II. Catégories de PBA de Coward et Franklin (1989, traduction libre, p. 60-61).

Catégorie	Critères
Paire quasi certaine (n = 128)	Une ou plusieurs des conditions suivantes sont respectées : <ol style="list-style-type: none"> 1. L'inventeur et l'auteur ont le même nom de famille et au moins deux initiales identiques (ex. Smith-JS et Smith-JS). 2. Le brevet a au moins deux inventeurs, qui se situent tous les deux dans la grappe du réseau et/ou dans la même spécialité (proximité intellectuelle). 3. Les titres du brevet et de la spécialité sont très similaires (proximité sémantique).
Paire très probable (n = 112)	Aucune des conditions ci-dessus n'est respectée, mais le titre du brevet et celui de la spécialité sont relativement semblables et il n'y a pas d'autre raison de douter de la validité de la paire.
Paire possible (n = 15)	Une ou plusieurs des conditions suivantes sont respectées : <ol style="list-style-type: none"> 1. Le nom de l'auteur correspond à plusieurs inventeurs (homonymie), mais aucune autre information ne permet de rejeter la paire. 2. Les titres du brevet et de la spécialité ne sont pas complètement différents, mais semblent toutefois représenter des focalisations intellectuelles distinctes.

Bien que Coward et Franklin (1989) utilisent un ensemble de méthodes et d'informations nous semblant pertinentes, on peut se demander si les critères utilisés pour identifier les « paires quasi certaines » sont suffisamment stricts. Par exemple, le fait qu'un auteur et un inventeur aient le même nom de famille et les deux mêmes initiales, lorsque ce critère est utilisé seul, risque de créer de nombreux faux positifs. Certains critères manquent aussi de précision. Par exemple, les auteurs n'expliquent pas ce qu'ils considèrent comme un titre de brevet très similaire au titre d'une spécialité. Malgré l'utilisation de critères potentiellement trop inclusifs et mal définis, l'étude de Coward et Franklin (1989) donne néanmoins l'exemple d'une méthode d'appariement des articles et des brevets pouvant servir d'inspiration pour le développement de méthodes potentiellement plus robustes.

1.4.1.4. Les techniques basées sur le contenu (techniques lexicales) et les techniques hybrides

Les techniques basées sur le contenu permettent un appariement plus précis des articles et des brevets et sont indispensables pour notre recherche. L'utilisation du contenu des brevets et des articles pour l'appariement peut être manuelle (le chercheur consulte les documents afin d'identifier ceux qui sont les plus similaires) ou automatisée (fouille de texte). Comme les techniques présentées ci-dessous en exemple combinent ces techniques avec celles basées sur l'identification d'auteurs-inventeurs, elles constituent également des exemples de techniques hybrides.

Ducor (2000), Murray (2002), ainsi que Murray et Stern (2007) ont identifié des PBA de façon manuelle (c.-à-d. en analysant le contenu des articles et des brevets). Le principal avantage de cette méthode est qu'elle offre une grande précision dans le repérage des paires. Cependant, c'est une méthode longue et coûteuse qui permet d'utiliser seulement de petits échantillons d'articles et de brevets, souvent limités à un domaine de recherche particulier. Par exemple, Ducor (2000) a identifié 40 PBA en biologie moléculaire, alors que Murray (2002) a ciblé deux chercheurs éminents dans le domaine du génie tissulaire et a apparié 53 articles et 11 brevets dont ils étaient tous les deux coauteurs et co-inventeurs. Murray et Stern (2007) ont, quant à eux, utilisé l'ensemble des 340 articles parus dans la revue *Nature Biotechnology* entre 1997 et 1999 inclusivement, et trouvé 169 brevets correspondants dans la base de données du USPTO. Pour la recherche des brevets, ils ont utilisé les noms des premiers auteurs, des derniers auteurs et des institutions apparaissant dans les articles. À partir des données du USPTO, Fehder, Murray et Stern (2014) ont utilisé une série de requêtes utilisant les noms des auteurs d'articles publiés dans les revues *Nature Biotechnology* et *Nature Materials* pour identifier des PBA potentielles. Ils ont ensuite examiné les articles et les brevets pour valider les paires formées, ce qui leur a permis d'identifier 525 PBA.

Les techniques basées sur le contenu peuvent aussi être automatisées, ce qui est utile lorsqu'on tente de former des PBA à partir d'un grand nombre d'articles et de brevets, comme en témoigne leur utilisation dans plusieurs études récentes (p. ex. Cassiman, Glenisson et Van Looy, 2007; Magerman, Van Looy, Baesens et Debackere, 2011; Magerman, Van Looy et Debackere, 2015; Magerman, Van Looy et Song, 2009). Par exemple, Magerman et al. (2015)

ont utilisé une technique basée à la fois sur le contenu et sur l'identification des auteurs-inventeurs. Ils ont appliqué trois critères de façon séquentielle pour identifier 584 PBA à partir de 88 248 brevets et 948 432 articles dans le domaine des biotechnologies. Ces critères sont : 1) le nombre de mots en commun entre les deux documents divisé par le nombre de mots du document le plus court est égal ou supérieur à 0,60; 2) le nombre de mots en commun entre les deux documents divisé par le nombre de mots du document le plus long est égal ou supérieur à 0,30; et 3) au moins un auteur de l'article figure dans la liste des inventeurs du brevet. Ils ont par ailleurs procédé à une validation de 250 PBA auprès d'experts du domaine.

Lissoni et al. (2013) ont utilisé l'ensemble des brevets (389) de la base de données KEINS³⁰ dont la demande a été déposée entre 1978 et 2001, se limitant aux domaines où les demandes de brevets sont les plus fréquentes (génie chimique, biologie, pharmacologie et électronique et télécommunications) ainsi qu'aux brevets dont au moins un inventeur est affilié à une université italienne. Ils ont ensuite utilisé le WoS pour identifier l'ensemble des 2 838 articles publiés entre 1975 et 2003 par les 218 inventeurs distincts identifiés à partir des 389 brevets. Les auteurs ont ensuite utilisé les critères suivants pour former des PBA dans cet ensemble de brevets et d'articles :

1. Au moins un des auteurs de l'article est aussi inventeur sur le brevet.
2. L'article a été publié à l'intérieur d'un délai de deux ans avant ou après l'année de dépôt de la demande de brevet.
3. Les publications sans résumé et les brevets ayant été déclarés par les inventeurs comme n'ayant pas fait l'objet d'une publication ont été exclus.

Ces critères ont permis d'identifier 6 810 PBA potentielles. La similarité cosinus de chacune des PBA a ensuite été calculée à partir des mots utilisés dans les titres et les résumés de l'article et du brevet. Les 10 % des paires les plus similaires, soit 680 PBA, ont finalement été retenues. Il est à noter qu'avec cette méthode, un article peut être apparié à un ou plusieurs brevets et vice-versa, ce qui complique l'analyse, mais permet de tenir compte du fait que les projets de recherche peuvent s'étaler sur une période plus ou moins longue et mener à plusieurs

³⁰ Cette base de données construite par Lissoni, Sanditov et Tarasconi (2006) contient un ensemble de brevets académiques (c.-à-d. dont au moins un inventeur est affilié à une université) et couvre la France, la Suède et l'Italie.

publications ou encore à plusieurs brevets, le premier cas étant plus fréquent que le second (Lissoni et al. 2013).

1.4.1.5. Utilité des techniques de liaison des articles et des brevets pour la formation de PBA

L'utilité respective des différentes techniques utilisées pour former des liens entre les articles et les brevets est présentée dans le Tableau III (p. 58) qui résume les principales propriétés de ces techniques. Bassecoulard et Zitt (2005) n'incluent pas dans ce tableau les techniques hybrides, présumément parce que la nature des relations, le rappel et la précision de ces techniques dépendront des choix de combinaison de techniques effectués.

Tableau III. Propriétés des techniques de liaison des articles et des brevets (adapté de Bassecoulard et Zitt, 2005, p.689, notre traduction)

Type de technique	Nature de la relation	Rappel	Précision
Basées sur les citations	Référence à l'article dans un brevet ou vice-versa	Faible	Élevée
Basées sur les individus actifs dans les deux champs	Articles et brevets partageant un ou plusieurs auteurs/inventeurs	Très faible	Très élevée
Basées sur les classifications thématiques	Appartenance des brevets et des articles à une même classe	Élevé	Faible
Basées sur le contenu	Similarité des contenus de champs texte libre des notices bibliographiques des articles et des brevets	Moyen à élevé	Faible à moyenne

Il est à noter que, alors que la présente étude s'intéresse à l'identification d'articles et de brevets issus du même travail de recherche, les niveaux de rappel et de précision associés à chaque type de technique présenté dans le Tableau III (p. 58) sont déterminés par Bassecoulard et Zitt (2005) dans une perspective d'identification de liens plus larges entre les articles et des brevets. Par exemple, pour Bassecoulard et Zitt (2005) les techniques basées sur les individus actifs dans les deux champs ont un rappel très faible, car elles ne prennent pas en compte les brevets et les articles qui sont liés conceptuellement (articles portant sur les mêmes sujets que

et/ou citant des brevets ou vice-versa). Dans notre recherche, ces techniques auront un rappel très élevé puisque nous nous intéressons spécifiquement aux articles et aux brevets qui sont le fruit d'une même recherche (et donc du travail des mêmes individus). Il est en effet peu probable que ces articles et brevets n'aient aucun auteur et inventeur en commun. À l'inverse, le niveau de précision de cette technique utilisée seule ne sera pas très élevé puisque l'inventeur d'un brevet peut également avoir publié de nombreux articles sans que tous ces articles soient issus de la même recherche que le brevet.

Le choix d'une technique hybride peut permettre de maximiser et d'équilibrer le rappel et la précision. D'ailleurs, comme nous le verrons au Chapitre 2, nous utilisons, dans la présente étude, une telle technique hybride combinant une technique basée sur l'activité dans les deux champs et une technique basée sur le contenu, similaire à la méthode utilisée par Magerman et al. (2015) et, dans une moindre mesure, à celle utilisée par Lissoni et al. (2013).

1.4.2. En résumé

Nous avons exploré dans cette section les différentes techniques utilisées pour lier les articles et les brevets. Ces techniques peuvent être basées sur les citations, sur les individus actifs dans les deux champs (les auteurs-inventeurs), sur les classifications thématiques, sur le contenu, ou encore une combinaison de ces informations. Nous avons présenté des exemples d'études ayant utilisé ces différentes techniques, puis avons discuté des avantages et désavantages de chacune des techniques globalement ainsi qu'en ce qui concerne le rappel et la précision pour l'identification de liens entre les articles et les brevets. Les niveaux de rappel et de précision dépendent de la nature du lien à identifier, et ainsi, les objectifs de notre recherche ont servi de guide dans le choix d'une méthode hybride pour effectuer l'appariement des articles et des brevets.

1.5. Résultats des études sur l'attribution des statuts d'auteur et d'inventeur réalisées à partir de paires brevet-article

Nous avons repéré dans la littérature trois études des pratiques d'attribution des statuts d'auteur et d'inventeur et ayant analysé des PBA à cette fin. Comme ces études sont similaires

à la nôtre à la fois au niveau de leur objet d'étude et au niveau de la méthode employée, nous en présentons ici les résultats en détail.

Dans l'échantillon de 40 PBA en biologie moléculaire utilisé par Ducor (2000), 38 avaient plus d'auteurs que d'inventeurs, la moyenne étant de dix auteurs et de trois inventeurs. Une autre observation intéressante est que dans la quasi-totalité des cas (37 sur 40), le dernier auteur était également inventeur, alors que le premier auteur était également inventeur dans seulement 26 cas. Ducor propose deux explications. La première est qu'il y a de nombreux auteurs dont la contribution ne justifierait pas de recevoir le statut d'inventeur, soit parce qu'ils ont effectué des tâches purement techniques, ou encore parce qu'il s'agit d'auteurs honorifiques³¹. La seconde explication est que certains auteurs n'auraient pas été nommés inventeurs malgré une contribution importante. Une autre explication possible de ces différences est que les critères d'attribution du statut d'inventeur seraient plus restrictifs que les critères d'attribution du statut d'auteur, mais Ducor rejette cette explication en soulignant la grande similarité entre les lignes directrices de l'ICJME concernant l'attribution du statut d'auteur et les critères d'octroi du statut d'inventeur dans la loi américaine. Il ajoute :

If a judge facing a patent with three inventors is provided with a corresponding paper having 10 authors and the ICMJE guidelines, chances are that the evidence will be considered sufficient that at least some inventors have been omitted. (Ducor 2000, 2)

Cela contraste avec une décision de la cour affirmant que le statut d'auteur ne permet pas, à lui seul, d'affirmer que la personne aurait dû être nommée inventeur³². Ducor a également observé que pour l'ensemble des sept articles ayant au moins un auteur de l'industrie, seuls ces auteurs avaient le statut d'inventeur sur le brevet correspondant. Il ne propose cependant pas d'explication en lien avec cette observation.

Le sondage effectué par Haeussler et Sauermann (2013) auprès de 2 191 chercheurs allemands et britanniques en biologie et en médecine a montré que 95 % des chercheurs avaient

³¹ En émettant l'hypothèse que des auteurs peuvent être exclus de la liste des inventeurs, Ducor (2000) laisse sous-entendre que les auteurs sont nommés avant les inventeurs. Or, comme les découvertes publiées ne peuvent pas, en principe, être brevetées par la suite, il est plus probable que les décisions concernant l'attribution du statut d'auteur aient plutôt lieu après les décisions concernant l'attribution du statut d'inventeur. Il est à noter que notre analyse des PBA est atemporelle et ne tient pas compte de ce temps écoulé entre la demande de brevet et la publication de l'article.

³² *Katz, In re*, 687 F.2d 450 (CCPA 1982)

le statut d'auteur, 92 % avaient le statut d'inventeur, et 89 % avaient les deux statuts. Ces résultats suggèrent que les répondants ont eu tendance à choisir un projet pour lequel ils avaient été nommés à la fois auteur et inventeur (Haeussler et Sauermann, 2013), ce qui crée un certain biais dans l'échantillon. Cela constitue une limite cependant peu problématique pour Haeussler et Sauermann (2013) qui s'intéressent aux types de contributions et aux facteurs sociaux qui mènent à l'obtention des statuts d'auteur et d'inventeur, et non pas, par exemple, aux différences entre le nombre d'auteurs et d'inventeurs dans les articles et les brevets. D'ailleurs, l'unité d'analyse de leur étude est les chercheurs qui ont contribué à un travail de recherche ayant mené à un article et un brevet, et non la paire article-brevet elle-même. Leurs résultats démontrent que les types de contribution et les facteurs sociaux menant à l'obtention du statut d'auteur sont variés. L'obtention du statut d'inventeur serait surtout liée aux contributions conceptuelles et serait également moins liée à des facteurs sociaux que l'obtention du statut d'auteur. Les limites de l'étude de Haeussler et Sauermann (2013) sont que les résultats ne peuvent être généralisés ni aux travaux ne menant pas à la fois à une publication et un brevet, ni à d'autres domaines que le domaine biomédical.

Lissoni et al. (2013) s'intéressent quant à eux aux facteurs qui déterminent l'attribution du statut d'auteur et d'inventeur au sein des groupes de recherche. Ils ont analysé 680 PBA divisés en quatre disciplines : la pharmacologie (n = 104), la biologie (n = 222), le génie chimique (n = 27) et l'électronique (n = 327). Dans leur étude, la proportion d'auteurs ayant obtenu le statut d'inventeur est la plus faible en pharmacologie (55,7 %) et la plus élevée en génie chimique, où le nombre moyen d'inventeurs des PBA est légèrement plus élevé que le nombre moyen d'auteurs. Leurs résultats montrent que les jeunes chercheurs et les femmes sont plus fréquemment exclus de la liste des inventeurs que les hommes et les chercheurs ayant plusieurs années d'expérience. Leurs résultats montrent également que les derniers auteurs obtiennent le plus souvent le statut d'inventeur suivi des premiers et des autres, et que le nombre de citations n'est pas significativement lié au statut d'inventeur.

Les trois études présentées ci-dessus comportent quelques limites importantes que nous présentons ci-dessous. Premièrement, deux d'entre elles ont analysé des PBA issues de disciplines spécifiques, soit la biologie moléculaire (Ducor, 2000) et la recherche biomédicale (Haeussler et Sauermann, 2013). Les résultats obtenus peuvent donc difficilement être

généralisés aux autres disciplines. Une seconde limite des études antérieures est qu'elles analysent des PBA issues de la recherche dans des pays particuliers, soit l'Allemagne et le Royaume-Uni (Haeussler et Sauermann, 2013), ainsi que l'Italie (Lissoni et al. 2013). Il est ainsi probable que les résultats de ces études ne soient pas généralisables à l'ensemble des pays. En conséquence aux deux limites que nous venons de souligner, mais également aux choix méthodologiques des auteurs, les études antérieures sont également limitées au niveau de la taille des échantillons utilisés. Ducor (2000) a analysé 40 PBA manuellement identifiées, alors que Lissoni et al. (2013) en ont analysé 680. Finalement, bien que Haeussler et Sauermann (2013) aient analysé un plus grand échantillon (2 191 PBA), l'utilisation d'un sondage pour identifier les PBA constitue une limite de leur étude puisqu'elle peut causer un biais dans l'échantillon.

Les analyses antérieures de PBA que nous venons de présenter ont permis d'illustrer les différences entre les pratiques d'attribution du statut d'auteur et les pratiques d'attribution du statut d'inventeur, ainsi que d'identifier des facteurs pouvant influencer ces pratiques et ainsi expliquer (en partie) ces différences. Les principaux éléments à retenir de ces études sont les suivants :

- Le statut hiérarchique est associé au statut d'inventeur, mais les contributions scientifiques passées (c.-à-d. la réputation) ne le sont pas (Haeussler et Sauermann, 2013; Lissoni et al., 2013).
- Les contributions techniques et la fourniture de données ou de matériel sont associées au statut d'auteur, mais non au statut d'inventeur (Haeussler et Sauermann, 2013).
- Les derniers auteurs sont plus souvent nommés inventeurs que les premiers auteurs, et les auteurs qui ne sont ni derniers ni premiers sont moins souvent dans la liste des inventeurs (Ducor, 2000; Lissoni et al., 2013).
- Les femmes sont plus souvent exclues de la liste d'inventeurs que les hommes (Lissoni et al. 2013).
- Les jeunes auteurs sont plus souvent exclus de la liste d'inventeurs que les auteurs avec plus d'expérience en recherche (Haeussler et Sauermann, 2013; Lissoni et al. 2013).

Conclusion

Nous venons, dans ce chapitre, de définir le cadre de référence de notre recherche, principalement formé des théories de Pierre Bourdieu et de Robert K. Merton, qui permettent de conceptualiser la science comme un espace social où les chercheurs font des investissements stratégiques de leurs capitaux, guidés par leur compréhension des normes et des enjeux de la science afin d'obtenir du capital scientifique et, ainsi, améliorer leur position dans la structure hiérarchisée du système. Nous avons aussi vu que les chercheurs réputés se voient souvent attribuer plus de crédit que leurs collègues moins réputés pour des contributions égales (Merton, 1968) et que les contributions des femmes en science sont souvent peu mises en visibilité (Rossiter, 1993).

Nous avons ensuite abordé le statut d'auteur, qui joue un rôle central dans ce système puisqu'une de ses fonctions principales est justement l'attribution du capital scientifique aux chercheurs. Nous avons présenté une revue de la littérature sur l'attribution du statut d'auteur, illustrant la diversité des formes de collaboration et des types de contribution pouvant mener à l'obtention du statut d'auteur. Nous avons vu que les pratiques d'attribution du statut d'auteur découlent à la fois des objets de recherche, de la structure du champ et des normes implicites qui y règnent, mais également des préférences, des intérêts et de l'autorité des chercheurs.

La troisième section a porté sur le brevet d'invention et le statut d'inventeur. Nous avons passé en revue la littérature et la jurisprudence mettant en lumière les critères plus stricts guidant l'attribution du statut d'inventeur comparativement à celui d'auteur. En effet, le statut d'inventeur est en principe attribué seulement aux individus ayant contribué aux aspects inventifs de la recherche alors que des contributions plus diverses mènent à l'obtention du statut d'auteur.

Puisque notre recherche vise à comparer les pratiques d'attribution des statuts d'auteur et d'inventeur à l'aide de PBA, nous avons aussi présenté une revue des méthodes d'appariement des articles et des brevets utilisées dans des travaux de recherche antérieurs. Nous avons également résumé les résultats de ces recherches lorsqu'elles portaient sur un thème proche de celui de notre étude. Ces résultats contribuent à montrer que les pratiques d'attribution du statut d'inventeur sont bel et bien différentes des pratiques d'attribution du statut d'auteur, et que les

deux sont influencées non seulement par la nature et l'ampleur des contributions des chercheurs, mais également par divers facteurs sociaux, dont le sexe et la réputation.

Dans notre étude, les articles et les brevets sont conceptualisés comme des biens symboliques qui génèrent pour les auteurs et les inventeurs du capital symbolique et, dans le cas des brevets seulement, du capital économique direct. En combinant les théories de Merton et de Bourdieu, nous conceptualisons l'attribution du statut d'auteur, d'une part, comme une pratique par laquelle le crédit est attribué aux chercheurs les plus méritants dans le but d'optimiser le fonctionnement du système de reconnaissance scientifique et, d'autre part, comme la manifestation de rapports de force et de stratégies individuelles à des fins de domination du champ. La recherche scientifique est de plus en plus un travail d'équipe, et ces équipes sont de plus en plus grandes, ce qui mène nécessairement une division quelconque du travail. La répartition des responsabilités et des tâches, ainsi que le lien entre la contribution, l'obtention du statut d'auteur, et l'ordre dans lequel les noms des auteurs apparaissent sur l'article, se trouvent alors à être le reflet à la fois des normes disciplinaires, des contributions respectives des membres de l'équipe, de leurs stratégies, ainsi que des rapports de force qui existent entre eux.

Nous avons dégagé de notre revue de la littérature un ensemble de facteurs pouvant jouer un rôle dans les pratiques de collaboration et d'attribution des statuts d'auteur et d'inventeur. Dans notre étude, nous nous intéressons à neuf de ces facteurs, que nous divisons en quatre niveaux : la discipline, la recherche, l'équipe, et l'individu.

Le premier niveau, la discipline, fait référence au contexte disciplinaire dans lequel une recherche est inévitablement située et qui détermine les enjeux (objets de recherche) qui sont d'intérêt pour les chercheurs (Merton, 1942, Bourdieu, 1976). Ces objets de recherche ainsi que la culture épistémique (Knorr Cetina, 1999) de la discipline peuvent imposer certains modes d'organisation du travail aux chercheurs qui seront potentiellement reflétés dans les listes d'auteurs et d'inventeurs des articles et des brevets.

Le deuxième niveau, la recherche, désigne le travail qui a été fait et ses produits finaux (l'article et le brevet). L'article est le bien symbolique (Bourdieu, 1976) qui officialise la contribution des chercheurs à l'inventaire des connaissances et leur permet d'obtenir la

reconnaissance de leurs pairs (Merton, 1957) et d'accroître leur capital symbolique. Le brevet permet quant à lui d'obtenir un gain potentiel en capital économique, mais aussi en capital symbolique dans la mesure où les brevets sont valorisés dans le monde académique (Sanberg et al., 2014; Stevens, Johnson et Sanberg, 2011). Notre revue de la littérature nous a permis d'identifier le gain potentiel en capital symbolique (c.-à-d. l'impact scientifique potentiel de la recherche) comme facteur au niveau de la recherche pouvant influencer l'attribution des statuts d'auteur et d'inventeur. En effet, Jabbehdari et Walsh (2017) ont trouvé que les articles hautement cités ont généralement moins d'auteurs dont la contribution a été minime.

L'équipe de recherche (troisième niveau) désigne l'ensemble des individus impliqués dans un travail de recherche. Nous incluons dans notre étude deux facteurs liés à l'équipe de recherche. Le premier est la taille de l'équipe, telle que mesurée par le nombre d'auteurs, celle-ci pouvant avoir un impact sur l'organisation du travail (Walsh et Lee, 2015) et donc sur la division du travail. Puisque Jabbehdari et Walsh (2017) ont trouvé que les articles en collaboration interinstitutionnelle ont moins de contributeurs ayant été exclus de la liste des auteurs, nous incluons dans notre étude la collaboration interinstitutionnelle comme deuxième facteur au niveau de l'équipe de recherche pouvant influencer la division du travail et l'attribution des statuts d'auteur et d'inventeur.

Les individus impliqués dans la recherche constituent le quatrième niveau de notre modèle. Toute collaboration présuppose une hiérarchie et une division des rôles où chaque individu occupe une position particulière (Gläser, 1963). Cette position est en principe reflétée dans l'attribution du statut d'auteur et nous tentons, dans notre étude, de déterminer dans quelle mesure cette position est liée à l'obtention du statut d'inventeur. Nous divisons cette idée de position en trois facteurs distincts, mais interreliés : la position dans la liste des auteurs, le rôle et la contribution. La position dans la liste des auteurs est en principe un reflet de la contribution respective d'un auteur (Zuckerman, 1968) et, dans certains cas, de son statut hiérarchique (Pontille, 2004, 2006). Par rôle, nous entendons la catégorie dans laquelle se situe un auteur dans les typologies de rôle dont nous avons discuté à la section 1.2.2.2.3. La position dans la liste des auteurs peut à la fois créer des distinctions artificielles ou masquer des distinctions réelles entre les auteurs d'un article (Mongeon et al. 2017). Nous utilisons donc la typologie de Baerlocher et al. (2007) reprise par Mongeon et al. (2017) pour diviser les auteurs d'un article

en trois groupes : les auteurs principaux, les auteurs superviseurs et les auteurs périphériques. En principe le statut d'inventeur revient de droit aux individus ayant fait une contribution inventive, la nature de la contribution effectuée par un auteur devrait donc être liée à sa nomination en tant qu'inventeur sur le brevet.

Nous avons discuté de l'effet Saint-Mathieu (Merton, 1968) qui fait en sorte que les contributions des chercheurs réputés ont tendance à être surestimées et, inversement, que les contributions des chercheurs les moins connus tendent à être sous-estimées. Nous avons également discuté de l'effet Matilda (Rossiter, 1993) selon lequel les contributions des femmes tendent à être sous-estimées. Nous incluons donc la réputation du chercheur ainsi que le sexe comme facteurs au niveau des individus pouvant influencer l'attribution des statuts d'auteur et d'inventeur. Le Tableau IV (p. 66) présente, en résumé, les quatre niveaux et les neuf facteurs pouvant influencer l'attribution des statuts d'auteur et d'inventeur que nous considérons dans notre étude.

Tableau IV. Facteurs influençant l'attribution des statuts d'auteur et d'inventeur.

Niveau	Facteurs
Discipline	Discipline
Recherche	Impact scientifique potentiel de la recherche
Équipe	Taille de l'équipe Collaboration interinstitutionnelle
Individu	Position dans la liste des auteurs Rôle Contribution Réputation Sexe

Ces différents niveaux et facteurs complètent le cadre théorique bourdieusien et mertonien pour constituer le cadre de référence qui guide et structure la recherche. Dans le chapitre suivant, nous détaillons les méthodes employées dans notre étude afin de déterminer la relation entre les neuf facteurs présentés ci-dessus et afin d'apporter des réponses aux questions de recherche que nous avons énoncées dans le chapitre d'introduction.

Chapitre 2. Méthodologie

Introduction

Dans ce chapitre, nous décrivons en détail les méthodes utilisées pour répondre à nos objectifs et nos questions de recherche, dont voici un rappel :

Objectif 1. Déterminer la relation entre les caractéristiques de la recherche et le nombre d'auteurs et d'inventeurs.

- QR1. Quelle est la relation entre la discipline de recherche et le ratio inventeurs/auteurs des PBA?
- QR2. Quelle est la relation entre l'impact scientifique potentiel de la recherche et le ratio inventeurs/auteurs des PBA?

Objectif 2. Déterminer la relation entre les caractéristiques de l'équipe de recherche et le nombre d'auteurs et d'inventeurs.

- QR3. Quelle est la relation entre la taille de l'équipe de recherche et le ratio inventeurs/auteurs des PBA?
- QR4. Quelle est la relation entre la collaboration interinstitutionnelle et le ratio inventeurs/auteurs des PBA?

Objectif 3. Déterminer la relation entre les contributions respectives des auteurs et l'obtention du statut d'inventeur.

- QR5. Quelle est la relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur?
- QR6. La relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur est-elle liée à la discipline?
- QR7. Quelle est la relation entre le rôle dans l'équipe de recherche et l'obtention du statut d'inventeur?
- QR8. Quelle est la relation entre le type de contribution à la recherche et l'obtention du statut d'inventeur?

Objectif 4. Déterminer la relation entre des caractéristiques individuelles des auteurs et l'obtention du statut d'inventeur.

QR9. Quelle est la relation entre la réputation d'un auteur et l'obtention du statut d'inventeur?

QR10. Quelle est la relation entre le sexe d'un auteur et l'obtention du statut d'inventeur?

Une approche méthodologique quantitative est adoptée pour apporter des réponses à ces questions. Le présent chapitre présente l'approche générale, les sources et le traitement des données, l'analyse des données et les moyens utilisés pour assurer le respect des critères de qualité de la recherche.

2.1. Approche méthodologique générale

Nous avons adopté un devis méthodologique de type corrélationnel prédictif puisque notre revue de la littérature a permis d'identifier des pistes d'hypothèses quant à l'influence de certains facteurs (variables prédictives) sur le ratio inventeurs/auteurs et l'obtention du statut d'inventeur (variables prédites) (Fortin et Gagnon, 2016). Bien que le devis corrélationnel ne permette pas de déterminer s'il existe un lien de causalité entre des variables, il permet néanmoins de déterminer dans quelle mesure il est possible de prédire la valeur d'une variable prédite à partir des valeurs d'une ou de plusieurs variables prédictives (Sproull, 1988).

Notre étude utilise principalement les méthodes bibliométriques. Alan Pritchard a introduit le terme bibliométrie en 1969 et l'a défini comme « l'application de méthodes mathématiques et statistiques aux livres et autres modes de communications » (Pritchard, 1969, p. 349, notre traduction). La bibliométrie comme champ de recherche s'intéresse principalement à l'étude de la science et de la technologie (Moed, 2005) à l'aide d'analyses statistiques de la production, de la diffusion et de l'utilisation des publications savantes et techniques (Baker et Lancaster, 1991).

La principale force de la bibliométrie est qu'elle permet d'analyser un large ensemble de données et d'obtenir des résultats potentiellement généralisables à l'ensemble de la population étudiée, et ce, en profitant de toute la puissance des tests statistiques offerts par les logiciels

d'analyse comme SPSS. Les principales bases de données bibliographiques utilisées en bibliométrie sont le Web of Science (WoS) de Clarivate Analytics et Scopus d'Elsevier (Mongeon et Paul-Hus, 2016). Ces bases de données ont la particularité d'indexer les références et les citations des articles, les adresses institutionnelles de tous les auteurs et les sources de financement de la recherche. De plus, elles couvrent l'ensemble des disciplines, mettant l'emphase sur les revues les plus importantes de chacune d'entre elles (Garfield, 1990); importance mesurée par le nombre de citations qu'elles reçoivent. Certaines disciplines sont toutefois mieux couvertes que d'autres (nous y reviendrons dans le paragraphe suivant). Les données de publication sont de ce fait facilement accessibles, dans la mesure où l'on possède un abonnement à l'une de ces bases de données.

Une des principales limites de la bibliométrie découle de la couverture inégale des différentes disciplines, les sciences sociales et les arts et humanités étant moins bien couverts que les sciences naturelles, le génie et la recherche biomédicale (Moed, 2005; Mongeon et Paul-Hus, 2016). Les publications en anglais sont également beaucoup mieux couvertes que les publications dans d'autres langues (Archambault, Vignola-Gagné, Côté, Larivière et Gingras, 2006; Mongeon et Paul-Hus, 2016). Ces limites sont cependant mitigées par le fait que notre étude porte sur la recherche dans les sciences naturelles et médicales, où les objets de recherche plus internationaux que dans les sciences sociales et humaines ont facilité l'adoption de l'anglais comme langue de publication un peu partout dans le monde (Gingras et Mosbah-Natanson, 2010). De plus, la couverture géographiquement et disciplinairement inégale du WoS est particulièrement problématique lorsque l'objectif est d'évaluer la production ou l'impact scientifique de chercheurs, groupes, institutions ou pays, parce que les publications dans la langue locale (autre que l'anglais) sont sous-représentées dans les bases de données (Archambault et al., 2006) et sont aussi moins citées (Desrochers et Larivière, 2016), par exemple. La couverture inégale du WoS ne constitue pas un problème en soi pour notre thèse puisqu'elle ne vise ni à mesurer ni à comparer la production scientifique ou le nombre de brevets des chercheurs ou de groupes de chercheurs.

Une limite davantage liée à notre recherche découle de l'usage de données bibliométriques en tant qu'observations indirectes de phénomènes. Selon Medawar (1963, cité dans MacRoberts et MacRoberts, 2017), l'article scientifique est un objet qui ne représente pas

adéquatement les processus intellectuels dont il découle. Selon Mitroff (1974, cité dans MacRoberts et MacRoberts, 2017), l'analyse des publications savantes serait un outil peu utile pour comprendre le fonctionnement de la science. La bibliométrie n'offre effectivement pas la même richesse et la même profondeur que des méthodes qualitatives, telles que l'entrevue, pour ce qui est de comprendre les motivations et les particularités des pratiques qui sous-tendent les grandes tendances; toutefois, elle permet de mesurer les phénomènes macrosociologiques, et est ainsi cohérente avec les cadres bourdieusien et mertonien utilisés dans cette thèse. Ainsi, nous analysons ici des résultats sur les pratiques de collaboration, de division du travail et d'attribution des statuts d'auteur et d'inventeur telles qu'elles sont observables à partir des articles scientifiques. Les données utilisées fourniront donc un portrait incomplet du fonctionnement interne des groupes de recherche et ne permettront pas de saisir directement les motivations des individus. Elles permettront néanmoins d'observer des tendances dans les pratiques de collaboration, de division du travail et d'attribution des statuts d'auteur et d'inventeur.

2.2. Sources et traitement des données

Dans cette section, nous présentons les sources de données utilisées ainsi que les traitements effectués afin de former les PBA et obtenir les valeurs pour chacune des variables de notre étude. Nous présentons ensuite les échantillons de PBA et d'auteurs qui seront analysés au Chapitre 3.

2.2.1. Sources de données

Pour notre recherche, nous utilisons les données du WoS comme source d'information sur les articles scientifiques. Ces données incluant le *Science Citation Index Expanded* (SCIE), le *Social Sciences Citation Index* (SSCI) et le *Arts & Humanities Citation Index* (A&HCI) sont hébergées sur un serveur SQL sous forme de base de données relationnelle par l'Observatoire des Sciences et des Technologies (OST) de l'Université du Québec à Montréal. Plusieurs types de documents (p. ex. actes de congrès, comptes rendus de livre, éditoriaux, lettres à l'éditeur) étant indexés dans le WoS, il importe ici de préciser que dans le cadre de notre recherche, le terme « article » réfère spécifiquement aux documents de type « article ». Ces derniers

constituent environ les deux tiers de l'ensemble des documents indexés dans le WoS. Si la couverture du WoS est limitée en sciences humaines et sociales, la base de données couvre un grand nombre de revues du domaine biomédical, des sciences naturelles et du génie, soit les domaines où l'on peut s'attendre à ce que les chercheurs obtiennent des brevets (Archambault et al., 2006; Mongeon et Paul-Hus, 2016). En effet, la couverture du WoS est estimée à 28 % et 33 % de l'ensemble des revues académiques publiées dans le monde dans le domaine biomédical et en sciences naturelles et génie, respectivement (Mongeon et Paul-Hus, 2016). La couverture de la base de données n'est pas exhaustive, mais l'objectif du WoS est d'avoir une couverture sélective des revues les plus importantes dans chaque discipline (Garfield, 1990), suivant la loi de Bradford (1934) selon laquelle la majorité des articles sur un sujet particulier sont publiés dans un petit nombre de revues principales. L'identification de ces revues principales est basée sur le nombre de références faites à ces revues dans la littérature (Moed, 2005). En analysant l'ensemble des références des articles de 2002 indexés dans le WoS, Moed (2005) a montré que 75 % des références citées dans les articles du WoS sont des articles de revues également indexées dans le WoS. Cette proportion varie d'une discipline à l'autre, allant de 17 % en arts et humanités à 92 % en biologie moléculaire et biochimie (Moed, 2005). Selon ces données de 2002, le génie est la discipline la moins bien couverte parmi celles qui sont d'intérêt pour notre étude, avec 46 % des références citées étant des articles de revues indexées dans le WoS. Il cependant difficile d'évaluer la couverture du WoS en termes d'articles, le nombre total d'articles par discipline publiés dans le monde demeurant inconnu.

Les notices bibliographiques du WoS comportent de nombreux champs, mais notre recherche n'en utilise que quelques-uns qui sont pertinents pour notre recherche, soit l'identifiant unique de l'article, le titre de l'article, le résumé, l'année de publication, la revue de publication, le facteur d'impact de la revue, le nom des auteurs, le prénom des auteurs et l'ordre des auteurs. Nous utilisons également quelques champs additionnels qui sont disponibles dans la base de données de l'OST et qui contiennent des informations calculées, dans plusieurs cas, à l'aide des champs des notices du WoS (voir l'Annexe 4, p. xxiv, pour une description des champs inclus dans l'analyse, et l'Annexe 5, p. xxvi, pour une représentation schématique de ces champs).

Pour ce qui est des informations sur les brevets, nous utilisons les données du *USPTO Patent Full-Text and Image Database* (PatFT) (<http://patft.uspto.gov/>), également hébergées dans une base de données relationnelle par l'OST. La base de données PatFT contient le texte intégral de l'ensemble des brevets (plus de cinq millions) octroyés par le USPTO depuis 1976.

Puisque la décision de protéger une invention dans un pays donné est principalement économique (Anderson, Cisló, Saavedra et Cameron, 2013), et que les États-Unis sont la plus importante économie dans le monde, la PatFT constitue une source de données relativement exhaustive sur la prise de brevets par les chercheurs. Par exemple, selon les statistiques de l'Organisation mondiale de la propriété intellectuelle (OMPI) (<http://www.OMPI.org/>), le USPTO octroie en moyenne 20,4 % de l'ensemble des brevets octroyés annuellement depuis 1985 dans le monde. Il s'agit en fait d'une sous-estimation de la proportion des inventions mondiales couvertes par la base de données PatFT puisqu'une même invention peut être brevetée dans plusieurs pays.

Dans la base de données relationnelle de l'OST, les données de la PatFT contiennent de nombreux champs, qui sont répartis dans des dizaines de tables. Dans le cadre de notre recherche, seuls quelques champs qui sont en lien avec notre recherche sont utilisés, soit le titre, le résumé, l'année de dépôt, le nombre d'inventeurs, et le nom des inventeurs (voir l'Annexe 4 [p. xxiv] et l'Annexe 5 [p. xxvi] pour plus de détails sur les champs utilisés dans notre étude et la structure de notre base de données relationnelle).

Ces deux sources de données contiennent l'ensemble des informations requises pour former les PBA et répondre à la plupart de nos questions de recherche. Cependant, la question de recherche QR8 requière des informations concernant les contributions respectives des auteurs des articles faisant l'objet d'une PBA, et celles-ci ne sont pas directement accessibles à partir des notices bibliographiques. Nous avons donc, avec l'aide de deux assistantes de recherche, collecté manuellement ces informations à partir de la version intégrale des articles, soit dans la déclaration des contributions lorsque cette section est présente dans les articles, ou dans la section des remerciements.

2.2.2. Traitement des données

Nous décrivons ci-dessous les méthodes utilisées pour 1) former les PBA, 2) identifier les caractéristiques des PBA, 3) collecter les informations sur le rôle ainsi que sur les contributions respectives des auteurs de chacun des articles, et enfin 4) dresser le profil des auteurs.

2.2.2.1. Formation des PBA

Nous avons collecté l'ensemble des notices bibliographiques des articles scientifiques indexés dans le WoS et publiés entre 1991 et 2016 à partir de la base de données de l'OST (N = 24 945 618). Nous avons fixé la limite inférieure de la plage temporelle à 1991 puisque les résumés des articles, que nous utilisons pour l'appariement des articles et des brevets, n'ont été systématiquement inclus dans les notices bibliographiques qu'à partir de cette date. La limite supérieure de la plage temporelle a été fixée à 2016, car les articles publiés en 2017 n'étaient pas disponibles dans la base de données de l'OST au moment de la collecte de nos données. Nous avons également recueilli toutes les notices bibliographiques des brevets d'invention de la base de données PatFT dont le dépôt a été effectué entre 1991 et 2015 (N = 4 052 421). La limite supérieure de la plage temporelle pour les brevets est fixée à 2015 parce que les données de 2016 et 2017 n'étaient pas disponibles au moment de la collecte.

Pour appairer les brevets et les articles issus de la même recherche, nous avons développé un algorithme s'inspirant de celui utilisé par Magerman et al. (2015) et que nous décrivons dans la présente section. Ce choix repose sur le fait que l'algorithme permet de créer des PBA automatiquement sur la base de mesures pouvant être appliquées à un très grand nombre de documents. Contrairement à la méthode utilisée par Haeussler et Sauermann (2013) qui nécessite de solliciter les chercheurs pour obtenir les PBA, ou encore la méthode manuelle de Ducor (2000) qui a fait des requêtes une à une dans des moteurs de recherche d'articles et de brevets, la méthode choisie n'a pas de contrainte au niveau de la quantité de données traitées. De plus, puisque la validité des PBA est en grande partie fondée sur des mesures automatisées, une telle méthode ne nécessite pas d'expertise ou de connaissances particulières des domaines de recherche des articles et des brevets appariés. Il s'agit d'une méthode hybride, selon la

typologie présentée au Chapitre 1, puisqu'elle utilise les trois éléments d'information suivants pour former les PBA :

1. l'année de publication des articles et l'année de dépôt des demandes de brevet,
2. le nom des auteurs et des inventeurs,
3. le contenu du titre et du résumé des articles et des brevets.

Dans les sections suivantes, nous présentons plus en détail les critères que nous avons utilisés pour former les PBA, ainsi que le processus par lequel nous avons validé les PBA formées par l'algorithme.

Nous avons limité l'appariement aux articles publiés au plus tôt un an avant la demande de brevet puisque, en principe, des résultats de recherche déjà publiés ne peuvent faire l'objet d'une demande de brevet puisqu'ils font partie du domaine public. Cependant, la loi américaine accorde un délai d'un an suivant la publication pour le dépôt d'une demande de brevet. Cela permet de réduire le nombre de PBA potentielles en éliminant les articles publiés plus d'un an avant la date de dépôt de la demande de brevet. Nous avons également choisi de conserver pour l'appariement uniquement les articles publiés dans les deux années suivant le dépôt de la demande de brevet, afin de limiter davantage le nombre de PBA potentielles. Par exemple, pour un brevet dont la demande a été déposée au cours de l'année 2005, nous ne conservons pour l'appariement que les articles publiés entre 2004 et 2007 inclusivement. Cette fenêtre de deux ans suivant la demande de dépôt du brevet est également utilisée par Murray (2002) et représente un délai raisonnable permettant d'éliminer des faux positifs sans éliminer trop de PBA valides.

Pour qu'un article et un brevet soient considérés comme une PBA potentielle par notre algorithme, au moins un auteur de l'article doit être inventeur du brevet. Dans la base de données PatFT, le nom des inventeurs a la forme « nom de famille; prénom » (p. ex. Lapointe; Jérémie), alors que dans la base de données du WoS, le nom des auteurs a la forme « nom de famille-initiales » (p. ex. Lapointe-J)³³. Il est donc nécessaire d'uniformiser le champ inventeur et le champ auteur afin d'identifier les articles dont un des auteurs figure dans la liste des inventeurs d'un brevet. Nous avons ainsi modifié le nom des inventeurs pour qu'ils prennent la forme

³³ Les prénoms entiers des auteurs ont commencé à être inclus dans les notices bibliographiques du WoS à partir de 2008 seulement, mais cette information n'est pas disponible pour tous les articles. Même après 2008, le champ prénom est parfois vide et contient parfois seulement des initiales.

« nom de famille-initiales ». Un nom d’auteur peut avoir plusieurs initiales (ex. Larose-DP), mais comme il n’y a pas de contrôle d’autorité effectué dans les notices du WoS et de la PatFT, il est possible que le nom d’un même individu soit inscrit avec une seule initiale dans une notice et avec plusieurs initiales dans d’autres. Afin de maximiser le rappel lors de l’appariement entre les noms d’auteur et les noms d’inventeur, nous avons donc choisi d’utiliser seulement la première initiale pour l’appariement initial. Ainsi, un article ayant comme auteur « Larose-DP » et un brevet ayant comme inventeur « Larose; Daniel » (Larose-D, dans sa forme modifiée) sont considérés comme constituant une PBA potentielle.

Les syntagmes nominaux (SN)³⁴ contenus dans le titre et le résumé d’un article et d’un brevet sont utilisés pour mesurer la similarité entre les deux documents. L’identification des SN a été effectuée à l’aide du module OpenNLP d’Apache (<http://incubator.apache.org/opennlp/>) un outil bien connu et largement utilisé. Le processus se décline en trois étapes. La première étape est l’analyse morphosyntaxique (identification des verbes, noms, adjectifs, etc.). Les SN sont ensuite identifiés par un filtre sélectionnant les séquences de mots comprenant exclusivement des noms et des adjectifs et se terminant par un nom (il s’agit de la forme type d’un syntagme nominal en anglais³⁵). Finalement, les SN au pluriel sont convertis au singulier. L’ensemble des étapes d’identification des SN dans les titres et les résumés de tous les articles du WoS et de tous les brevets de la PatFT ont été effectuées par le Centre for Science and Technology Studies (CWTS) de l’Université de Leyde. Le Tableau V (p. 76) présente un exemple d’identification des SN dans le titre d’un article.

³⁴ Un syntagme est défini comme « un groupe d’éléments qui forment une unité dans une organisation hiérarchisée » (<http://www.larousse.fr/dictionnaires/francais/syntagme/76215>. Consulté le 12 juillet 2017).

³⁵ Bien que cette méthode soit adaptée à la langue anglaise (en français, par exemple, les SN peuvent se terminer par un adjectif) il ne s’agit pas d’une limite dans notre cas puisque les notices bibliographiques du WoS sont entièrement en anglais, même pour les articles en d’autres langues.

Tableau V. Exemple d'extraction des SN dans le titre d'un article.

Titre de l'article	Syntagme nominal (SN)	Occurrences
Analysis of transport phenomena in cellulose diacetate membranes. 4. Dependence of membrane desalination properties of annealing temperature - molecular analysis	Analysis	1
	Transport phenomena	1
	Cellulose diacetate membrane	1
	Dependence	1
	Membrane desalination property	1
	Temperature	1
	Molecular analysis	1
Total		7

Chaque SN en commun entre un article et un brevet constitue un lien entre les deux documents. Ainsi, le nombre de liens (SN en commun) permet de mesurer la similarité textuelle entre un brevet et un article. Nous utilisons une méthode similaire à celle de Magerman et al. (2015) présentée au Chapitre 1 afin de calculer un score de similarité pour chaque PBA. Ce score correspond à la moyenne de deux valeurs, soit 1) le nombre de SN en commun divisé par le nombre total de SN dans le titre et le résumé de l'article et 2) le nombre de SN en commun divisé par le nombre total de SN dans le titre et le résumé du brevet, ce qu'illustre la formule mathématique suivante :

$$Similarité = \frac{x(y + z)}{2yz}$$

où

x est le nombre de SN en commun dans le titre et le résumé de l'article et du brevet.

y est le nombre total de SN dans le titre et le résumé du brevet.

z est le nombre total de SN dans le titre et le résumé de l'article.

Pour déterminer le score de similarité minimum requis pour qu'un article et un brevet soient identifiés comme une PBA potentielle, nous avons procédé à une série de formation et de validation de PBA. Nous avons fixé un seuil élevé (Similarité $\geq 0,90$), puis une assistante de recherche a effectué une vérification manuelle des 67 PBA ainsi formées en suivant une série de deux étapes que nous avons préalablement élaborées :

1. Les paires auteur-inventeur ont été validées à l'aide des prénoms. Lorsque le prénom de l'auteur n'était pas disponible dans la notice de l'article, il a été obtenu en

consultant l'article (si disponible). Une PBA est rejetée lorsqu'aucune de ses paires auteur-inventeur n'est valide.

2. Les titres (et les résumés au besoin) de l'article et du brevet sont utilisés pour confirmer que les deux documents traitent du même sujet. Une PBA est rejetée lorsque les deux documents ne semblent pas traiter du même sujet.

L'ensemble des 67 PBA identifiées avec un seuil minimum de similarité de 0,90 étant valide, nous avons procédé à une seconde étape d'appariement en réduisant le seuil à 0,80, puis procédé à une nouvelle validation manuelle, puis réduit le seuil à 0,70, et ainsi de suite. Le Tableau VI (p. 77) présente les résultats de la validation des PBA en fonction du score de similarité.

Tableau VI. Nombre et proportion de PBA valides en fonction du score de similarité.

Similarité	Nombre de PBA potentielles	PBA valides	Proportion (%)
>=0,90	67	67	100,0
0,80-0,89	221	221	100,0
0,70-0,79	310	310	100,0
0,60-0,69	484	484	100,0
0,50-0,59	786	786	100,0
0,40-0,49	2 636	1 706	64,7
0,30-0,39	30 290	14 209	46,9
Total	34 794	17 783	51,1

Le processus a été interrompu après la vérification manuelle des 30 290 PBA dont le score de similarité est situé entre 0,30 et 0,39 et dont 46,9 % ont été identifiés comme étant valides. La précision relativement faible de l'algorithme à ce niveau de similarité (46,9 %), ainsi que la tendance observée de décroissance de la proportion de PBA valides lorsque le score de similarité diminue, laissent présumer que la majorité des PBA sous ce seuil seront des faux positifs. De plus, en incluant les PBA avec un score de similarité relativement faible, nous risquerions d'inclure dans nos données un certain nombre d'articles et de brevets traitant de sujets semblables, mais n'étant pas réellement issus de la même recherche, ce qui pourrait réduire la validité des résultats de notre étude.

En somme, les critères utilisés pour la formation des 17 783 PBA indiqués dans le Tableau VI (p. 77) sont les suivants :

1. L'article est publié au plus tôt un an avant et au plus tard deux ans après le dépôt de la demande de brevet;
2. La vérification des noms et prénoms des auteurs et inventeurs permet d'identifier au moins un des auteurs de l'article figurant parmi les inventeurs du brevet;
3. Le score de similarité de la PBA est supérieur ou égal à 0,30;
4. La vérification manuelle des titres des articles et brevets appariés indique que les deux documents ont un contenu similaire.

L'appariement est basé sur les noms des individus apparaissant sur les deux documents, sur leur proximité temporelle, ainsi que sur la similarité de leur contenu. Il peut donc arriver que deux articles soient appariés au même brevet (et vice-versa), mais que certains de ces appariements soient basés sur une plus grande similitude, tant au niveau du contenu de l'article et du brevet (score de similarité) qu'au niveau des listes d'auteurs et d'inventeurs (nombre d'inventeurs figurant dans la liste des auteurs). Prenons l'exemple des chercheurs A, B et C ayant inventé une nouvelle méthode de formation de cristaux. Ils publient un article X dont A, B et C sont les auteurs, et un brevet Y où il est indiqué qu'A et C sont inventeurs. Donc X et Y forment une PBA. En toute logique, puisque A, B et C font de la recherche sur les cristaux, il est possible que nous repérions dans le WoS d'autres articles signés par A, B, ou C et possiblement en collaboration avec d'autres chercheurs, et dont le contenu sera similaire à X. Par exemple, A peut avoir écrit un autre article (Z) sur les cristaux en collaboration avec les chercheurs E et F. Dans un tel cas, il est possible que l'algorithme forme, en plus de la PBA YX, une seconde PBA YZ. Or, il est clair qu'ici, seule la PBA YX nous intéresse. Donc, afin de maximiser la validité de nos PBA en nous assurant autant que possible que les articles et les brevets formant les PBA ne sont pas seulement similaires, mais effectivement issus de la même recherche, nous avons effectué, dans l'ordre présenté, une série d'étapes supplémentaires visant à ce que chaque article soit apparié à un seul brevet et vice-versa :

1. Nous conservons la PBA pour laquelle le nombre d'auteurs qui sont aussi inventeurs est le plus élevé.

2. Nous conservons la PBA dont le taux de similarité est le plus élevé.
3. S'il demeure des cas où plusieurs articles sont appariés à un seul brevet ou vice-versa, nous conservons une PBA aléatoire.

Dans certains cas, il se peut qu'un ou plusieurs inventeurs des PBA ne figurent pas parmi les auteurs de l'article. Prenons par exemple, une PBA formée d'un article X dont les auteurs sont A, B, et C et d'un brevet Y dont les inventeurs sont C, D et E. Il est possible qu'A, B, C, D et E aient travaillé ensemble sur la recherche ayant menée au brevet, mais que les inventeurs D et E, pour une raison quelconque, n'aient pas participé à la publication et ne soient pas auteurs. Cependant, il est aussi possible que le chercheur C ait travaillé avec A et B sur un projet ayant mené à l'article X, et sur un autre projet avec D et E ayant mené au brevet Y. Comme il est difficile de déterminer de quel cas il s'agit, nous avons choisi d'éliminer autant que possible ce type d'ambiguïté en ne conservant que les PBA dont tous les inventeurs figurent dans la liste des auteurs. Suite à l'ensemble de ces étapes, notre jeu de données final comporte un total de 6 312 PBA.

Notre méthode diffère de celle de Magerman et al. (2015) à quelques égards. D'abord, elle utilise les SN au lieu des mots, ce qui la rend plus précise à cet égard. Le calcul du score de similarité et la détermination du seuil d'inclusion diffèrent également. Margeman et al. (2015) considère une PBA valide lorsque le nombre de mots en commun entre les deux documents divisé par le nombre de mots du document le plus court est égal ou supérieur à 0,60, et lorsque le nombre de mots en commun entre les deux documents divisé par le nombre de mots du document le plus long est égal ou supérieur à 0,30. Notre méthode utilise plutôt la moyenne des deux valeurs et est donc plus inclusive et augmente le risque de faux positif. Ce risque est cependant mitigé par l'étape de validation manuelle des PBA. De plus, nous ne fixons pas de seuils minimums pour chacun des documents, mais plutôt un seul seuil pour le score de similarité, soit 0,30. Notre méthode est à cet égard plus inclusive que celle de Magerman et al. (2015) dont la moyenne minimale des deux mesures utilisées est de 0,45. En réduisant ainsi le seuil de similarité pour l'identification des PBA, nous obtenons un rappel supérieur à celui de Magerman et al. (2015), mais la précision de notre algorithme est en contrepartie réduite. Cette perte de précision est cependant contrecarrée par la série d'étapes supplémentaires de validation et de sélection des PBA que nous venons de décrire.

Maintenant que nous avons identifié un ensemble de PBA pour notre étude, nous sommes en mesure de recueillir les informations sur les articles et les brevets qui nous permettront de répondre à nos questions de recherche QR1 à QR4 (discipline, impact scientifique potentiel, taille de l'équipe de recherche et collaboration interinstitutionnelle), ainsi que les informations sur les auteurs qui nous permettront de répondre aux questions de recherche QR5 à QR10 (position dans la liste des auteurs, discipline, rôle, contribution, sexe et réputation). Dans les sections suivantes, nous décrivons la façon dont nous avons opérationnalisé ces concepts et recueilli ces données.

2.2.2.2. Discipline des PBA (QR1 et QR6)

Nous avons utilisé la classification disciplinaire de la NSF pour attribuer automatiquement une discipline à chacun des articles et, par le fait même, à la PBA dont l'article fait partie. Cette classification comporte 144 spécialités, regroupées en quatorze disciplines et deux grands champs disciplinaires (sciences sociales et humanités d'une part, sciences naturelles, d'autre part). Dans la base de données WoS hébergée par l'OST, une discipline unique est assignée à chacune des revues, mais non aux articles individuels. Si la discipline attribuée à la revue reflète généralement bien le contenu de cette dernière, il est néanmoins possible qu'une revue publie des articles provenant de diverses disciplines, ce qui rend la tâche d'identifier la discipline de l'article un peu plus complexe. C'est le cas, par exemple, de revues multidisciplinaires comme *PNAS*, *Nature*, *Science* et *PLoS ONE*. Ainsi, afin de maximiser la précision de la classification disciplinaire des PBA, nous avons déterminé la discipline d'un article en identifiant la discipline citée le plus souvent dans l'article. Par exemple, un article citant dix articles publiés dans des revues de chimie et cinq articles publiés dans des revues de physique sera considéré comme un article de chimie. Cette méthode présuppose 1) qu'il est plus probable que les articles cités dans un article de chimie soient également des articles de chimie que des articles d'une autre discipline, et 2) qu'il est plus probable qu'un article publié dans une revue de chimie soit un article de chimie qu'un article d'une autre discipline. Ainsi, cette méthode permet de catégoriser un article à partir de plusieurs observations (les références) plutôt qu'une seule (la revue dans laquelle l'article est publié), ce qui réduit le risque d'erreur. Dans le cas où un article cite deux disciplines à parts égales dont une des deux correspond à la discipline de la revue dans laquelle l'article est publié, nous avons attribué à l'article cette discipline. Dans

31 cas où aucune des deux disciplines citées à parts égales ne correspond à la discipline de la revue dans laquelle l'article est publié, nous avons consulté l'article afin de lui attribuer manuellement une discipline.

Le Tableau VII (p. 81) présente le nombre de PBA identifiées par discipline. Afin d'obtenir des échantillons disciplinaires de taille suffisante pour notre analyse, nous divisons les disciplines en quatre groupes disciplinaires : recherche biomédicale, sciences naturelles, génie, et sciences sociales. Aucune PBA n'a été repérée en sciences humaines, d'où l'absence de ce groupe disciplinaire dans la liste.

Tableau VII. Nombre de PBA, d'articles et de brevets par discipline.

Groupe disciplinaire	Discipline	Nombre de PBA	Proportion (%)
Recherche biomédicale	Médecine clinique	1 332	21,1%
	Recherche biomédicale	1 691	26,8%
	Santé	5	0,1%
	<i>Sous-total</i>	<i>3 028</i>	<i>48,0%</i>
Sciences naturelles	Chimie	854	13,5%
	Physique	953	15,1%
	Biologie	109	1,7%
	Sciences de la terre et de l'espace	43	0,7%
	Mathématique	26	0,4%
	<i>Sous-total</i>	<i>1 985</i>	<i>31,4%</i>
Génie	Génie	1 279	20,3%
	<i>Sous-total</i>	<i>1 279</i>	<i>20,3%</i>
Sciences sociales	Champs professionnels	12	0,2%
	Sciences sociales	3	0,0%
	Psychologie	5	0,1%
	<i>Sous-total</i>	<i>20</i>	<i>0,3%</i>
Total		6 312	100,0%

Seulement 20 PBA ont été classées dans le groupe disciplinaire « sciences sociales ». Comme il s'agit d'un très petit nombre de PBA, et comme les phénomènes sociaux peuvent plus difficilement faire l'objet d'un brevet, nous avons consulté les articles afin de reclasser les PBA dans un des trois autres groupes disciplinaires. Les vingt PBA ont ainsi été reclassées en génie.

Dans le reste du texte, le terme « discipline » fait donc référence aux trois groupes disciplinaires restants. Cette information est consignée dans la variable « gr_disc ».

2.2.2.3. Impact scientifique potentiel de la recherche (QR2)

Tel que mentionné dans la conclusion du Chapitre 1, nous utilisons le facteur d'impact de la revue comme mesure de l'impact scientifique potentiel d'un article au moment de sa publication. Cette information est disponible directement dans la base de données de l'OST. Pour chaque article, nous utilisons le facteur d'impact (FI) calculé avec une fenêtre de citations de deux ans³⁶, et consignons cette information dans la variable « fi_2ans ».

2.2.2.4. Taille de l'équipe de recherche (QR3)

Bien qu'il puisse arriver que des individus ayant contribué à un travail de recherche ne figurent pas parmi les auteurs (remerciés ou auteurs fantômes) ou vice-versa (auteurs honorifiques), les termes « équipe » ou « équipe de recherche » font référence aux auteurs d'un article dans le présent document. Ainsi, la taille de l'équipe se mesure par le nombre d'auteurs de l'article et cette information est consignée, pour chaque article, dans la variable « nb_auteurs ».

2.2.2.5. Collaboration interinstitutionnelle (QR4)

Dans la présente recherche, un article est considéré comme étant le fruit d'une collaboration interinstitutionnelle lorsqu'au moins deux institutions différentes apparaissent dans la liste des adresses de l'article. Pour chaque article, cette information est consignée dans la variable dichotomique « multi_institutions ».

2.2.2.6. Position dans la liste des auteurs (QR5)

Comme nous l'avons vu dans le Chapitre 1, l'ordre des auteurs d'un article est en partie guidé par des normes disciplinaires et peut donc nous renseigner sur la nature et l'ampleur de la contribution respective de chacun des auteurs d'un article. La première et la dernière position

³⁶ Le FI d'une revue pour une année donnée est égal au nombre moyen de citations reçues au cours de l'année en question par les articles publiés dans les deux années précédentes. Par exemple, un FI de 3,875 en 2015 signifie que les articles publiés dans la revue en 2013 et 2014 ont été cités en moyenne 3,875 fois en 2015.

dans la liste sont généralement les plus importantes (sauf lorsque les auteurs sont en ordre alphabétique), et constituent donc en quelque sorte deux pôles de capital. Nous avons donc créé une variable « position » ayant sept modalités (premier, deuxième, troisième, milieu, antépénultième, pénultième et dernier). Une de ces modalités est attribuée à chacun des auteurs de chacun des articles en fonction de la position qu'il occupe par rapport au début ou à la fin de la liste. Par exemple, lorsqu'un auteur est à plus de trois positions du début et de la fin de la liste, la variable prend la modalité « milieu ». Cette modalité est donc utilisée seulement pour les articles ayant sept auteurs ou plus. Le Tableau VIII (p. 83) présente l'exemple d'un article à huit auteurs.

Tableau VIII. Exemple des modalités de la variable « position » pour chacun des auteurs d'un article ayant huit auteurs.

Position de l'auteur dans la liste	Position dans la liste des auteurs (position)
1	Premier
2	Deuxième
3	Troisième
4	Milieu
5	Milieu
6	Antépénultième
7	pénultième
8	Dernier

Pour un article ayant quatre auteurs, seules les modalités « premier », « deuxième », « pénultième », et « dernier » seront attribuées. Les articles signés par trois ou cinq auteurs sont des cas particuliers puisqu'un des auteurs est à distance égale de la première et de la dernière position. Dans ces cas, les auteurs du centre ont été divisés aléatoirement en deux groupes égaux pour éviter la surreprésentation d'un groupe par rapport à l'autre. Pour les articles à trois auteurs nous avons attribué à un de ces groupes la modalité « deuxième » et à l'autre, la modalité « pénultième ». Pour les articles à cinq auteurs, nous avons attribué à un groupe la modalité « troisième » la modalité « antépénultième » à l'autre groupe.

2.2.2.7. Rôle (QR7)

Comme nous l'avons vu au Chapitre 1 (section 1.2.2.2.3), Baerlocher et al. (2007) ont proposé une typologie pour catégoriser les auteurs en recherche biomédicale, dont nous nous inspirons pour attribuer un des trois rôles suivants aux auteurs de chacun des articles de notre jeu de données :

- **Principal** : auteurs apparaissant généralement au début de la liste et ayant effectué des tâches importantes.
- **Périphérique** : auteurs apparaissant généralement au milieu de la liste, et dont les contributions sont de moindre ampleur ou importance.
- **Superviseur** : auteurs apparaissant généralement à la fin de l'article et ayant effectué des tâches de supervision.

Au-delà du premier auteur (qui est typiquement un auteur principal) et du dernier auteur (qui est typiquement un auteur superviseur) (Pontille, 2004, 2006), il est difficile d'identifier le rôle des auteurs. Comme nous l'avons vu au Chapitre 1, l'ordre alphabétique peut être utilisé pour signaler une distinction entre les contributions d'un auteur ou d'un groupe d'auteurs (lorsque seule une partie de la liste des auteurs est en ordre alphabétique) (Mongeon et al., 2017; Zuckerman, 1968). Ainsi, il peut arriver que la liste des auteurs d'un article contienne une séquence d'auteurs en ordre alphabétique au milieu de l'article, et permette ainsi de mettre en valeur la contribution des auteurs qui apparaissent avant ou après la séquence alphabétique. Nous opérationnalisons donc les trois types de rôles de la façon suivante :

- **Principal** : tous les auteurs apparaissant en début de liste et avant une séquence alphabétique que nous qualifions « d'intentionnelle » (voir plus bas).
- **Périphérique** : tous les auteurs faisant partie d'une séquence alphabétique intentionnelle précédée et suivie d'au moins un auteur.
- **Superviseur** : tous les auteurs apparaissant en fin de liste et après une séquence d'auteurs en ordre alphabétique intentionnelle.

Afin de nous assurer que la séquence alphabétique est le fruit de la décision de l'équipe de recherche quant à l'ordre des auteurs (et non du hasard), et qu'il s'agit donc effectivement d'un moyen utilisé par les chercheurs pour signaler une distinction entre différents groupes

d'auteurs, nous ne tenons compte que des séquences alphabétiques ayant moins de 5 % de chance d'être dues au hasard. Cette probabilité est fonction du nombre total d'auteurs d'un article et du nombre d'auteurs formant la séquence ordre alphabétique (voir Mongeon et al. [2017] pour plus de détail sur le calcul de cette probabilité). Ce seuil étant fixé à 5 %, seules les séquences alphabétiques d'au moins cinq auteurs sont prises en compte. En conséquence, la détection de l'ordre alphabétique partiel est limitée aux articles ayant sept auteurs et plus puisque nous ne considérons que les articles avec au moins un auteur avant et après la séquence. Les articles de moins de 7 auteurs ne sont pas codés et ne sont pas inclus dans l'analyse visant à répondre à QR7. Le rôle attribué à chacun des auteurs d'un article est consigné dans la variable « rôle ».

Le Tableau IX (p. 86) donne l'exemple d'une liste d'auteurs comportant une séquence alphabétique à partir de laquelle nous pouvons identifier les auteurs principaux, périphériques et superviseurs. On voit qu'il y a clairement un sous-groupe d'auteurs au centre de l'article qui se distinguent des deux premiers et des deux derniers auteurs du fait qu'ils forment une séquence alphabétique, cette dernière étant par ailleurs assez longue (cinq auteurs) pour qu'on puisse présumer qu'il s'agit d'un choix délibéré de la part des auteurs. On remarque que les deux premiers et les derniers auteurs dans l'exemple sont également en ordre alphabétique, mais forment des séquences distinctes de celle formée par les auteurs de la troisième à la septième position parce que Demirkol précède Ozer dans l'ordre alphabétique, et parce que Gokcay précède Shin. Les séquences formées par les deux premiers et les deux derniers auteurs sont ignorées, car elles comportent seulement deux auteurs et ont donc une probabilité de 50 % d'être le fruit du hasard.

Tableau IX. Exemple d'article comportant une séquence d'auteurs périphériques listés en ordre alphabétique.

Position dans la liste des auteurs	Nom de l'auteur	Rôle
1	Aoyama, Y	Principal
2	Ozer, I	
3	Demirkol, M	Périphérique
4	Ebara, T	
5	Murase, T	
6	Podskarbi, T	
7	Shin, YS	
8	Gokcay, G	Superviseur
9	Okubo, M	

Parmi nos PBA, nous avons identifié un total de 92 articles contenant une séquence en ordre alphabétique intentionnelle précédée d'au moins un auteur principal et suivi d'au moins un auteur superviseur. Le Tableau X (p. 86) présente le nombre et la proportion d'auteurs de chaque rôle et pour chaque discipline. L'échantillon de PBA utilisé pour l'analyse de la relation entre le rôle et le statut d'inventeur est grandement réduit par rapport au 6 312 PBA utilisées pour l'analyse précédente. Puisque nous analysons les PBA de chaque discipline séparément, nous avons choisi de limiter l'analyse aux PBA en recherche biomédicale (N = 61) pour assurer la puissance statistique des tests qui seront effectués pour répondre à QR7.

Tableau X. Nombre et proportion des auteurs par rôle et par discipline pour les articles en ordre alphabétique partiel.

Discipline	Principal		Périphérique		Superviseur		Total
	N	%	N	%	N	%	N
Recherche biomédicale	119	22,4	291	54,8	121	22,8	531
Sciences naturelles	60	38,7	56	36,1	39	25,2	155
Génie	9	39,1	10	43,5	4	17,4	23
Total	188	26,5	357	50,4	164	23,1	709

2.2.2.8. Contribution à la recherche (QR8)

De plus en plus de revues (surtout dans le domaine biomédical) exigent que les auteurs déclarent leurs contributions respectives aux articles publiés. Ces déclarations deviennent une

source d'information supplémentaire pour déterminer qui a fait quoi. Dans notre thèse, le concept de contribution est défini comme le type de contribution effectué par l'auteur dans un article tel qu'indiqué dans la déclaration des contributions.

Comme ces déclarations ne figurent pas dans les notices bibliographiques du WoS, nous avons consulté la version intégrale des articles afin d'en extraire la contribution de chacun des auteurs. Cette information se trouve parfois dans une section de l'article dédiée aux contributions des auteurs et parfois dans la section des remerciements. Puisqu'il s'agit d'un processus plutôt long, nous avons choisi de cibler ici les articles du domaine biomédical. La déclaration des contributions est d'ailleurs souvent exigée par les revues de cette discipline, alors que cette pratique est moins fréquente dans les autres disciplines. Sur les 3 028 articles formant les PBA en recherche biomédicale, 112 comprenaient une déclaration de la contribution des auteurs, ce qui nous a permis d'identifier les contributions des 994 auteurs de ces articles. Au total, 2 437 libellés ont été recueillis, dont 454 libellés distincts. Vingt-et-un (21) articles n'ont pas été traités, l'accès au plein texte et donc aux déclarations de contributions n'étant pas possible via les bibliothèques de l'Université de Montréal.

Les libellés utilisés pour décrire les contributions respectives des auteurs peuvent varier d'un article à l'autre. Par exemple, certains libellés sont plus spécifiques (p. ex. « did the pp. 28-30 modeling and contributed to its analysis ») alors que d'autres sont plus généraux (p. ex. « wrote the paper », « designed the research », « performed the research »). Nous avons donc classé chacun des libellés de contribution dans une des six catégories suivantes, tirées de l'étude de Larivière et al. (2016) :

- Conception
- Réalisation
- Analyse
- Rédaction
- Matériel
- Autres

La classification de l'ensemble des libellés de contribution a été effectuée par une assistante de recherche. Afin de vérifier la validité de la classification, un test intercodeur a été

effectué avec l'aide d'une assistance de recherche sur un échantillon aléatoire de 117 libellés, soit 25,8 % des 454 libellés distincts. Un taux d'accord de 87,2 % et un kappa de Cohen (1960) de 85 % ont été obtenus, ce qui est au-delà du taux de 75 % généralement jugé satisfaisant (Landis et Koch 1977). Le Tableau XI (p. 88) présente le nombre de libellés de contribution par catégorie.

Tableau XI. Nombre et proportion de libellés de contribution par catégorie.

Catégorie	Nombre de libellés (N)	Proportion des libellés (%)
Conception	342	14,0
Réalisation	731	30,0
Analyse	424	17,4
Rédaction	615	25,2
Matériel	174	7,1
Autres	151	6,2
Total	2 437	99,9

Comme un auteur peut avoir effectué plusieurs contributions pour un même article, cette information est consignée dans six variables dichotomiques représentant chacune un type de contribution.

2.2.2.9. Réputation (QR9)

Dans notre étude, le concept de réputation d'un auteur équivaut à son capital scientifique accumulé, que nous mesurons à l'aide du nombre de citations reçues par un auteur à l'année du dépôt de la demande de brevet. La variable « nb_citations » contient cette information pour chaque auteur de chaque article.

Cela requiert que nous identifions d'abord l'ensemble des articles publiés par chacun des auteurs des articles formant les PBA. Or, l'identification de l'ensemble des publications d'un individu n'est pas une tâche simple. Cela requiert un travail énorme de désambiguïsation des noms d'auteurs, à cause notamment des problèmes d'homonymie (plusieurs chercheurs peuvent avoir le même nom). La tâche est par ailleurs grandement alourdie par le fait que le prénom complet des auteurs n'est pas inclus dans les notices bibliographiques du WoS avant 2008 (seules les initiales le sont). Le nombre de noms d'auteurs distincts pour l'ensemble des PBA étant trop élevé pour faire une désambiguïsation manuelle, nous avons utilisé une méthode

de désambiguïsation automatique développée par Caron et van Eck (2014)³⁷. Cet algorithme se base sur de nombreuses informations disponibles dans les notices bibliographiques, ainsi que sur l'analyse des réseaux de collaboration et de cocitations, pour regrouper les publications appartenant à un individu spécifique. Outre le nombre relativement faible d'alternatives, le choix de cet algorithme a été motivé par le fait qu'il a été beaucoup utilisé par les chercheurs dans le domaine de la bibliométrie (p. ex. Archambault, Mongeon et Larivière, 2017; Larivière et Costas, 2016; Sugimoto et al., 2017). Il est également implémenté dans le logiciel de visualisation de réseaux VosViewer. De plus, il était beaucoup plus simple d'utiliser les données déjà désambiguïsées plutôt que de traiter l'ensemble de nos données à l'aide d'un autre algorithme.

Comme chacun des inventeurs des PBA est associé à un des auteurs, la désambiguïsation des inventeurs a donc été effectuée en même temps que celle des auteurs. Le Tableau XII (p. 89) présente le nombre d'auteurs et d'inventeurs identifiés dans l'ensemble de notre jeu de données une fois cette désambiguïsation effectuée. La désambiguïsation automatique des auteurs n'est évidemment pas aussi précise qu'une méthode manuelle. Cependant, notre objectif n'étant pas de produire un classement précis des chercheurs selon leur production scientifique, mais bien d'avoir une idée de l'ampleur de leur contribution respective à l'avancement des connaissances, la désambiguïsation produite par l'algorithme nous apparaît suffisante à cette fin.

Tableau XII. Nombre d'auteurs distincts et d'inventeurs distincts par discipline.

Discipline	Nombre d'auteurs distincts	Nombre d'inventeurs distincts
Recherche biomédicale	17 110	6 164
Sciences naturelles	7 574	4 304
Génie	3 859	2 679
Total	28 543	13 147

À partir de ces données de publications désambiguïsées, nous avons donc pu calculer le nombre de citations reçues par les articles auxquels les auteurs ont contribué. Plus spécifiquement à l'aide d'une requête dans la base de données WoS de l'OST, nous calculons pour chaque auteur de chaque PBA le nombre de citations reçues par tous les articles publiés à

³⁷ Les données de publications désambiguïsées ont été fournies directement par CWTS.

une date antérieure ou égale à l'année du dépôt de la demande de brevet. Nous utilisons cette année plutôt que l'année de publication de l'article puisque l'objectif est de déterminer la relation entre la réputation du chercheur et l'attribution du statut d'inventeur, celle-ci ayant lieu au moment du dépôt de la demande de brevet. Une limite de cette mesure est que nous ne pouvons utiliser que les articles de l'auteur qui sont indexés dans le WoS, et que le nombre de citations reçues par ces articles est également limité aux citations provenant d'autres articles indexés dans le WoS.

2.2.2.10. Sexe (QR10)

Nous attribuons un sexe aux auteurs à partir de leur prénom. Tel que mentionné plus tôt, les prénoms des auteurs sont seulement (et partiellement) disponibles dans le WoS à partir de 2008, ce qui rend difficile l'attribution d'un sexe aux auteurs avant cette date. Cette difficulté a été contournée par une série d'étapes. Premièrement, le prénom d'un auteur peut être obtenu de la liste des inventeurs du brevet lorsque l'auteur en question figure également dans la liste des inventeurs. Deuxièmement, comme la mesure de la production et de l'impact scientifiques des auteurs requiert l'identification de l'ensemble des articles de chacun des auteurs, il suffit d'un seul article où le prénom est disponible pour déterminer le sexe de l'auteur en question. Finalement, dans les cas où le prénom d'un auteur demeure inconnu à la suite des deux étapes précédentes, nous avons consulté la liste des auteurs dans la version intégrale des articles afin d'y repérer les prénoms. Le sexe « inconnu » est attribué aux auteurs dont le prénom n'a pu être identifié.

Nous avons ensuite attribué un sexe à chacun des auteurs et inventeurs à partir d'une liste de 130 645 prénoms et leur sexe (75 486 prénoms masculins, 51 566 prénoms féminins). Cette liste de noms a été générée à l'aide de l'algorithme Wiki-Gendersort (Bérubé, Ghiasi et Larivière, en préparation) appliqué à l'ensemble des articles du WoS pour lesquels le prénom des auteurs est disponible. Nous avons utilisé cette liste pour sa facilité d'accès (une table liant les prénoms et le sexe est disponible dans la base de données de l'OST). L'article de Bérubé et al. (en préparation) montre également que l'algorithme qu'ils utilisent performe de façon comparable aux autres méthodes utilisées dans la littérature. En effet, le Tableau XIII (p. 91)

tiré de l'article de Bérubé et al. (en préparation) montre que l'algorithme Wiki-Gendersort attribue correctement le sexe dans plus de 97 % des cas pour l'ensemble des 130 645 prénoms.

Tableau XIII. Évaluation de l'algorithme Wiki-Gendersort à l'aide de quatre bases de données associant un sexe à des prénoms (adapté de Bérubé et al. [2017]).

Base de données	Noms identifiés (%)	Sexe correctement attribué (%)
GenderChecker	90,65	98,39
Gender.c	95,85	98,95
NamSor	89,39	97,07
U.S. Census	97,28	99,91

L'algorithme recherche chacun des prénoms dans Wikipédia (édition anglaise) et génère une liste de pages. Il extrait ensuite les termes du texte d'introduction de chacune de ces pages et attribue un sexe à l'individu présenté dans chaque page en comptant le nombre d'occurrences des déterminants masculins *he* et *his* et des déterminants féminins *she* et *her*. Lorsque le même sexe est attribué à 75 % des individus ayant un même prénom, ce sexe est attribué au prénom. Les prénoms pour lesquels le seuil de 75 % n'est pas atteint sont placés dans la catégorie « inconnu ». Une des limites de cette méthode basée sur les probabilités est qu'il est possible qu'un sexe (féminin ou masculin) soit attribué à un prénom qui est en fait mixte (p. ex. l'algorithme considère le prénom « Claude » comme étant masculin bien qu'il puisse aussi être féminin). Le Tableau XIV (p. 91) présente les résultats du processus d'attribution d'un sexe aux auteurs de notre jeu de données.

Tableau XIV. Nombre et proportion des auteurs par sexe et par discipline.

Discipline	Hommes		Femmes		Inconnu		Total
	N	%	N	%	N	%	N
Recherche biomédicale	11 166	60,0	4 191	22,5	3 255	17,5	18 612
Sciences naturelles	5 481	65,7	1 002	12,0	1 862	22,3	8 345
Génie	2 816	66,3	460	10,8	969	22,8	4 245
Total	19 463	62,4	5 654	18,1	6 092	19,5	31 202

La proportion plutôt élevée (19,5 %) des auteurs dont le sexe n'a pu être identifié s'explique en partie par le fait que les prénoms des chercheurs n'ont pas tous pu être identifiés malgré les efforts que nous avons déployés. De plus, l'algorithme Wiki-Gendersort a de la

difficulté à assigner un sexe à certains prénoms asiatiques, qui sont d'ailleurs très nombreux dans notre jeu de données. Ainsi, nous ne considérons dans cette thèse que les auteurs pour lesquels un sexe a pu être identifié. Nous consignons cette information dans la variable « sexe » qui a deux modalités possibles : « homme » ou « femme ».

2.2.3. Échantillon

Les 6 312 PBA identifiées et les 36 839 auteurs des articles composant ces PBA constituent nos populations de PBA et de chercheurs. Alors que certaines analyses peuvent être effectuées sur l'ensemble de notre jeu de données (QR1 à QR6 et QR9), d'autres n'ont été effectuées que sur des échantillons de convenance (QR7, QR8 et QR10), soit sur les sous-ensembles d'observations pour lesquels certaines informations (p. ex. sexe, contribution, rôle) étaient disponibles. Le Tableau XV (p. 92) présente la taille des échantillons utilisés pour chacune des questions de recherche.

Tableau XV. Taille des échantillons pour chaque question de recherche.

QR	Unité d'analyse	Recherche biomédicale		Sciences naturelles		Génie		Total	
		PBA	Auteurs	PBA	Auteurs	PBA	Auteurs	PBA	Auteurs
1, 2, 3, 4	PBA	3 028	s.o.	1 985	s.o.	1 299	s.o.	6 312	s.o.
5, 6, 9, 10	Auteurs	3 017	15 361	1 948	6 488	1 248	3 278	6 213	25 127
7, 9, 10	Auteurs	61	450	s.o.	s.o.	s.o.	s.o.	61	450
8, 9, 10	Auteurs	112	843	s.o.	s.o.	s.o.	s.o.	112	843

Note : Les questions de sexe (QR9) et de réputation (QR10) se répètent dans la première colonne du tableau, car elles ne sont jamais considérées seules, mais en combinaison avec la position dans la liste des auteurs (QR5) et la discipline (QR6), le rôle (QR7) ou la contribution (QR8).

Comme nous nous intéressons à la valeur prédictive distincte des caractéristiques individuelles des auteurs et de leurs contributions (voir la section suivante portant sur l'analyse des données pour plus de détails), les trois dernières lignes du Tableau XV (p. 92) présentent le nombre de PBA et d'auteurs résultant du chevauchement entre les échantillons utilisés pour plusieurs des questions de recherche. Par exemple, pour mesurer la valeur prédictive distincte du rôle (QR7), de la réputation (QR9) et du sexe (QR10), nous utilisons un sous-ensemble des 709 auteurs pour lesquels le rôle a été identifié, soit les 595 auteurs pour lesquels le sexe a

également pu être identifié. La réputation des auteurs étant identifiée pour l'ensemble des auteurs, son inclusion dans l'analyse n'a pas pour effet de réduire l'échantillon.

2.2.4. En résumé

Les notices bibliographiques des articles du WoS et des brevets du USPTO ont été obtenues directement de l'OST via un serveur SQL. Ces données ont permis de former 6 312 PBA à l'aide d'une méthode basée sur les auteurs-inventeurs et sur le contenu des articles et des brevets. Afin de répondre aux questions de recherche QR1 à QR4 ainsi que QR6, nous avons identifié pour chaque PBA le nombre d'auteurs, la discipline, le fait que l'article est une collaboration interinstitutionnelle ou non, et le facteur d'impact de la revue calculé sur une période de 2 ans. Comme ces informations sont disponibles pour l'ensemble des articles, nous sommes en mesure d'utiliser l'ensemble des PBA dans nos analyses pour ces questions de recherche.

L'algorithme de désambiguïsation des auteurs de Caron et van Eck (2014) a ensuite permis d'identifier 31 204 auteurs et 13 472 inventeurs distincts à partir de ces PBA. Des données sur la contribution respective de chacun de ces auteurs aux articles qu'ils ont publiés ont été recueillies afin de répondre aux questions de recherche QR5, QR7 et QR8. Ces données sont 1) la position dans la liste des auteurs d'un article, 2) le rôle de l'auteur dans l'équipe de recherche tel que déterminé à l'aide des séquences d'auteurs en ordre alphabétique partiel, et 3) la contribution à la recherche telle qu'indiquée dans la déclaration de contribution des auteurs d'un article. Ces deux dernières informations n'étant pas disponibles pour l'ensemble des articles, des échantillons de convenance sont utilisés pour l'analyse de la relation entre ces variables et le statut d'inventeur.

Les données désambiguïsées fournies par l'algorithme de Caron et van Eck (2014) nous ont aussi permis de collecter, pour chaque auteur le nombre de citations reçues (la réputation) en date de dépôt de la demande de brevet (QR9). Le sexe des auteurs (QR10) a quant à lui été déterminé à l'aide de l'algorithme de Bérubé et al. (en préparation). Un sexe (Homme ou Femme) a pu être attribué à 63,5 % des 31 204 auteurs distincts.

L'ensemble de ces opérations de collecte et de traitement de données nous a permis de constituer des échantillons de PBA et d'auteurs, présentés dans le Tableau XV (p. 92), et utilisés

pour répondre à nos questions de recherche. La section suivante présente les méthodes d'analyse de données utilisées.

2.3. Analyse des données

Nous présentons dans les prochaines sections les méthodes utilisées pour analyser les données recueillies dans le but d'apporter des réponses à nos questions de recherche. Le devis corrélationnel prédictif se prête à la formulation d'hypothèses (Fortin et Gagnon, 2016). Nous avons donc formulé un ensemble d'hypothèses de recherche que nous présentons dans la prochaine section. Les deux sections suivantes présentent les méthodes utilisées pour vérifier ces hypothèses. Nous illustrons finalement, dans la dernière section, l'adéquation entre les hypothèses de recherche, les variables prédictives et prédites qui leur sont associées, les sources de données utilisées, et le type d'analyse effectuée.

2.3.1. Hypothèses de recherche

Notre recherche porte sur deux unités d'analyse, soit les PBA elles-mêmes et les auteurs des articles, et nous avons une hypothèse générale ainsi qu'un ensemble d'hypothèses spécifiques pour chacune d'elles. Ces hypothèses sont tirées de notre revue de la littérature sur les pratiques de collaboration et d'attribution des statuts d'auteur et d'inventeur, ainsi que des résultats d'études antérieures sur les PBA que nous avons présentés au Chapitre 1 (section 1.5).

Des études antérieures ont montré que les pratiques de collaboration et d'attribution du statut d'auteurs varient en fonction de la discipline (Knorr Cetina, 1999; Larivière et al., 2016; Pontille, 2004); que l'impact scientifique d'une recherche est lié à un nombre moins important d'auteurs « spécialistes » (Jabbadheri et Walsh, 2017); que plus le nombre d'auteurs est élevé plus les équipes ont tendance à avoir une structure bureaucratique (Walsh et Lee, 2015); et que la collaboration interinstitutionnelle est associée à une augmentation du nombre d'auteurs Jabbadheri et Walsh (2017). Notre hypothèse générale concernant les PBA est donc que le ratio inventeurs/auteurs des PBA est déterminé par un ensemble de facteurs, dont la taille de l'équipe de recherche, l'impact scientifique potentiel de la recherche, et la discipline. De cette hypothèse générale découlent les quatre hypothèses spécifiques suivantes :

HS₁ Le ratio inventeurs/auteurs diffère entre les disciplines.

- HS₂** Plus le facteur d'impact de la revue de publication de l'article formant une PBA est élevé, plus le ratio inventeurs/auteurs sera élevé.
- HS₃** Plus le nombre d'auteurs d'une PBA est élevé, plus le ratio inventeurs/auteurs sera faible.
- HS₄** Les PBA issues d'une collaboration interinstitutionnelle auront un ratio inventeurs/auteurs plus faible que les PBA dont tous les auteurs sont affiliés à la même institution.

Nous avons vu que les premières et dernières positions sont généralement les plus importantes dans les disciplines où la recherche s'effectue en laboratoire (Pontille, 2004, 2006); que les auteurs peuvent être divisés en trois groupes : auteurs principaux, superviseurs et périphériques) (Baerlocher et al., 2007; Mongeon et al., 2017); que le statut d'inventeur est en principe réservé aux individus qui ont participé à la conception de l'invention³⁸; que les contributions des femmes en science ont tendance à être sous-estimées (Rossiter, 1993); que des disparités femmes-hommes existent à tous les niveaux du système de reconnaissance de la science (Larivière et al. 2013); et que le capital va au capital (Bourdieu, 1975; Merton 1968). Nous émettons donc l'hypothèse générale que le fait qu'un auteur obtienne ou non le statut d'inventeur est déterminé à la fois par sa contribution, son rôle, sa réputation et son sexe. De cette hypothèse générale découlent sept hypothèses spécifiques additionnelles.

- HS₅** Plus un auteur est près de la première ou dernière position dans la liste des auteurs de l'article, plus la probabilité qu'il soit inventeur est grande.
- HS₆** La relation entre la position dans la liste des auteurs et l'obtention du statut d'inventeur diffère en fonction de la discipline.
- HS₇** Les auteurs avec un rôle de supervision ont une probabilité plus grande d'être inventeurs que les auteurs principaux et les auteurs périphériques.
- HS₈** Les auteurs principaux ont une probabilité plus grande d'être inventeurs que les auteurs périphériques.
- HS₉** Les auteurs ayant contribué à la conception de l'étude ont une probabilité plus grande d'être inventeurs.
- HS₁₀** Les hommes ont une probabilité plus grande que les femmes d'être inventeurs.

³⁸ *Hardee, In re*, 223 USPQ 1122 (Comm'r Pat 1984).

HS₁₁ Plus le nombre de citations d'un auteur est élevé, plus la probabilité qu'il soit inventeur est grande.

Le devis corrélatif prédictif prescrit l'utilisation de la régression (Fortin et Gagnon, 2016; Schoumaker, 2013) pour vérifier nos hypothèses de recherche (c.-à-d. déterminer dans quelle mesure les variables prédictives permettent de prédire les variables prédites). Les deux sections suivantes présentent plus en détail les méthodes utilisées pour l'analyse des données afin de vérifier les deux groupes d'hypothèses (HS₁ à HS₄, et HS₅ à HS₁₁). Les analyses sont effectuées à l'aide du logiciel SPSS d'IBM, version 24.

2.3.2. Prédiction du ratio inventeurs/auteurs (HS₁ à HS₄)

Cette partie de l'analyse est effectuée sur l'ensemble des 6 312 PBA et consiste à déterminer dans quelle mesure le ratio inventeurs/auteurs d'une PBA peut être prédit par un ensemble de variables prédictives, soit la discipline, le facteur d'impact de la revue de publication, le nombre d'auteurs de l'article, et la collaboration interinstitutionnelle. Le test statistique choisi pour vérifier les hypothèses HS₁ à HS₄ est la régression linéaire multiple, qui permet d'étudier les relations entre plusieurs variables prédictives et une variable prédite, et qui fournit une équation permettant de prédire la valeur de la variable prédite à partir des valeurs des variables prédictives (Fortin et Gagnon, 2016). L'équation de la régression linéaire multiple prend la forme suivante :

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_h X_h + \varepsilon$$

Où Y est la valeur de la variable prédite,

X_h est la valeur d'une variable prédictive,

β_h est le coefficient de corrélation partiel qui mesure la variation d'Y pour une variation d'une unité de X_h et lorsque toutes les autres variables prédictives sont maintenues constantes,

α est la valeur d'Y lorsque toutes les variables prédictives ont la valeur 0,

et ε est le résidu (c.-à-d. la différence entre la valeur de Y prédite par le modèle et la valeur réelle observée). (Schoumaker, 2013)

En d'autres termes, la régression linéaire multiple permet de déterminer l'existence et l'ampleur d'une relation statistiquement significative entre plusieurs variables prédictives et une variable prédite tout en contrôlant l'effet d'autres variables (Schoumaker, 2013). La capacité prédictive d'un modèle de régression (c.-à-d. de l'ensemble des variables considérées) dans son ensemble est évaluée à l'aide d'un coefficient de corrélation multiple au carré (R^2) qui indique à quel point les variations observées dans la variable prédite peuvent être prédites par les variables prédictives (Fortin et Gagnon, 2016).

La régression linéaire multiple est une méthode d'analyse appropriée dans notre cas étant donné 1) que nous avons une seule variable prédite (le ratio inventeurs/auteurs), 2) qu'il s'agit d'une variable continue, 3) que nous considérons plusieurs variables prédictives (Sproull, 1988). Selon Sproull (1988) les variables prédictives d'un modèle de régression multiple doivent toutes être continues. Or, deux de nos variables prédictives sont nominales (la collaboration interinstitutionnelle et la discipline). Ces variables peuvent être incluses dans le modèle à condition de les transformer en variables dichotomiques pouvant prendre la valeur 0 ou 1 (Schoumaker, 2013). La variable discipline sera donc transformée en trois variables (recherche biomédicale, sciences naturelles, et génie) prenant la valeur 1 si la PBA est classée dans cette discipline ou 0 sinon. La variable « multi_institution » étant déjà dichotomique, elle n'a pas à être transformée.

Dans un modèle de régression, on distingue généralement les variables prédictives des variables de contrôle (aussi appelées variables de confusion ou variables intermédiaires) (Schoumaker, 2013). Ces variables sont introduites et agissent de la même façon que les variables prédictives dans le modèle, la distinction dépendant uniquement des hypothèses formulées (Schoumaker, 2013). Dans notre cas, l'année de publication est utilisée comme variable de contrôle (voir Figure 1, p. 98) puisqu'il s'agit d'une variable pouvant exercer une influence sur les autres variables du modèle, mais qui ne fait pas l'objet d'une de nos questions ou hypothèses de recherche. Elle apparaîtra donc au côté des variables prédictives (gr_disc, nb_auteurs, fi_2ans, et multi_institution) dans la table des coefficients de la régression, et elle fera partie de l'équation servant à prédire le ratio inventeurs/auteurs, mais ne fera pas l'objet d'une discussion au même titre que les variables prédictives.

La Figure 1 (p. 98) présente un schéma du modèle de régression linéaire utilisé. Les noms abrégés des variables (entre parenthèses dans la figure) sont ceux qui sont utilisés dans la base de données que nous utilisons pour notre recherche (voir l'Annexe 5 [p. xxvi] pour l'ensemble des variables).



Figure 1. Relation entre les variables du modèle visant à prédire le ratio inventeurs/auteurs des PBA (HS₁ à HS₄).

Au lieu d'utiliser la régression standard, où toutes les variables sont simultanément entrées dans le modèle, nous avons choisi d'utiliser la régression hiérarchique (aussi appelée régression pas à pas, ou régression séquentielle), où les variables sont entrées une à une ou par groupe (Tabachnick et Fidell, 2013). Ceci permet de déterminer si l'ajout de variables supplémentaires améliore significativement la capacité prédictive (R^2) du modèle, et ainsi d'identifier le modèle le plus simple, conformément au principe de parcimonie voulant que l'on évite de surcharger un modèle de régression de variables dont l'apport au coefficient R^2 est minime (Schoumaker, 2013). Le choix de l'ordre d'inclusion des variables est fondé sur la théorie et la littérature antérieure, les variables prédictives étant jugées par le chercheur comme étant le plus fortement liées à la variable prédite étant entrées en premier dans le modèle. Cette opération ne change pas la capacité prédictive (R^2) du modèle, mais permet d'observer l'apport distinctif de l'ajout de chaque variable ou groupe de variable.

Dans le chapitre 3, les modèles de régression linéaire sont accompagnés de tableaux et figures descriptifs visant à dresser un portrait plus visuel des échantillons utilisés et des relations entre les variables.

2.3.3. Prédiction du statut d'inventeur (HS₅ à HS₁₁)

Cette partie de l'analyse vise à déterminer dans quelle mesure l'attribution (ou non) du statut d'inventeur aux auteurs des PBA peut être prédite par un ensemble de variables prédictives liées à leur contribution et/ou à leur rôle dans l'équipe de recherche (position dans la liste des auteurs, rôle et contribution), à leur réputation, et à leur sexe. Le test statistique utilisé est la régression logistique, puisque la variable prédite (le statut d'inventeur) est une variable dichotomique (la régression linéaire nécessite que la variable prédite soit une variable continue). La régression logistique est très similaire à la régression linéaire. La principale différence, outre la variable prédite dichotomique, est qu'au lieu de chercher à prédire la valeur de la variable prédite, l'objectif est de prédire la probabilité qu'une observation (dans notre cas un auteur) soit dans l'une des deux catégories de la variable prédite (dans notre cas, le statut d'inventeur) à l'aide d'une ou de plusieurs variables prédictives (Fortin et Gagnon, 2016). La formule de la régression logistique, qui ressemble beaucoup à celle de la régression linéaire multiple, est la suivante :

$$\ln \left[\frac{p}{(1-p)} \right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Nous avons donc à la droite de la formule la plupart des mêmes éléments que dans la régression linéaire, soit une constante α , des coefficients de corrélation partiels β_h , et des variables prédictives X_h . Cependant, au lieu de calculer une valeur pour la variable prédite, la constante, les coefficients, et les variables prédictives permettent de calculer le logit de la probabilité qu'un auteur soit inventeur (c.-à-d. la probabilité que la variable prédite ait la valeur 1), et non cette probabilité directement. L'expression $\frac{p}{(1-p)}$ représente le rapport de cotes (*odds ratio*), qui est donc, dans notre cas, la probabilité qu'un auteur soit inventeur (p) divisé par la probabilité qu'il ne le soit pas ($1-p$). Par exemple, si $p = 0,75$, le rapport de cote sera de 3, ce qui signifie qu'un auteur a 3 fois plus de chance d'être inventeur que de ne pas l'être (sa cote est de 3 contre 1). Le Tableau XVI (p. 100), tiré de Schoumaker (2013, p. 262), permet de mieux comprendre comment interpréter la probabilité, le rapport de cotes et le logit (p).

Tableau XVI. Exemple de valeurs de probabilités, rapport de cotes et logit (p).

p	$1-p$	rapport de cotes $\left(\frac{p}{(1-p)}\right)$	logit (p) $\left(\ln\left[\frac{p}{(1-p)}\right]\right)$
0,00	1,00	0,00	$-\infty$
0,10	0,90	0,11	-0,95
0,25	0,75	0,33	-0,48
0,50	0,50	1,00	0,00
0,75	0,25	3,00	0,48
0,90	0,10	9,00	0,95
1,00	0,00	$+\infty$	$+\infty$

Le logit (p) permet donc de transformer la variable prédite en une valeur continue, c'est pourquoi la procédure utilisée dans la régression logistique est la même que dans la régression linéaire (seule l'interprétation des résultats diffère).

L'évaluation globale d'un modèle de régression logistique peut s'effectuer à l'aide du R^2 (comme dans la régression linéaire), mais on l'appelle plutôt pseudo- R^2 dans ce cas-ci parce qu'il est conceptuellement différent du R^2 du modèle linéaire. En effet, le pseudo- R^2 n'est pas une mesure de la variance expliquée par le modèle, mais une comparaison du modèle complet à la constante α (Schoumaker, 2013). Une autre méthode pour évaluer la capacité prédictive globale d'un modèle de régression logistique est la proportion des cas classés correctement, qui s'obtient en comparant la classification réelle des cas à la classification effectuée par le modèle. Plus la proportion des cas correctement classés est élevée, plus le modèle est bon (Schoumaker, 2013). Nous avons choisi d'utiliser cette deuxième mesure dans notre étude parce qu'elle est tout aussi valide, mais plus simple à interpréter que le pseudo- R^2 .

Puisque les informations sur le rôle et sur la contribution à la conception de l'étude ne sont disponibles que pour des sous-ensembles de données, nous utilisons trois échantillons et modèles de régression logistique distincts pour tester nos hypothèses de recherche HS₅ à HS₁₁. Le Tableau XVII (p. 101) présente les variables incluses dans chacun des modèles de régression logistique utilisés. La variable « position » fait l'objet de l'hypothèse HS₅ et est ainsi analysée dans le premier modèle. Elle est utilisée comme variable de contrôle dans les deux autres modèles qui visent plutôt à déterminer la relation entre le rôle (modèle 2) et la contribution (modèle 3) en contrôlant pour le rôle.

Tableau XVII. Hypothèses vérifiées, variables prédites et variables prédictives pour chacun des modèles globaux de régression logistique.

Modèle	Hypothèses vérifiées	Variables prédites	Variables prédictives	Variables de contrôle
1	HS ₅ , HS ₆ , HS ₁₀ et HS ₁₁	Inventeur	position sexe nb_citations gr_disc	s.o.
2	HS ₇ , HS ₈ , HS ₁₀ et HS ₁₁	Inventeur	rôle homme nb_citations	position
3	HS ₉ , HS ₁₀ et HS ₁₁	Inventeur	analyse conception matériel réalisation rédaction autres sexe nb_citations	position

La Figure 2 (p. 101), la Figure 3 (p. 102) et la Figure 4 (p. 102) fournissent des représentations visuelles de ces trois modèles de régression logistique.

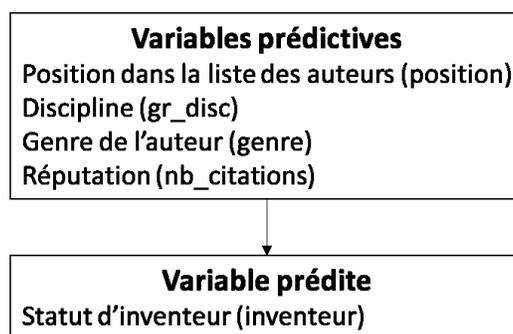


Figure 2. Modèle de régression logistique visant à prédire l'obtention du statut d'inventeur par les auteurs des PBA à partir de la position dans la liste des auteurs, de la discipline, du sexe et de la réputation (HS₅, HS₆, HS₁₀ et HS₁₁).

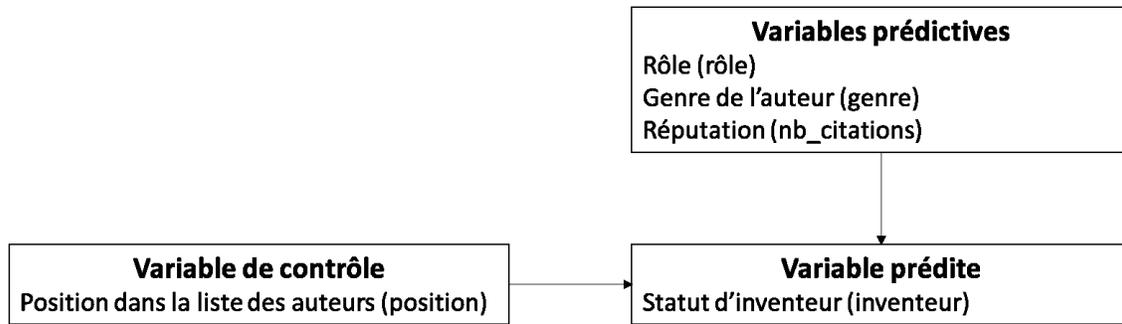


Figure 3. Modèle de régression logistique visant à prédire l'obtention du statut d'inventeur par les auteurs des PBA à partir du rôle, du sexe et de la réputation (HS₇, HS₈, HS₁₀ et HS₁₁).

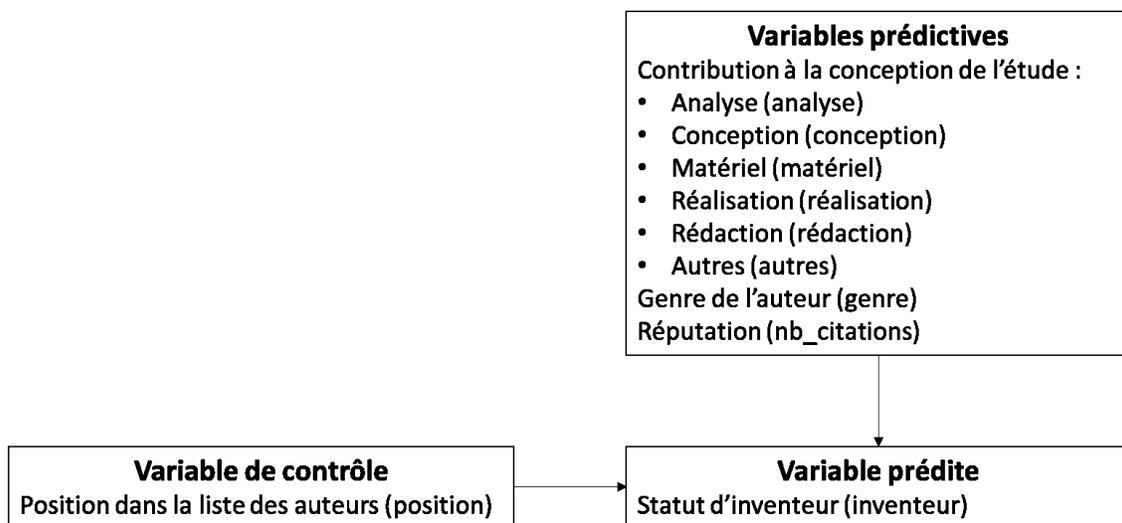


Figure 4. Modèle de régression logistique visant à prédire l'obtention du statut d'inventeur les auteurs des PBA à partir de la contribution, du sexe et de la réputation (HS₉, HS₁₀ et HS₁₁).

Comme dans l'analyse précédente visant à valider les hypothèses de recherche HS₁ à HS₄, nous utilisons la régression logistique hiérarchique afin de voir si l'ajout de chacune des variables permet d'améliorer significativement la capacité prédictive du modèle. Les modèles de régression logistique sont aussi accompagnés de figures illustrant la relation entre les variables et la proportion d'auteurs ayant obtenu le statut d'inventeur.

2.3.4. Liens entre les hypothèses de recherche, les variables, les sources de données, et les types d'analyse

Le Tableau XVIII (p. 103) présente les liens entre les hypothèses de recherche, les variables prédictives, contrôles et prédites, et la source des données utilisées pour les vérifier.

Tableau XVIII. Liens entre les hypothèses de recherche, les sources de données et le type d'analyse.

Hypothèse	Variable prédite	Variables prédictives ou de contrôle	Source de données	Type d'analyse
HS ₁ à HS ₄	ratio inventeurs/auteurs	fi_2ans	PatFT, WoS	Régression linéaire multiple
		gr_disc	PatFT, OST	
		nb_auteurs	PatFT, OST	
		multi_institution	PatFT, OST	
		année_pub	OST	
HS ₅ à HS ₁₁	inventeur	position	PatFT, WoS	Régression logistique
		rôle	PatFT, WoS, chercheur principal (ordre alphabétique partiel)	
		analyse conception matériel réalisation rédaction autres	PatFT, WoS, Assistantes de recherche (déclarations des contributions)	
		sexe	OST*	
		nb_citations	OST/CWTS	
		gr_disc	PatFT, OST	

* Les données utilisées pour l'attribution du sexe ont été générées par l'algorithme de Bérubé et al. (en préparation) et ont été obtenues à partir de la base de données de l'OST.

2.3.5. En résumé

Dans cette section, nous avons décrit les tests statistiques utilisés pour analyser nos données et tester nos hypothèses de recherche. Pour les hypothèses HS₁ à HS₄, nous utilisons un modèle de régression linéaire où les variables prédictives sont l'impact scientifique potentiel de la recherche, la discipline, le nombre d'auteurs et le fait que l'article soit ou non issu d'une collaboration interinstitutionnelle. Le modèle permet de déterminer dans quelle mesure ces

variables permettent de prédire le ratio inventeurs/auteurs des PBA. Puisque la variable prédite (le statut d'inventeur) pour les hypothèses HS₅ à HS₁₁ est dichotomique, nous utilisons la régression logistique pour les tester. Les variables prédictives sont dans ce cas-ci liées à la contribution et au rôle du chercheur (position, rôle et contribution) et aux caractéristiques individuelles des auteurs (nb_citations et sexe). Dans les deux cas, nous utilisons la régression de type hiérarchique qui permet d'insérer les variables une à une dans le modèle et de déterminer l'impact de leur inclusion dans le modèle sur la capacité prédictive de ce dernier.

2.4. Qualité de la recherche

La qualité d'une recherche quantitative s'examine en fonction d'un ensemble de critères. Deux de ces critères font consensus entre les auteurs : la validité interne et la validité externe (Fortin et Gagnon, 2016; Guba et Lincoln, 1989; Pickard, 2007; Shadish, Cook et Campbell, 2002; Silipigni Connaway et Radford, 2017). À ces deux critères, Guba et Lincoln (1989) et Pickard (2007) ajoutent la fidélité et l'objectivité, alors que de leur côté, Cook et Campbell (1979), Fortin et Gagnon (2016) et Silipigni Connaway et Radford (2017) ajoutent la validité de conclusion statistique et la validité de construit. Dans les sections suivantes, nous présentons brièvement chacun de ces critères, ainsi que les moyens utilisés pour en assurer le respect.

2.4.1. Validité de conclusion statistique

La validité de conclusion statistique est le « degré de certitude des conclusions tirées des tests statistiques sur les relations entre les variables ou sur les différences entre les groupes » (Fortin et Gagnon, 2016, p. 175). Shadish et al. (2002) identifient deux principales menaces à la validité de conclusion statistique. La première est la faible puissance statistique des tests effectués. La puissance statistique visée est de 95 % et elle est maximisée en effectuant la plupart de nos tests statistiques sur un grand nombre d'observations, soit 6 312 PBA pour QR1 à QR4 et 25 127, 450, et 843 auteurs pour QR6, QR7 et QR8, respectivement. La deuxième menace est le non-respect des postulats liés aux tests statistiques utilisés. Les postulats de la régression linéaire et de la régression logistique sont les suivants (Fox, 2016; Gelman et Hill, 2007; Tabachnick et Fidell, 2013) :

1. La relation entre chacune des variables prédictives continues et la variable prédite (ou le logit de la variable prédite dans le cas de la régression logistique) doit être linéaire.
2. Les observations doivent être indépendantes les unes des autres (indépendance des erreurs).
3. Il ne doit pas y avoir de valeurs extrêmes pouvant affecter la capacité prédictive du modèle.
4. Les variables prédictives ne doivent pas être fortement corrélées entre elles (absence de multicollinéarité).

Les trois postulats suivants s'appliquent seulement à la régression linéaire :

5. Les valeurs résiduelles sont normalement distribuées pour chacune des valeurs de la variable prédite (normalité des erreurs).
6. La relation entre les valeurs résiduelles et la variable prédite doit être linéaire (linéarité des erreurs).
7. La variance des valeurs résiduelles doit être similaire pour chacune des valeurs de la variable prédite (homocédasticité).

Pour assurer la validité de conclusion statistique de nos analyses, nous procéderons à une vérification du respect de chacun de ces postulats pour chacun de nos modèles de régression.

2.4.2. Validité interne

La validité interne désigne la véracité des résultats obtenus, leur adéquation avec la réalité. Notre devis étant corrélationnel prédictif, il s'agit de s'assurer que les effets observés sur les variables prédites (le ratio inventeurs/auteurs et le statut d'inventeur) sont causés par les variables prédictives du modèle (p. ex. le sexe, le statut, la contribution) et non par d'autres variables non prises en compte dans l'étude (Fortin et Gagnon, 2016; Guba et Lincoln, 1989).

La validité interne de notre recherche est assurée de deux façons. D'abord, nous avons utilisé les informations à notre disposition pour identifier les caractéristiques des chercheurs (réputation et sexe) ainsi que la nature et l'ampleur des contributions respectives des auteurs de chacun des articles. Cela permet de contrôler un maximum de facteurs lors des analyses statistiques. De plus, nous assurons la validité interne de notre recherche en utilisant les tests statistiques appropriés et en mettant l'accent sur l'ampleur des effets observés (et pas seulement sur la significativité statistique des résultats) dans la présentation des résultats. Des précautions

sont en effet de mise dans l'interprétation des relations observées entre les variables prédictives et prédites. Les différences entre les groupes peuvent être statistiquement significatives sans pour autant être assez importantes pour ressenties ou remarquées en pratique.

Selon Creswell (2009), une bonne connaissance du phénomène étudié et des variables utilisées permet d'augmenter la validité interne d'une recherche. En ce sens, la revue de la littérature présentée au Chapitre 1 et le fait que l'auteur de la présente recherche ait antérieurement participé à des projets de recherche sur les pratiques d'attribution du statut d'auteur et sur la division du travail au sein des équipes de recherche (Larivière et al., 2016; Mongeon et al., 2017; Paul-Hus et al., 2017) permettent d'améliorer la validité interne de notre recherche.

2.4.3. Validité de construit

La validité de construit est l'adéquation entre les variables prédictives et prédites et les construits théoriques qu'elles mesurent (Fortin et Gagnon, 2016). Parmi les nombreuses menaces à la validité de construit décrites par Shadish et al. (2002), les suivantes s'appliquent à notre recherche :

1. La description inadéquate des construits.
2. L'utilisation d'une seule mesure pour opérationnaliser les construits.
3. L'utilisation d'une seule méthode.

Afin d'assurer la validité de construit de notre étude, nous avons présenté au Chapitre 1 une revue exhaustive de la littérature portant sur les concepts et les phénomènes qui nous intéressent, et nous avons clairement expliqué et justifié leur opérationnalisation. Nous utilisons également plusieurs mesures pour opérationnaliser la notion de contribution, soit la position dans la liste des auteurs (position), le fait d'être à l'intérieur ou à l'extérieur d'un sous-ensemble d'auteurs listés en ordre alphabétique (rôle), et la contribution indiquée dans les déclarations de contributions des auteurs (contribution). L'utilisation d'une seule méthode (la bibliométrie) demeure une limite de notre étude. En effet, l'étude des pratiques d'attribution des statuts d'auteur et d'inventeur à l'aide d'autres méthodes quantitatives (p. ex. enquête) ou qualitatives (p. ex. étude de cas) pourrait éventuellement confirmer ou nuancer les résultats de notre étude.

2.4.4. Validité externe

La validité externe fait référence à la possibilité de généraliser les résultats de l'analyse à l'ensemble de la population et à d'autres contextes ou périodes que ceux pris en compte dans l'étude (Fortin et Gagnon, 2016; Guba et Lincoln, 1989; Shadish et al., 2002). Un des obstacles à la validité externe est la généralisation des résultats à différentes périodes sans tenir compte des différences entre ces périodes pouvant avoir un effet sur les résultats (Fortin et Gagnon, 2016; Shadish et al., 2002). Nous neutralisons cette menace, d'une part, en analysant des données couvrant une période de vingt-cinq ans et, d'autre part, en introduisant l'année du dépôt du brevet comme variable de contrôle dans nos modèles de régression. D'autres menaces potentielles découlent de l'échantillonnage (la sélection des cas étudiés) et du choix du milieu étudié. Le fait que nous effectuons nos analyses sur un grand nombre d'observations (l'ensemble de la population accessible, dans la plupart des cas), contribue à neutraliser ces menaces et à maximiser la validité externe de notre étude.

2.4.5. Fidélité

La fidélité de la recherche fait référence au fait que d'autres chercheurs puissent la reproduire et obtenir les mêmes résultats (Guba et Lincoln, 1989; Pickard, 2007). Notre recherche est en grande partie peu vulnérable aux facteurs pouvant en réduire la fidélité (p. ex. le codage par des humains) puisqu'une grande partie des données est tirée directement des notices bibliographiques des articles et des brevets. Cependant, l'identification des contributions respectives des auteurs des articles a nécessité un codage manuel. Ce codage a été vérifié au moyen d'un test intercodeur dont le résultat (taux d'accord de 89,7 %) permet d'avoir confiance en sa validité. En effet, un taux d'accord au-delà de 80 % est généralement jugé satisfaisant (Fortin et Gagnon, 2016). De plus, nous avons vérifié que le critère de fidélité était respecté en documentant en détail la méthode d'appariement des articles et des brevets, le processus de validation (manuel en partie) des PBA ainsi que les méthodes de collecte des autres données afin que d'autres chercheurs puissent être en mesure de fidèlement reproduire notre recherche.

2.4.6. Objectivité

Le critère de l'objectivité fait référence au fait que la recherche n'est pas influencée par les biais, valeurs, motivations, intérêts ou préjugés du chercheur (Guba et Lincoln, 1989; Pickard, 2007). L'objectivité de notre recherche a été assurée par l'utilisation d'un devis méthodologique quantitatif se prêtant peu aux interprétations subjectives ou à l'influence du chercheur sur les mesures effectuées. Il demeure néanmoins une influence possible du chercheur dans le choix des mesures. Cette source de biais potentiel est contrôlée par la revue de la littérature ainsi que la justification et la transparence des choix effectués. La documentation détaillée de notre recherche et son évaluation par des collègues et pairs contribuent également à en assurer l'objectivité. De plus, certains des moyens utilisés pour assurer la validité de construit, notamment la revue exhaustive de la littérature, et la fidélité de l'étude aident aussi à en assurer l'objectivité.

Conclusion

Dans ce chapitre, nous avons détaillé l'ensemble de nos choix et processus méthodologiques visant à apporter des réponses à nos questions de recherche. Notre recherche utilise les méthodes bibliométriques et a donc utilisé comme principale source de données les notices bibliographiques des articles du WoS et des brevets de la PatFT du USPTO fournies par l'OST. Nous avons décrit les opérations de traitement des données effectuées afin de former les PBA, ainsi que pour recueillir l'ensemble des variables requises pour notre analyse.

Notre devis méthodologique étant de type corrélationnel prédictif, nous avons formulé un ensemble de onze hypothèses de recherche et avons décrit les tests statistiques utilisés pour les vérifier, soit la régression linéaire et la régression logistique. Nous avons ensuite présenté les liens entre les hypothèses de recherche, les variables prédictives et prédites, la source de données utilisées pour obtenir ces variables, et les tests statistiques utilisés pour vérifier nos hypothèses.

Dans la dernière section, nous avons ensuite présenté un ensemble de critères servant à évaluer la qualité de la recherche quantitative. Les principales menaces à la validité de nos choix

et outils méthodologiques ont été exposées, ainsi que les moyens utilisés pour neutraliser ces menaces autant que possible.

Chapitre 3. Résultats

Introduction

La présentation des résultats est divisée en deux sections. La première porte sur les prédicteurs du ratio inventeurs/auteurs des PBA. Elle couvre donc les deux premiers objectifs de la thèse et apporte des réponses aux questions de recherche QR1 à QR4 et aux hypothèses de recherche HS₁ à HS₄. La seconde section est subdivisée en trois sous-sections et présente la relation entre la probabilité qu'un auteur figure parmi les inventeurs du brevet et la position dans la liste des auteurs, le rôle, la contribution, le sexe et la réputation. Elle couvre donc les objectifs 3 et 4, les questions de recherche QR5 à QR10 et les hypothèses de recherche HS₅ à HS₁₁. Chaque section débute par un rappel des objectifs et des questions de recherche. Dans chaque section, nous présentons des résultats descriptifs, les résultats des tests statistiques effectués, et validons les hypothèses associées à cette partie de l'analyse.

3.1. Prédicteurs du ratio inventeurs/auteurs

Nous nous intéressons dans cette première section à la relation entre les listes d'auteurs et d'inventeurs des articles et des brevets formant les PBA. Plus particulièrement, nous cherchons à illustrer la relation entre un ensemble de caractéristiques des PBA et le ratio inventeurs/auteurs. Voici en rappel les objectifs et questions de recherche auxquels cette analyse vise à répondre.

Objectif 1. Déterminer la relation entre les caractéristiques de la recherche et le nombre d'auteurs et d'inventeurs.

- QR1. Quelle est la relation entre la discipline de recherche et le ratio inventeurs/auteurs des PBA?
- QR2. Quelle est la relation entre l'impact scientifique potentiel de la recherche et le ratio inventeurs/auteurs des PBA?

Objectif 2. Déterminer la relation entre les caractéristiques de l'équipe de recherche et le nombre d'auteurs et d'inventeurs.

QR3. Quelle est la relation entre la taille de l'équipe de recherche et le ratio inventeurs/auteurs des PBA?

QR4. Quelle est la relation entre la collaboration interinstitutionnelle et le ratio inventeurs/auteurs des PBA?

3.1.1. Le ratio inventeurs/auteurs en fonction du nombre d'auteurs, de la collaboration interinstitutionnelle, du facteur d'impact de la revue de publication et de la discipline

Dans cette section nous analysons la relation entre le ratio inventeurs/auteurs des PBA, les caractéristiques de l'équipe de recherche (nombre d'auteurs et nombre d'institutions) et les caractéristiques de la recherche (discipline, facteur d'impact de la revue de publication). Nous présentons d'abord quelques résultats descriptifs puis nous utilisons la régression linéaire pour déterminer l'ampleur et la significativité statistique de chacun de ces prédicteurs, lorsque tous les autres prédicteurs sont constants et en contrôlant également pour l'année de publication de l'article.

3.1.1.1. Résultats descriptifs

Le Tableau XIX (p. 112) présente les statistiques descriptives sur le nombre d'auteurs des PBA en fonction de la discipline, de la collaboration interinstitutionnelle et du facteur d'impact de la revue dans laquelle l'article a été publié. Le facteur d'impact d'une revue doit être parmi les 10 % les plus élevés de la discipline pour qu'elle soit considérée comme une revue à haut facteur d'impact (haut FI). Nous observons que dans chaque discipline le nombre moyen d'auteurs des PBA est le plus faible pour les PBA sans collaboration interinstitutionnelle dont l'article est publié dans une revue à bas facteur d'impact, et le plus élevé pour les PBA en collaboration interinstitutionnelle dont l'article est publié dans une revue à haut FI. De plus, les distributions ont en commun une asymétrie positive, indiquant une distribution décalée à gauche de la médiane et une queue de distribution s'étalant à la droite. Dans l'ensemble, les coefficients d'asymétrie sont au-delà du seuil de normalité de ± 2 (Field, 2009), c'est pourquoi nous avons

inclus la médiane dans nos statistiques descriptives³⁹. Les coefficients d'asymétrie varient aussi grandement d'un groupe de PBA à l'autre [0,836-6,571] et sont généralement plus faibles en génie. Le nombre d'auteurs médian par article diffère cependant d'une discipline à l'autre. Il est en effet le moins élevé en génie [2,84-4,46] et le plus élevé en recherche biomédicale [4,32-9,35], les sciences naturelles se situant entre les deux [3,41-5,67].

Tableau XIX. Statistiques descriptives sur le nombre d'auteurs des PBA par discipline en fonction de la collaboration interinstitutionnelle et du facteur d'impact de la revue de publication.

Discipline	Groupe	PBA		Nombre d'auteurs				
		N	Min	Max	Méd.	Moy.	Écart type	Asym.
Génie	Uni-institution - bas FI	634	2	15	3	2,84	1,142	3,051
	Uni-institution - haut FI	58	2	8	3	3,57	1,365	0,836
	Multi-institution - bas FI	536	2	13	3	3,58	1,543	1,750
	Multi-institution - haut FI	71	2	11	4	4,46	2,076	0,998
	<i>Ensemble</i>	1 299	2	15	3	3,27	1,465	2,071
Recherche biomédicale	Uni-institution - bas FI	820	2	27	4	4,32	2,820	3,278
	Uni-institution - haut FI	48	2	12	4	4,58	2,277	1,162
	Multi-institution - bas FI	1 906	2	42	6	6,54	3,752	2,169
	Multi-institution - haut FI	254	2	44	8	9,35	5,744	2,247
	<i>Ensemble</i>	3 028	2	44	5	6,15	3,969	2,455
Sciences naturelles	Uni-institution - bas FI	935	2	17	3	3,41	1,827	2,643
	Uni-institution - haut FI	66	2	14	3	3,61	2,411	2,750
	Multi-institution - bas FI	852	2	26	4	4,89	2,408	1,889
	Multi-institution - haut FI	132	2	44	5	5,67	4,095	6,571
	<i>Ensemble</i>	1 985	2	44	4	4,20	2,451	3,771

Le Tableau XX (p. 113) présente les statistiques descriptives sur le nombre d'inventeurs des PBA en fonction de la discipline, de la collaboration interinstitutionnelle et du facteur d'impact de la revue dans laquelle l'article a été publié. Nous observons que, comparativement au nombre d'auteurs présenté dans le Tableau XIX (p. 112), le nombre d'inventeurs moyen et médian par brevet est beaucoup plus stable entre chacun des groupes et entre les disciplines. En

³⁹ Nous utilisons des moyennes dans le texte et les figures de la section 3.1.1. Nous avons fait ce choix car 1) la moyenne et la médiane du nombre d'auteurs sont très proches et l'usage de l'une ou de l'autre n'affecte pas les tendances observées; 2) seules les distributions du nombre d'auteurs sont asymétriques (les distributions du nombre d'inventeurs du Tableau XX (p. 112) et du ratio inventeurs/auteurs des PBA du Tableau XXI (p. 113) sont approximativement normales), et 3) la régression linéaire est un modèle basé sur la moyenne. Nous utilisons d'ailleurs le logarithme du nombre d'auteurs dans les modèles de régression de la section 3.1.1.2.

effet, alors que le nombre moyen d'auteurs se situe entre 2,84 et 9,35 pour l'ensemble des disciplines et groupes, le nombre moyen d'inventeurs se situe quant à lui entre 2,19 et 3,25. Comme le nombre d'auteurs, le nombre d'inventeurs a tendance à être plus élevé pour les PBA en collaboration interinstitutionnelle ainsi que pour les PBA dont l'article est publié dans une revue à haut FI. Les distributions du nombre d'inventeurs des PBA sont aussi moins asymétriques que les distributions du nombre d'auteurs. En effet, les coefficients d'asymétrie du Tableau XX (p. 113) se situent entre 0,375 et 2,056, alors qu'ils se situent entre 0,836 et 6,571 dans le Tableau XIX (p. 112). En fait, seul le coefficient d'asymétrie du groupe « Uni-institution - bas FI » (2,056) est au-delà du seuil de normalité de ± 2 (Field, 2009).

Tableau XX. Statistiques descriptives sur le nombre d'inventeurs des PBA par discipline en fonction de la collaboration interinstitutionnelle et du facteur d'impact de la revue de publication.

Discipline	Groupe	PBA		Nombre d'inventeurs				
		N	Min	Max	Méd.	Moy.	Écart type	Asym.
Génie	Uni-institution - bas FI	634	1	6	2	2,21	0,911	0,961
	Uni-institution - haut FI	58	1	5	2	2,57	1,126	0,625
	Multi-institution - bas FI	536	1	6	2	2,54	1,055	0,611
	Multi-institution - haut FI	71	1	8	3	2,92	1,350	1,126
	<i>Ensemble</i>	1299	1	8	2	2,40	1,028	0,897
Recherche biomédicale	Uni-institution - bas FI	820	1	8	2	2,19	1,126	1,517
	Uni-institution - haut FI	48	1	6	2	2,60	1,300	0,852
	Multi-institution - bas FI	1906	1	10	2	2,39	1,269	1,341
	Multi-institution - haut FI	254	1	8	2	2,69	1,395	1,001
	<i>Ensemble</i>	3028	1	10	2	2,36	1,25	1,350
Sciences naturelles	Uni-institution - bas FI	935	1	13	2	2,37	1,198	2,056
	Uni-institution - haut FI	66	1	4	2	2,42	0,929	0,347
	Multi-institution - bas FI	852	1	10	2	2,59	1,316	1,296
	Multi-institution - haut FI	132	1	7	3	3,25	1,356	0,375
	<i>Ensemble</i>	1985	1	13	2	2,52	1,272	1,512

Le Tableau XXI (p. 114) présente les statistiques descriptives sur le ratio inventeurs/auteurs des PBA en fonction de la discipline, de la collaboration interinstitutionnelle et du facteur d'impact de la revue dans laquelle l'article a été publié. Une seule tendance semble ici commune à chacune des trois disciplines : pour un même niveau de facteur d'impact, le ratio inventeurs/auteurs moyen et médian est moins élevé pour les PBA issues de collaborations

interinstitutionnelles. Cette tendance est cependant moins marquée en génie qu'en recherche biomédicale et en sciences naturelles. Les coefficients d'asymétrie se situant entre -0,980 et 1, les distributions sont approximativement normales (Field, 2009). Dans l'ensemble, les ratios inventeurs/auteurs moyens et médians sont les plus faibles en recherche biomédicale et les plus élevés en génie. Cela est peu surprenant étant donné que, comme nous venons de le voir, le nombre d'auteurs moyen et médian est le plus élevé en recherche biomédicale et le moins élevé en génie (Tableau XIX, p. 112) et le nombre moyen d'inventeurs est beaucoup plus stable entre les disciplines (Tableau XX, p. 113).

Tableau XXI. Statistiques descriptives sur le ratio inventeurs/auteurs des PBA par discipline en fonction de la collaboration interinstitutionnelle et du facteur d'impact de la revue de publication.

Discipline	Groupe	PBA		Ratio inventeurs/auteurs				
		N	Min	Max	Méd.	Moy.	Écart type	Asym.
Génie	Uni-institution - bas FI	634	0,125	1	1,000	0,816	0,253	-0,945
	Uni-institution - haut FI	58	0,200	1	0,792	0,755	0,264	-0,518
	Multi-institution - bas FI	536	0,077	1	0,833	0,765	0,270	-0,701
	Multi-institution - haut FI	71	0,111	1	0,750	0,724	0,279	-0,435
	<i>Ensemble</i>	1299	0,077	1	1	0,787	0,264	-0,789
Recherche biomédicale	Uni-institution - bas FI	820	0,038	1	0,500	0,601	0,293	0,160
	Uni-institution - haut FI	48	0,125	1	0,571	0,641	0,294	0,011
	Multi-institution - bas FI	1906	0,045	1	0,400	0,454	0,280	0,674
	Multi-institution - haut FI	254	0,043	1	0,308	0,371	0,247	1,000
	<i>Ensemble</i>	3028	0,038	1	0,429	0,490	0,291	0,534
Sciences naturelles	Uni-institution - bas FI	935	0,071	1	0,889	0,755	0,276	-0,595
	Uni-institution - haut FI	66	0,071	1	1,000	0,792	0,282	-0,980
	Multi-institution - bas FI	852	0,038	1	0,571	0,599	0,286	0,148
	Multi-institution - haut FI	132	0,125	1	0,646	0,650	0,265	0,016
	<i>Ensemble</i>	1985	0,038	1	0,667	0,682	0,290	-0,232

La Figure 5 (p. 115) présente la relation entre le nombre d'auteurs et le ratio inventeurs/auteurs des PBA dans chacune des disciplines et dans l'ensemble. Nous observons une décroissance du ratio inventeurs/auteurs à mesure que le nombre d'auteurs augmente, ce qui suggère qu'un plus grand nombre d'auteurs ne se traduit pas systématiquement par une augmentation équivalente du nombre d'inventeurs pour le brevet correspondant. De plus, nous observons qu'indépendamment du nombre d'auteurs, le ratio inventeurs/auteurs est

généralement le plus élevé en génie, suivi des sciences naturelles et de la recherche biomédicale. Nous évitons de tirer des conclusions de l'inversion des tendances entre le génie et les sciences naturelles à partir des articles ayant sept auteurs, puisque 97 % des PBA en génie ont moins de sept auteurs. Les différences disciplinaires au niveau du nombre d'auteurs, d'inventeurs et du ratio inventeurs/auteurs présentées dans le Tableau XIX (p. 112), le Tableau XX (p. 113) et le Tableau XXI (p. 114) auraient pu porter à attribuer les différences disciplinaires au niveau du ratio inventeurs/auteurs moyen au fait que le nombre moyen d'auteurs varie davantage entre les disciplines que le nombre moyen d'inventeurs. Cependant, la Figure 5 (p. 115) montre que pour les articles ayant entre deux et six auteurs (78,4 % de l'ensemble des PBA sont dans cette plage) le ratio inventeurs/auteurs moyen est le plus faible en recherche biomédicale et le plus élevé en génie, indépendamment du nombre d'auteurs.

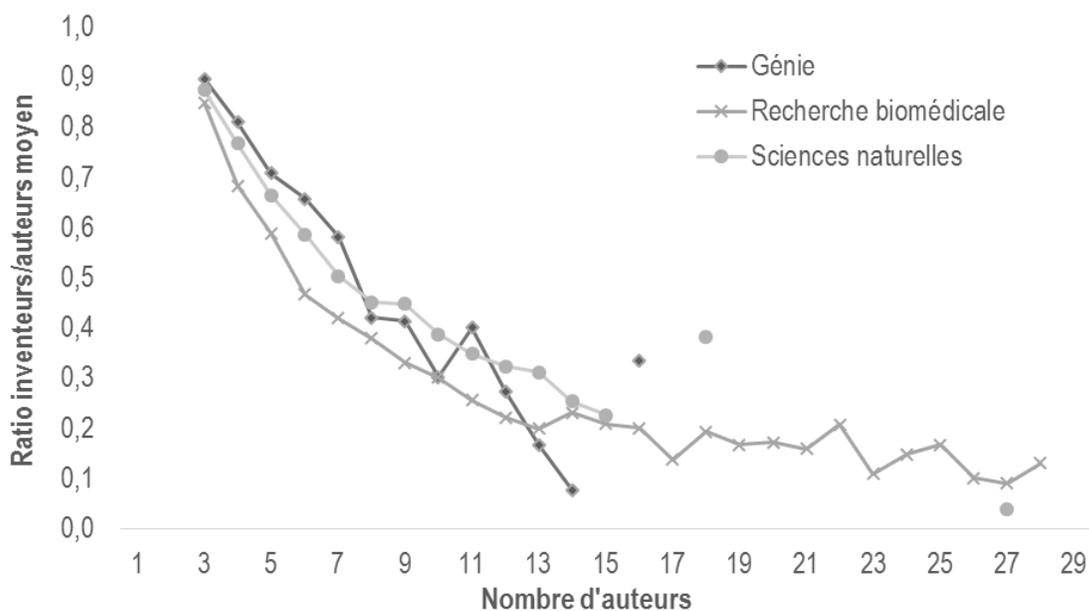


Figure 5. Ratio inventeurs/auteurs moyen en fonction du nombre d'auteurs et par discipline.

La Figure 6 (p. 116), la Figure 7 (p. 117) et la Figure 8 (p. 117) présentent la relation entre le nombre d'auteurs et le ratio inventeurs/auteurs pour les PBA résultant d'une collaboration interinstitutionnelle ainsi que pour celles dont l'article est publié dans une revue à haut FI en génie, recherche biomédicale et en sciences naturelles, respectivement. Les articles en collaboration interinstitutionnelle semblent peu différer de l'ensemble, ce qui suggère qu'en maintenant le nombre d'auteurs constant, la collaboration interinstitutionnelle est faiblement

liée au ratio inventeurs/inventeurs. Le ratio inventeurs/auteurs semble positivement lié au facteur d'impact de la revue de publication en sciences naturelles puisque la courbe tend à demeurer au-dessus de la moyenne pour l'ensemble des PBA de cette discipline. Ce n'est cependant pas le cas en génie et recherche biomédicale où la courbe pour les PBA dont l'article est publié dans une revue à haut FI demeure très près de la moyenne pour l'ensemble.

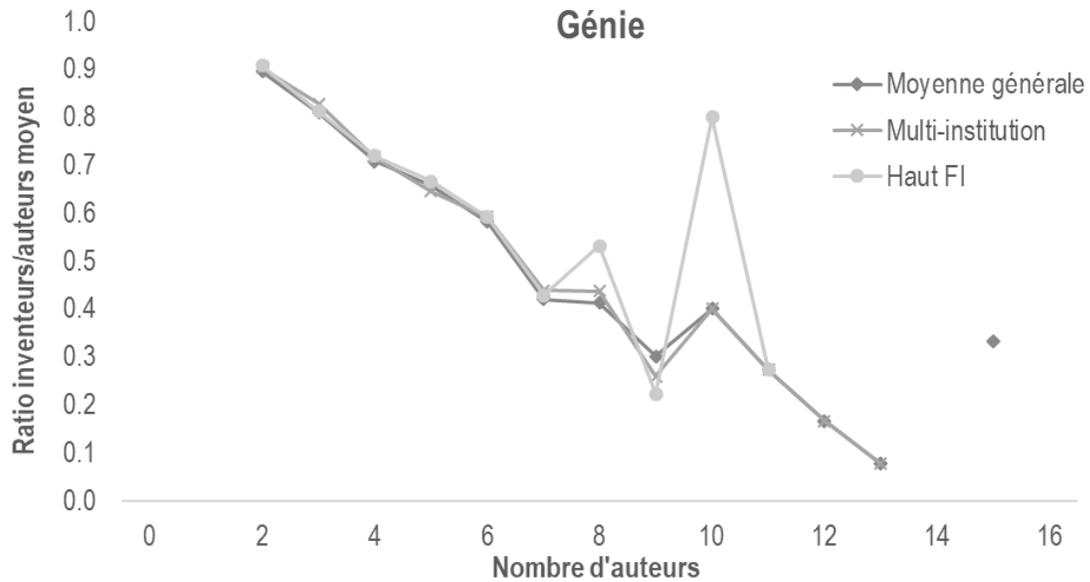


Figure 6. Ratio inventeurs/auteurs moyen en génie en fonction du nombre d'auteurs pour l'ensemble des articles, les articles en collaboration interinstitutionnelle et les articles publiés dans une revue à haut FI.

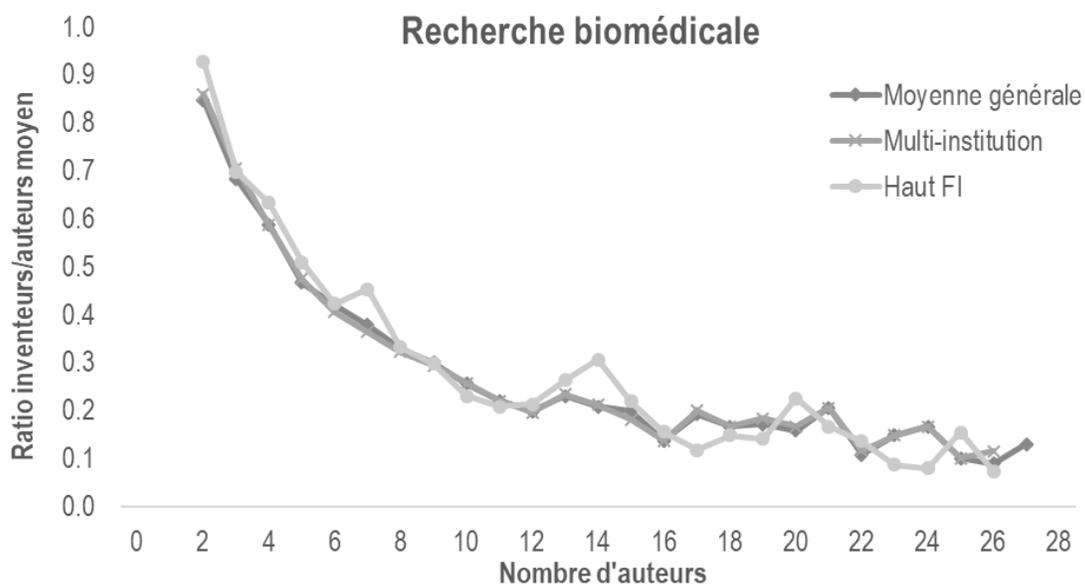


Figure 7. Ratio inventeurs/auteurs moyen en recherche biomédicale en fonction du nombre d'auteurs pour l'ensemble des articles, les articles en collaboration interinstitutionnelle et les articles publiés dans une revue à haut FI.

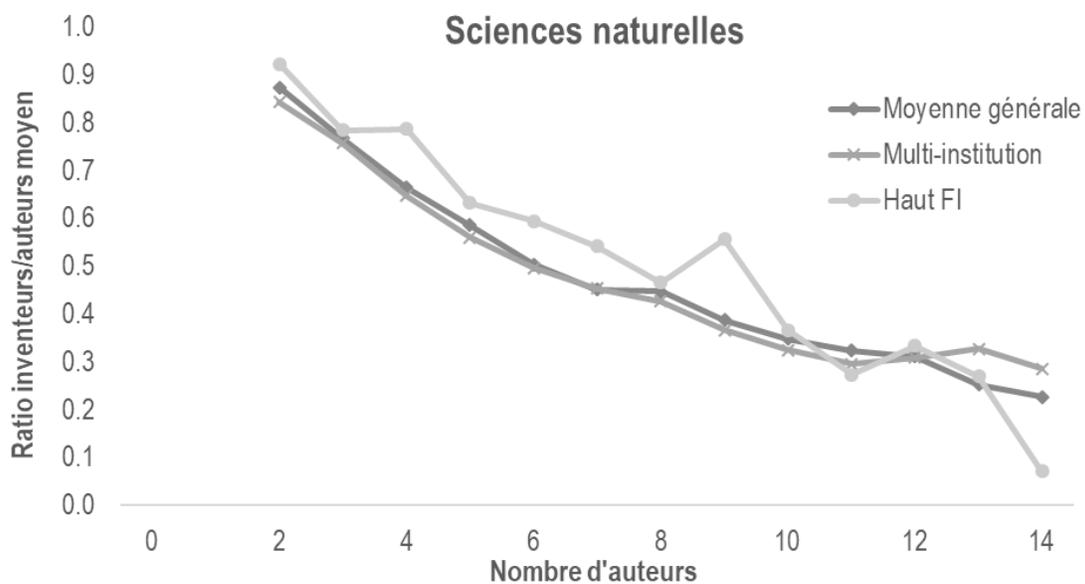


Figure 8. Ratio inventeurs/auteurs moyen en sciences naturelles, en fonction du nombre d'auteurs pour l'ensemble des articles, les articles en collaboration interinstitutionnelle et les articles publiés dans une revue à haut FI.

Dans la section suivante, nous présentons les résultats de l'analyse de régression linéaire visant à déterminer l'ampleur et la significativité statistique des relations entre les variables prédictives et prédites en contrôlant pour l'année de publication de l'article.

3.1.1.2. Modèle de régression linéaire

Nous utilisons un modèle de régression linéaire multiple hiérarchique afin de vérifier l'ampleur et la signification statistique de la relation entre le ratio inventeurs/auteurs et le nombre d'auteurs, la collaboration interinstitutionnelle, le facteur d'impact de la revue et la discipline. Pour cette analyse, la variable « discipline » est divisée en trois variables dichotomiques (« génie », « biomed », « sc_nat »). Nous utilisons également le logarithme du nombre d'auteurs (variable « log(nb_auteurs) »). Cette transformation de la variable « nb_auteurs » vise à forcer la linéarité de la relation entre la cette variable et la variable prédite, et ainsi à respecter ce postulat de la régression linéaire. Nous incluons également l'année de publication de l'article (année_pub) comme variable de contrôle.

3.1.1.2.1. Vérification des postulats de la régression linéaire multiple

Afin d'assurer la validité des conclusions statistiques tirées de notre analyse, nous avons vérifié que notre jeu de données respectait les postulats de la régression linéaire mentionnés au chapitre précédent et dont voici un rappel :

1. La relation entre chacune des variables prédictives continues et la variable prédite doit être linéaire.
2. Les observations doivent être indépendantes les unes des autres (indépendance des erreurs).
3. Il ne doit pas y avoir de valeurs extrêmes pouvant affecter la capacité prédictive du modèle.
4. Les variables prédictives ne doivent pas être fortement corrélées entre elles (absence de multicolinéarité).
5. Les valeurs résiduelles sont normalement distribuées pour chacune des valeurs de la variable prédite (normalité des erreurs).
6. La relation entre les valeurs résiduelles et la variable prédite doit être linéaire (linéarité des erreurs).
7. La variance des valeurs résiduelles doit être similaire pour chacune des valeurs de la variable prédite (homocédasticité).

Nous avons d'abord identifié comme valeur extrême les observations dont la valeur du résidu de Student supprimé était supérieure à ± 3 écarts types ainsi que les cas dont la valeur de la statistique de distance de Cook était supérieure à 1. Au total, onze valeurs extrêmes ont été identifiées et retirées de l'analyse (N = 6 301).

Afin de nous assurer que les distributions des variables indépendantes continues sont normales, nous utilisons le logarithme du nombre d'auteurs « $\log(\text{nb_auteurs})$ » dans le modèle. L'Annexe 6 (p. xxvii) présente un ensemble de diagrammes et de tableaux produit par SPSS comme extraits de l'analyse de régression et utilisés pour vérifier le respect des autres postulats. L'inspection visuelle des diagrammes de dispersion a permis de confirmer que la relation entre la variable prédite et chacune des variables prédictives continues (ce postulat ne concerne pas les variables nominales) est linéaire ou approximativement linéaire. De plus, l'inspection visuelle du tracé P-P normal et de l'histogramme des résidus standardisés a permis de confirmer que les résidus sont approximativement normalement distribués. Pour vérifier l'absence de multicolinéarité, nous nous sommes assuré qu'aucun des coefficients de corrélation de Pearson entre les variables prédictives ou de contrôle n'était supérieur à 0,8. Nous avons également consulté le facteur d'inflation de la variance (VIF), une statistique fournie par SPSS et indiquant un problème potentiel de multicolinéarité lorsque sa valeur se rapproche de 10. Dans notre cas, le VIF le plus élevé est 1,448 (pour la variable « génie »), ce qui confirme l'absence de multicolinéarité entre les variables de notre modèle. L'indépendance des erreurs a été vérifiée à l'aide de la statistique de Durbin-Watson. Celle-ci peut prendre une valeur allant de 0 à 4, l'indépendance parfaite des erreurs se situant au point central, soit la valeur 2. La valeur obtenue est de 1,326, ce qui confirme le respect de ce postulat. L'homocédasticité a été confirmée par l'inspection visuelle du diagramme de dispersion mettant en relation les résidus standardisés et les valeurs prédites non standardisées.

3.1.1.2.2. Significativité statistique et capacité prédictive du modèle

Le modèle de base de notre régression linéaire hiérarchique inclut seulement le logarithme du nombre d'auteurs comme prédicteur du ratio inventeurs/auteurs et l'année de publication comme variable de contrôle. Nous ajoutons successivement les variables disciplinaires (« génie », « biomed » et « sc_nat »), la collaboration interinstitutionnelle

(« multi_institutions ») et le facteur d'impact de la revue (« haut_fi ») afin de déterminer l'apport respectif de chacune de ces variables à la capacité prédictive du modèle global. Comme les trois variables disciplinaires correspondent en fait aux trois modalités possibles de la variable « discipline », l'une de ces modalités est utilisée comme modalité de référence. Ici, la modalité « biomed » est la référence, et elle n'apparaît donc pas dans les prédicteurs du Tableau XXII (p. 120) et du Tableau XXIV (p. 122). Nous utilisons « biomed » comme modalité de référence, car les résultats précédents (Tableau XXI, p. 114) suggèrent qu'elle est associée à des ratios inventeurs/auteurs généralement plus faibles que les autres disciplines, ce qui permettra d'obtenir des rapports de cotes supérieurs à 1 et plus faciles à interpréter (Rizzi, 2013). Le Tableau XXII (p. 120) montre que la signification statistique des différences dans la statistique F calculée pour chacun des modèles est inférieure à 0,05, ce qui indique que chaque variable améliore significativement le modèle global. Cependant, la différence de R² entre chacun des modèles montre que l'apport des variables « multi_institutions » et « haut_fi » est relativement faible puisqu'ils ajoutent respectivement moins de 0,05 % et 0,1 % à la capacité prédictive du modèle. La prise en compte de la discipline a cependant un impact plus important sur la capacité prédictive du modèle (1,7 %).

Tableau XXII. Qualité de l'ajustement et valeur prédictive du modèle de régression linéaire multiple hiérarchique visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI.

Modèle	R	R ²	R ² ajusté	Erreur standard	Variations				
					Δ R ²	Δ F	ddl1	ddl2	Sig. Δ F
1	0,656 ^a	0,430	0,430	0,234164	0,430	2 380,264	2	6 298	0,000
2	0,669 ^b	0,448	0,447	0,230671	0,017	97,096	2	6 296	0,000
3	0,669 ^c	0,448	0,447	0,230612	0,000	4,221	1	6 295	0,040
4	0,670 ^d	0,449	0,449	0,230320	0,001	16,961	1	6 294	0,000

a. Prédicteurs : (Constante), année_pub, log(nb_auteurs)

b. Prédicteurs : (Constante), année_pub, log(nb_auteurs), génie, sc_nat

c. Prédicteurs : (Constante), année_pub, log(nb_auteurs), génie, sc_nat, multi_institutions

d. Prédicteurs : (Constante), année_pub, log(nb_auteurs), génie, sc_nat, multi_institutions, haut_fi

Le Tableau XXIII (p. 121) présente les résultats de l'analyse de variance (ANOVA) pour le modèle incluant l'ensemble des prédicteurs et indique que le modèle est statistiquement significatif ($F = 856,123$; $p < 0,001$).

Tableau XXIII. Significativité statistique du modèle de régression linéaire visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI.

	Somme des carrés	ddl	Carré moyen	F	Sig.
Régression	272,491	6	45,415	856,123	0,000
Résidu	333,881	6 294	0,053		
Total	606,372	6 300			

3.1.1.2.3. Coefficients des prédicteurs

Le Tableau XXIV (p. 122) présente les coefficients de régression non standardisés (B) et standardisés (Bêta) pour l'ensemble des prédicteurs du modèle. Le principal prédicteur du modèle est le logarithme du nombre d'auteurs qui est fortement et négativement lié au ratio inventeurs/auteurs ($B = -0,778$; $p < 0,001$). Le deuxième principal prédicteur est la discipline. Les coefficients des variables « génie » ($B = 0,103$; $p < 0,001$) et « sc_nat » ($B = 0,073$; $p < 0,001$) montrent que le ratio inventeurs/auteurs tend à être plus élevé dans ces disciplines, comparativement à la recherche biomédicale. Selon le modèle, la collaboration interinstitutionnelle est négativement liée au ratio inventeurs/auteurs ($B = -0,014$; $p < 0,05$), alors qu'un haut FI est positivement lié au ratio inventeurs/auteurs ($B = 0,041$; $p < 0,001$). Dans l'ensemble, le modèle montre que, en ordre d'importance, le nombre d'auteurs est le plus fort prédicteur du ratio inventeurs/auteurs, suivi de la discipline, du FI de la revue et de la collaboration interinstitutionnelle. La force de la relation entre ces deux dernières variables et le ratio inventeurs/auteurs est cependant beaucoup plus faible (deux ordres de grandeur) que celle de la relation entre le logarithme du nombre d'auteurs et le ratio inventeurs/auteurs.⁴⁰

⁴⁰ Comme la force (ou l'importance) d'une relation est difficile à évaluer de façon absolue, l'importance d'un prédicteur dans un modèle est interprétée relativement aux autres prédicteurs du même modèle dans l'ensemble du chapitre.

Tableau XXIV. Coefficients du modèle de régression linéaire visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI.

Variables	B	ES	Bêta	t	Sig.	Intervalle de confiance à 95,0 % pour B	
						Inférieur	Supérieur
Constante	-5,837	1,043		-5,594	0,000	-7,882	-3,791
année_pub	0,003	0,001	0,063	6,592	0,000	0,002	0,004
log(nb_auteurs)	-0,778	0,014	-0,596	-53,679	0,000	-0,807	-0,750
Génie	0,103	0,008	0,134	12,160	0,000	0,086	0,119
sc_nat	0,073	0,007	0,109	10,403	0,000	0,059	0,087
multi_institutions	-0,014	0,007	-0,023	-2,205	0,027	-0,027	-0,002
haut_fi	0,041	0,010	0,039	4,118	0,000	0,021	0,060

Comme le modèle compare les PBA en génie et en sciences naturelles aux PBA en recherche biomédicale (« biomed » ayant été choisie comme modalité de référence de la variable « discipline »), il ne permet pas de déterminer s'il existe des différences significatives entre les PBA en génie et en sciences naturelles. Nous avons donc effectué une nouvelle régression excluant les PBA en recherche biomédicale (Tableau XXV, p. 122), ce qui a permis de comparer les PBA en génie et en sciences naturelles. Le coefficient de la variable « sc_nat » permet de confirmer que, *ceteris paribus*, le ratio inventeurs/auteurs tend à être inférieur en sciences naturelles comparativement au génie (B = -0,032; p < 0,001).

Tableau XXV. Coefficients du modèle de régression linéaire visant à prédire le ratio inventeurs/auteurs des PBA à l'aide du nombre d'auteurs, de la discipline, de la collaboration interinstitutionnelle et du FI (recherche biomédicale exclue).

Variables	B	ES	Bêta	T	Sig.	Intervalle de confiance à 95,0 % pour B	
						Inférieur	Supérieur
Constante	-5,837	1,559		-4,213	0,000	-9,627	-3,512
année_pub	0,004	0,001	0,076	4,944	0,000	0,002	0,005
log(nb_auteurs)	-0,742	0,024	-0,514	-31,490	0,000	-0,788	-0,696
sc_nat	-0,032	0,009	-0,056	-3,623	0,000	-0,050	-0,015
multi_institutions	-0,019	0,009	-0,033	-2,067	0,039	-0,036	-0,001
haut_fi	0,036	0,014	0,037	2,467	0,007	0,007	0,064

Afin de mieux observer les différences disciplinaires, nous analysons chacune des disciplines individuellement afin de déterminer si la significativité de la relation entre les différents prédicteurs et le ratio inventeurs/auteurs est indépendante de la discipline. Les trois modèles disciplinaires sont présentés dans le Tableau XXVI (p. 123) et confirment que le nombre d'auteurs est le principal prédicteur dans chacune des disciplines. La collaboration interinstitutionnelle est un prédicteur significatif en sciences naturelles ($B = -0,043$; $p < 0,001$), mais non en génie ($B = 0,014$; $p = 0,303$) et en recherche biomédicale ($B = -0,010$; $p = 0,270$). De plus, la relation entre le FI de la revue et le ratio inventeurs/auteurs demeure significative en recherche biomédicale ($B = 0,045$; $p < 0,01$) et en sciences naturelles ($B = 0,061$; $p < 0,01$), mais non en génie ($B = -0,001$; $p = 0,955$).

Tableau XXVI. Coefficient de la régression linéaire visant à prédire le ratio inventeurs/auteurs pour chaque discipline à l'aide du nombre d'auteurs de la collaboration interinstitutionnelle et du FI.

	Variables	B	ES	Bêta	t	Sig.	Intervalle de confiance à 95,0 % pour B	
							Inférieur	Supérieur
génie	Constante	-7,028	2,578		-2,726	0,006	-12,086	-1,970
	année_pub	0,004	0,001	0,081	3,160	0,002	0,002	0,007
	log(nb_auteurs)	-0,722	0,042	-0,457	-17,276	0,000	-0,804	-0,640
	multi_institutions	0,014	0,014	0,027	1,031	0,303	-0,013	0,041
	haut_fi	-0,001	0,023	-0,001	-0,056	0,955	-0,046	0,043
biomed	Constante	-5,114	1,418		-3,606	0,000	-7,895	-2,333
	année_pub	0,003	0,001	0,061	4,351	0,000	0,002	0,004
	log(nb_auteurs)	-0,803	0,018	-0,669	-44,300	0,000	-0,838	-0,767
	multi_institutions	-0,010	0,009	-0,016	-1,104	0,270	-0,029	0,008
	haut_fi	0,045	0,014	0,046	3,261	0,001	0,018	0,071
sc_nat	Constante	-6,450	1,954		-3,301	0,001	-10,282	-2,618
	année_pub	0,004	0,001	0,075	3,868	0,000	0,002	0,006
	log(nb_auteurs)	-0,739	0,029	-0,527	-25,698	0,000	-0,795	-0,682
	multi_institutions	-0,043	0,012	-0,074	-3,594	0,000	-0,066	-0,019
	haut_fi	0,061	0,019	0,062	3,238	0,001	0,024	0,097

3.1.1.2.4. Vérification des hypothèses

Les hypothèses testées dans cette section de l'analyse, à l'aide d'un modèle de régression linéaire multiple, sont les suivantes :

HS₁ Plus le facteur d'impact de la revue de publication de l'article formant une PBA est élevé, plus le ratio inventeurs/auteurs sera élevé.

Alors que le modèle général (Tableau XXIV, p. 122) prédit un ratio inventeurs/auteurs plus élevé pour les PBA dont l'article est publié dans une revue à haut facteur d'impact, nous observons cependant que cette relation n'est pas significative en génie lorsque nous modélisons la relation entre les variables prédictives et le ratio inventeurs/auteurs pour chaque discipline individuellement (Tableau XXVI, p. 123). Nos analyses permettent donc d'accepter partiellement l'hypothèse HS₁.

HS₂ Le ratio inventeurs/auteurs est influencé par la discipline.

Les résultats présentés au Tableau XXIV (p. 122) et au Tableau XXV (p. 122) montrent que, *ceteris paribus*, le ratio inventeurs/auteurs moyen diffère significativement d'une discipline à l'autre. En effet, comparativement à la recherche biomédicale, le ratio inventeurs/auteurs est significativement plus élevé en génie et en sciences naturelles, le ratio le moins élevé étant en recherche biomédicale et le plus élevé en génie (Tableau XXIV, p. 122). Le modèle excluant les PBA en recherche biomédicale (Tableau XXV, p. 122) montre que le ratio inventeurs/auteurs est aussi significativement plus faible en sciences naturelles qu'en génie. Ces résultats permettent d'accepter l'hypothèse HS₂.

HS₃ Plus le nombre d'auteurs d'une PBA est élevé, plus le ratio inventeurs/auteurs sera faible.

Nos résultats permettent également d'accepter l'hypothèse HS₃. En effet, le logarithme du nombre d'auteurs des articles est fortement et négativement lié au ratio inventeurs/auteurs dans tous les modèles du Tableau XXVI (p. 123). De plus, le Tableau XXII (p. 120) montre que le logarithme du nombre d'auteurs d'un article explique à lui seul 43 % de la variance du ratio inventeurs/auteurs (contre 44,9 % pour le modèle incluant l'ensemble des prédicteurs). Le nombre d'auteurs est donc le principal prédicteur du ratio inventeurs/auteurs.

HS₄ Les PBA issues d'une collaboration interinstitutionnelle auront un ratio inventeurs/auteurs plus faible que les PBA dont tous les auteurs sont affiliés à la même institution.

Le coefficient de la variable «multi_institutions» est négatif et statistiquement significatif dans les modèles incluant la discipline comme prédicteur (Tableau XXIV, p. 122;

Tableau XXV, p. 122). Cependant, lorsque nous considérons les disciplines individuellement (Tableau XXVI, p. 123), le coefficient demeure significatif en sciences naturelles, mais non en génie et en recherche biomédicale. Il est également relativement bas, suggérant une relation plutôt faible entre la collaboration interinstitutionnelle et le ratio inventeurs/auteurs. En somme, nos résultats montrent que les PBA issues d'une collaboration interinstitutionnelle ont un ratio inventeurs/auteurs légèrement plus faible que les autres PBA en sciences naturelles seulement. L'hypothèse HS₄ est donc partiellement acceptée.

3.1.2. En résumé

Dans cette première partie de l'analyse, nous avons présenté des résultats descriptifs et construit un modèle de régression linéaire afin de vérifier nos quatre hypothèses concernant les liens entre les caractéristiques des travaux et des équipes de recherche et le ratio inventeurs/auteurs des PBA. Les résultats ont permis de démontrer que le principal prédicteur du ratio inventeurs/auteurs est le logarithme du nombre d'auteurs, ce qui est logique étant donné que le nombre d'auteurs sert de dénominateur dans le calcul de ce ratio. Nous observons également l'existence de différences disciplinaires concernant la relation entre le nombre d'auteurs et d'inventeurs des PBA. En effet, toutes choses étant égales par ailleurs, on trouve le ratio inventeurs/auteurs le plus élevé en génie et le moins élevé en recherche biomédicale. Le lien observé entre la collaboration interinstitutionnelle et le ratio inventeurs/auteurs est relativement faible. Il est statistiquement significatif pour l'ensemble ainsi qu'en sciences naturelles, mais n'atteint pas le seuil de significativité en génie et en recherche biomédicale. De même, la publication dans une revue à haut FI est positivement et significativement liée au ratio inventeurs/auteurs dans tous les modèles, sauf en génie.

3.2. Prédicteurs du statut d'inventeur

Nous nous intéressons dans cette seconde section à la relation entre le statut d'auteur et celui d'inventeur. Voici en rappel les objectifs et questions de recherche auxquels cette analyse vise à répondre.

Objectif 3. Déterminer la relation entre les contributions respectives des auteurs et l'obtention du statut d'inventeur.

- QR5. Quelle est la relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur?
- QR6. La relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur est-elle liée à la discipline?
- QR7. Quelle est la relation entre le rôle dans l'équipe de recherche et l'obtention du statut d'inventeur?
- QR8. Quelle est la relation entre le type de contribution à la recherche et l'obtention du statut d'inventeur?

Objectif 4. Déterminer la relation entre des caractéristiques individuelles des auteurs et l'obtention du statut d'inventeur.

- QR9. Quelle est la relation entre la réputation d'un auteur et l'obtention du statut d'inventeur?
- QR10. Quelle est la relation entre le sexe d'un auteur et l'obtention du statut d'inventeur?

L'analyse est effectuée en trois sections, où nous présentons des résultats descriptifs et des modèles de régression logistique montrant la relation entre le statut d'inventeur et 1) la position dans la liste des auteurs, 2) le rôle et 3) la contribution à la recherche. Chacune de ces analyses inclut également les caractéristiques individuelles, soit le sexe et la réputation. Les différences entre les disciplines sont quant à elles considérées uniquement dans la première analyse puisque les échantillons de PBA en sciences naturelles et en génie disponibles pour répondre à QR7 sont trop petits, et parce que les données utilisées pour répondre à QR8 ont seulement été collectées pour les PBA en recherche biomédicale.

3.2.1. L'obtention du statut d'inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation

Nous analysons dans cette section l'ensemble des auteurs des PBA dont le sexe a pu être identifié. Nous présentons d'abord quelques résultats descriptifs mettant en relation les trois

variables prédictives (position, sexe, réputation) et la variable prédite (inventeur). Par la suite, nous présentons les résultats de l'analyse de régression logistique effectuée afin de vérifier les hypothèses de recherche HS₅, HS₆, HS₁₀, et HS₁₁.

3.2.1.1. Résultats descriptifs

La Figure 9 (p. 128), la Figure 10 (p. 128) et la Figure 11 (p. 129) présentent, pour le génie, la recherche biomédicale et les sciences naturelles respectivement, la proportion d'auteurs ayant obtenu le statut d'inventeur en fonction de leur position dans la liste des auteurs. Nous divisons les auteurs en quatre groupes selon leur sexe (hommes ou femmes) et selon leur réputation (peu cité, hautement cité), un auteur étant considéré hautement cité lorsque le nombre de citations reçues est au-dessus de la médiane globale pour l'ensemble des PBA de la discipline sur toute la période étudiée⁴¹. Les figures montrent qu'à quelques exceptions près, les hommes obtiennent plus souvent le statut d'inventeur que les femmes pour une même position dans la liste des auteurs et un même niveau de citations. Les exceptions sont surtout en génie, au niveau des auteurs peu cités occupant la deuxième et l'antépénultième position, ainsi qu'au niveau des auteurs hautement cités occupant la première, l'antépénultième et la pénultième position. En recherche biomédicale, les derniers auteurs hautement cités en recherche biomédicale font aussi exception, une proportion similaire d'hommes et de femmes ayant obtenu le statut d'inventeur. En sciences naturelles, la proportion d'hommes et de femmes ayant obtenu le statut d'inventeur est similaire pour les auteurs hautement cités occupant la troisième et l'antépénultième position. Les auteurs hautement cités sont aussi, à quelques exceptions près, plus souvent inventeurs que les auteurs peu cités du même sexe et occupant la même position dans la liste des auteurs. Les exceptions sont les femmes occupant la deuxième position en génie, les hommes occupant l'antépénultième position en génie, ainsi que les hommes et les femmes au milieu de la liste des auteurs en sciences naturelles.

⁴¹ Nous utilisons la médiane du nombre de citations pour l'ensemble de la période et non pour l'année afin de nous assurer que le nombre d'observations est assez élevé ainsi que pour éviter une trop grande variation du seuil d'inclusion dans le groupe « hautement cité » d'une année à l'autre. La médiane est utilisée plutôt que la moyenne afin de former des groupes de taille égale.

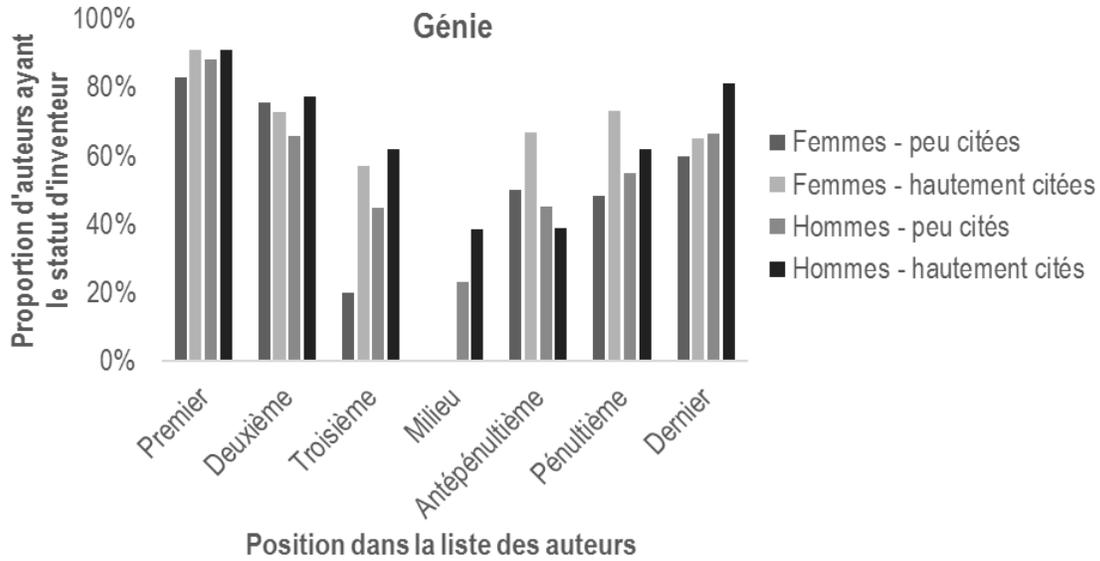


Figure 9. Proportion d’auteurs en génie ayant le statut d’inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.

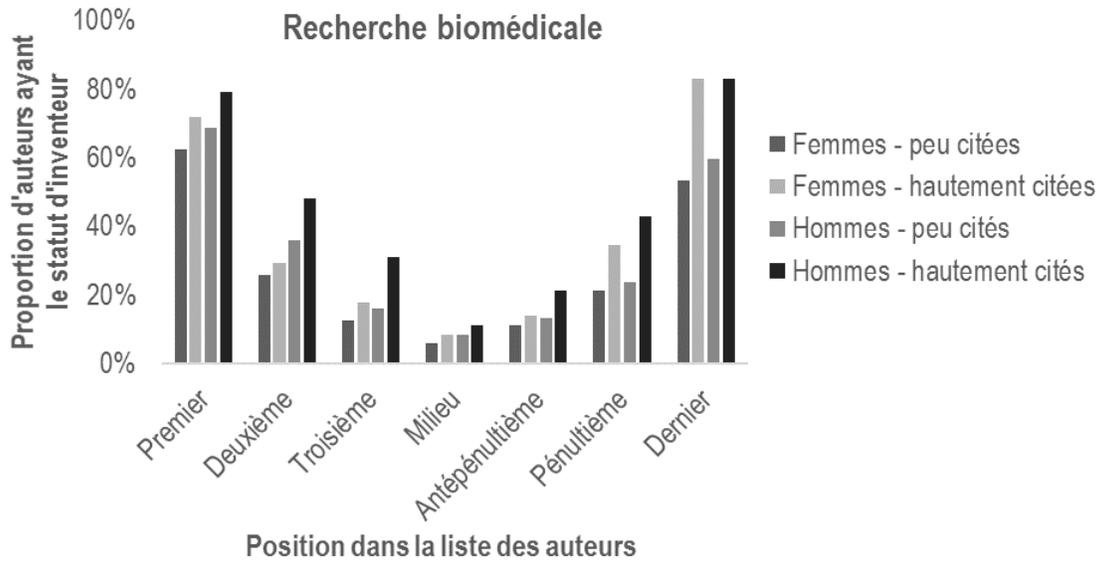


Figure 10. Proportion d’auteurs en recherche biomédicale ayant le statut d’inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.

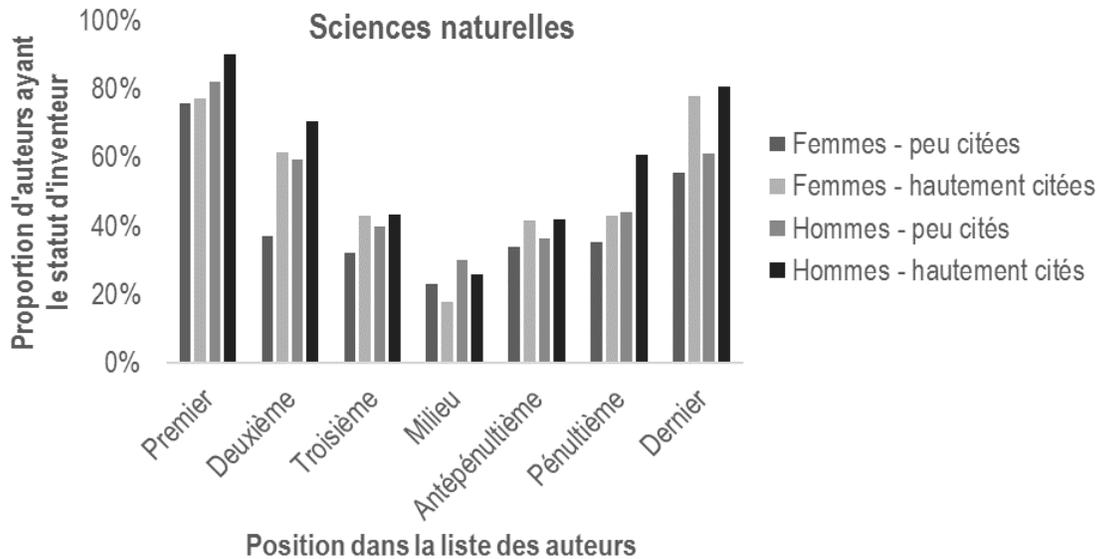


Figure 11. Proportion d’auteurs en sciences naturelles ayant le statut d’inventeur en fonction de la position dans la liste des auteurs, du sexe et de la réputation.

Nous remarquons aussi des différences disciplinaires, notamment en ce qui concerne la relation entre la position dans la liste des auteurs et l’obtention du statut d’inventeur. Alors que dans toutes les disciplines, la proportion d’inventeurs diminue à mesure que l’on s’éloigne de la première ou de la dernière position, l’ampleur de cette diminution diffère d’une discipline à l’autre. En génie, notamment, la proportion d’inventeurs demeure relativement élevée pour chacune des positions dans la liste des auteurs comparativement à la recherche biomédicale, les sciences naturelles se situant entre ces deux extrêmes. Les femmes au milieu de la liste des auteurs en génie sont ici l’exception, aucune d’entre elles n’ayant obtenu le statut d’inventeur. Ce résultat se doit cependant d’être interprété avec prudence puisqu’il est basé sur un très petit nombre d’observations (parmi les auteurs du milieu en génie, seulement six sont des femmes).

Dans la section suivante, nous introduisons l’ensemble de ces variables dans un modèle de régression logique afin de mieux comprendre leurs interrelations et afin, plus particulièrement, de tester leur significativité statistique.

3.2.1.2. Régression logistique

Notre analyse comporte quatre modèles de régression logistique (un par discipline et un pour l’ensemble) visant à prédire l’obtention du statut d’inventeur par un auteur à partir de sa

position dans liste des auteurs, de son sexe (homme ou femme) et de sa réputation (hautement cité ou peu cité).

3.2.1.2.1. Vérification des postulats de la régression logistique

Voici en rappel les postulats à respecter pour assurer la validité de conclusion statistique d'une régression logistique :

1. La relation entre chacune des variables prédictives continues et la variable prédite (ou le logit de la variable prédite dans le cas de la régression logistique) doit être linéaire.
2. Les observations doivent être indépendantes les unes des autres (indépendance des erreurs).
3. Les variables prédictives ne doivent pas être fortement corrélées entre elles (absence de multicollinéarité).
4. Il ne doit pas y avoir de valeurs extrêmes pouvant affecter la capacité prédictive du modèle.

Le premier postulat ne s'applique pas puisque nos modèles ne comportent que des variables nominales. Le respect du deuxième postulat est assuré par la conception de l'étude qui fait en sorte que les observations sont indépendantes les unes des autres (le fait qu'un auteur soit inventeur ne réduit pas la probabilité des autres auteurs du même article d'être inventeurs). Aucune valeur extrême (résidu de Student supprimé supérieur à ± 3 écarts types) n'a été identifiée⁴². L'absence de multicollinéarité a été vérifiée à l'aide des coefficients de Pearson. Aucun des coefficients de corrélation de Pearson entre les variables prédictives ne s'est révélé supérieur à 0,8, ce qui confirme l'absence de multicollinéarité. L'Annexe 6 (p. xxxi) présente les tableaux de corrélation pour cette analyse ainsi que pour les analyses des sections 3.2.2 et 3.2.3.

3.2.1.2.2. Qualité de l'ajustement et capacité prédictive des modèles

Nous construisons quatre modèles, soit un par discipline et un pour l'ensemble des PBA. Les modèles de base incluent seulement la variable « position », à laquelle nous ajoutons successivement le sexe et la réputation afin de déterminer l'apport respectif de chaque variable

⁴² La distance de Cook est utile pour détecter les valeurs extrêmes dans la régression linéaire et non dans la régression logistique. C'est pourquoi nous l'avons utilisé dans la partie 3.1.1.2, mais non dans l'ensemble des analyses de la section 3.2.

à la capacité prédictive du modèle. Les modalités « biomed » et « milieu » sont utilisées comme référence pour la discipline et la position, respectivement, puisque comme nous l'avons vu à la section 3.2.1.1, ce sont les groupes pour lesquels le ratio inventeurs/auteurs est en moyenne le plus bas. Le Tableau XXVII (p. 131) montre que les différences de khi carré (χ^2) sont statistiquement significatives ($p < 0,001$) pour chaque modèle, ce qui indique que chacune des variables améliore significativement le modèle global. Le sexe a cependant un apport moindre (environ 0,6 % selon le pseudo- R^2 de Nagelkerke) à la capacité prédictive du modèle que la discipline (2,5 %) et la réputation (1,6 %).

Tableau XXVII. Qualité de l'ajustement et capacité prédictive des modèles de régression logistique visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs

Modèle	Qualité de l'ajustement			Différence avec modèle précédent			Capacité prédictive	
	χ^2	ddl	Sig.	$\Delta \chi^2$	Δ ddl	Sig. $\Delta \chi^2$	-2LV	R_N^2
1 ^a	7 197,895	6	0,000	7 197,895	6	0,000	27 618,617	0,332
2 ^b	7 837,937	8	0,000	640,042	2	0,000	26 978,575	0,357
3 ^c	7 993,218	9	0,000	155,281	1	0,000	3 435,422	0,363
4 ^d	8 394,818	10	0,000	401,600	1	0,000	26 421,694	0,379

a. Prédicteurs : (Constante), premier, deuxième, troisième, antépénultième, pénultième, dernier

b. Prédicteurs : (Constante), premier, deuxième, troisième, antépénultième, pénultième, dernier, génie, sc_nat

c. Prédicteurs : (Constante), premier, deuxième, troisième, antépénultième, pénultième, dernier, génie, sc_nat, homme

d. Prédicteurs : (Constante), premier, deuxième, troisième, antépénultième, pénultième, dernier, génie, sc_nat, homme, hautement_cité

3.2.1.2.3. Coefficients et rapports de cotes des prédicteurs

Le Tableau XXVIII (p. 132) présente les coefficients de régression (B) et les rapports de cotes (RC) pour l'ensemble des prédicteurs du modèle global. Les coefficients des différentes modalités de la variable position reflètent les résultats descriptifs de la Figure 11 (p. 129) : ils sont tous statistiquement significatifs ($p < 0,001$) et sont plus élevés pour les premiers et derniers auteurs, diminuant graduellement à mesure que l'on s'éloigne de ces deux pôles. De plus, la probabilité d'obtenir le statut d'inventeur est plus élevée en génie (RC = 2,472; $p < 0,001$) et en sciences naturelles (RC = 1,993; $p < 0,001$), comparativement à la recherche biomédicale. Le fait d'être hautement cité augmente la probabilité d'être inventeur (RC = 1,886; $p < 0,001$). Le

sexe est le plus faible prédicteur du modèle, mais le fait d'être un homme augmente néanmoins la probabilité d'être inventeur de façon significative (RC = 1,461; $p < 0,001$). Dans l'ensemble la première et la dernière position dans la liste des auteurs sont les meilleurs prédicteurs de l'obtention du statut d'auteur, avec un RC d'un ordre de grandeur plus élevé que celui des autres prédicteurs.

Tableau XXVIII. Coefficients du modèle de régression logistique global visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs, de la discipline, du sexe et de la réputation.

Variables	B	ES	Wald	ddl	Sig.	RC	Intervalle de confiance à 95 % pour RC	
							Inférieur	Supérieur
premier	3,238	0,061	2786,178	1	0,000	25,491	22,603	28,748
deuxième	1,872	0,061	931,620	1	0,000	6,499	5,763	7,329
troisième	0,959	0,074	167,024	1	0,000	2,608	2,255	3,016
antépénultième	0,694	0,075	85,620	1	0,000	2,002	1,728	2,319
pénultième	1,477	0,061	585,219	1	0,000	4,379	3,885	4,936
dernier	2,836	0,060	2 234,629	1	0,000	17,044	15,154	19,171
génie	0,905	0,048	357,578	1	0,000	2,472	2,251	2,715
sc_nat	0,690	0,035	380,765	1	0,000	1,993	1,860	2,136
homme	0,379	0,037	102,156	1	0,000	1,461	1,357	1,572
hautement_cité	0,634	0,032	395,549	1	0,000	1,886	1,772	2,007
Constante	-2,794	0,059	2 215,739	1	0,000	0,061		

Pour vérifier la significativité statistique de la différence entre les coefficients des variables « génie » et « sc_nat », nous avons reproduit le modèle en utilisant les sciences naturelles comme discipline de référence (Tableau XXIX, p. 133). Cela nous a permis de confirmer que, *ceteris paribus*, les auteurs en génie ont significativement plus de chance d'être inventeurs que les auteurs en sciences naturelles (RC = 1,314; $p < 0,001$).

Tableau XXIX. Coefficients du modèle de régression logistique excluant la recherche biomédicale et visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs, de la discipline, du sexe et de la réputation.

Variables	B	ES	Wald	ddl	Sig.	RC	Intervalle de confiance à 95 % pour RC	
							Inférieur	Supérieur
premier	2,817	0,109	662,167	1	0,000	16,723	13,494	20,725
deuxième	1,588	0,108	214,601	1	0,000	4,896	3,958	6,055
troisième	0,712	0,129	30,568	1	0,000	2,038	1,583	2,623
antépénultième	0,562	0,130	18,781	1	0,000	1,754	1,360	2,261
pénultième	1,074	0,107	100,370	1	0,000	2,927	2,372	3,611
dernier	1,903	0,104	335,339	1	0,000	6,705	5,469	8,219
génie	0,273	0,051	28,745	1	0,000	1,314	1,189	1,452
homme	0,369	0,064	33,348	1	0,000	1,446	1,276	1,638
hautement cité	0,546	0,048	128,643	1	0,000	1,726	1,570	1,896
Constante	-1,614	0,109	218,087	1	0,000	0,199		

Afin de mieux observer les différences entre les disciplines, nous avons également généré trois modèles de régression logistique distincts, soit un pour chacune des disciplines. Les coefficients des prédicteurs des modèles disciplinaires sont présentés dans le Tableau XXX (p. 134). Les coefficients montrent qu'en génie et en sciences naturelles, les premiers auteurs sont beaucoup plus souvent inventeurs que les derniers auteurs, alors que ce n'est pas le cas en recherche biomédicale où la probabilité d'être inventeur est un peu plus élevée pour les derniers auteurs que pour les premiers. Les hommes sont plus souvent inventeurs que les femmes en recherche biomédicale (RC = 1,428; $p < 0,001$) et en sciences naturelles (RC = 1,555; $p < 0,001$). Toutefois, aucune relation statistiquement significative n'est observée entre le sexe et le statut d'inventeur en génie (RC = 1,222; $p = 0,091$). Finalement, les auteurs hautement cités sont plus souvent inventeurs dans chaque discipline, *ceteris paribus*. Ces modèles confirment que la première et la dernière position dans la liste d'auteurs sont les meilleurs prédicteurs de l'obtention du statut d'inventeur, avec des RC un ordre de grandeur plus élevé que les autres prédicteurs.

Tableau XXX. Coefficients du modèle de régression logistique pour chaque discipline visant à prédire le statut d'inventeur en fonction de la position dans la liste des auteurs, du sexe, et de la réputation.

Variables	B	ES	Wald	ddl	Sig.	RC	Intervalle de confiance à 95 % pour RC		
							Inférieur	Supérieur	
génie	premier	3,089	0,291	112,445	1	0,000	21,957	12,405	38,862
	deuxième	1,912	0,291	43,107	1	0,000	6,767	3,824	11,975
	troisième	1,010	0,333	9,199	1	0,002	2,746	1,430	5,274
	antépénultième	0,714	0,332	4,629	1	0,031	2,043	1,066	3,915
	pénultième	1,314	0,289	20,652	1	0,000	3,720	2,111	6,557
	dernier	1,955	0,283	47,854	1	0,000	7,065	4,060	12,294
	homme	0,200	0,119	2,849	1	0,091	1,222	,968	1,541
	hautement_cité	0,507	0,085	35,211	1	0,000	1,660	1,404	1,963
	Constante	-1,374	0,297	21,375	1	0,000	0,253		
biomed	premier	3,364	0,076	1983,778	1	0,000	28,898	24,922	33,509
	deuxième	1,896	0,077	611,808	1	0,000	6,657	5,728	7,736
	troisième	0,989	0,093	112,144	1	0,000	2,688	2,238	3,227
	antépénultième	0,655	0,095	46,997	1	0,000	1,924	1,596	2,320
	pénultième	1,590	0,076	435,721	1	0,000	4,902	4,222	5,691
	dernier	3,408	0,077	1979,726	1	0,000	30,206	25,995	35,099
	homme	0,357	0,047	57,371	1	0,000	1,428	1,303	1,567
	hautement_cité	0,664	0,043	238,381	1	0,000	1,942	1,785	2,113
	Constante	-2,948	0,072	1669,592	1	0,000	0,052		
sc_nat	premier	2,745	0,121	517,191	1	0,000	15,559	12,281	19,711
	deuxième	1,501	0,119	158,991	1	0,000	4,488	3,554	5,667
	troisième	0,657	0,141	21,602	1	0,000	1,929	1,462	2,545
	antépénultième	0,546	0,143	14,694	1	0,000	1,727	1,306	2,284
	pénultième	1,022	0,118	75,515	1	0,000	2,778	2,206	3,498
	dernier	1,961	0,114	293,625	1	0,000	7,104	5,677	8,890
	homme	0,441	0,076	33,328	1	0,000	1,555	1,338	1,806
	hautement_cité	0,549	0,058	88,353	1	0,000	1,731	1,544	1,941
	Constante	-1,647	0,120	189,234	1	0,000	0,193	27,722	35,147

3.2.1.2.4. Vérification des hypothèses

Le modèle de régression logistique utilisé dans cette partie de l'analyse permet de tester les hypothèses suivantes :

HS₅ Plus un auteur est près de la première ou dernière position dans la liste des auteurs de l'article, plus la probabilité qu'il soit inventeur est grande.

Les résultats présentés au Tableau XXVIII (p. 132), au Tableau XXIX (p. 133) et au Tableau XXX (p. 134) montrent que dans chacune des disciplines étudiées les rapports de cotes sont très élevés pour les premiers et derniers auteurs et décroissent à mesure que l'on s'éloigne de ces deux pôles. L'hypothèse HS₅ peut donc être acceptée.

HS₆ La relation entre la position dans la liste des auteurs et l'obtention du statut d'inventeur diffère en fonction de la discipline.

Les différences observées entre les disciplines dans le Tableau XXVIII (p. 132) et le Tableau XXX (p. 134) permettent d'accepter l'hypothèse HS₆. Nous remarquons en effet que le rapport de cotes varie d'une discipline à l'autre pour la même position dans la liste des auteurs. De plus, le Tableau XXX (p. 134) montre que comparativement aux auteurs du milieu, les premiers auteurs sont plus souvent nommés inventeurs que les derniers auteurs en génie et en sciences naturelles. La tendance contraire est observée en recherche biomédicale, où les derniers auteurs sont plus souvent inventeurs que les premiers auteurs.

HS₁₀ Les hommes ont une probabilité plus grande que les femmes d'être inventeurs.

Le modèle de régression logistique global (Tableau XXVIII, p. 132) montre que les femmes sont moins souvent nommées inventrices que les hommes. Cependant, lorsque nous considérons les disciplines individuellement (Tableau XXX, 134), nous observons que le sexe n'est pas un prédicteur significatif en génie. Puisque la relation entre le sexe et l'obtention du statut d'inventeur en recherche biomédicale sera testée à nouveau dans les deux analyses suivantes (section 3.2.2 et 3.2.3), nous nous contentons pour l'instant d'accepter provisoirement l'hypothèse HS₁₀ pour les PBA de cette discipline. Cependant, nous pouvons dès maintenant rejeter l'hypothèse HS₁₀ pour les PBA en génie, et l'accepter pour les PBA en sciences naturelles, puisque ces deux disciplines ne sont pas prises en compte dans les analyses des sections 3.2.2 et 3.2.3.

HS₁₁ Plus le nombre de citations d'un auteur est élevé, plus la probabilité qu'il soit inventeur est grande.

Les rapports de cotes de la variable « hautement_cité » montrent que la réputation des auteurs est positivement et significativement liée à l'obtention du statut d'inventeur pour les PBA dans l'ensemble (Tableau XXVIII, p. 132), ainsi que dans chaque discipline (Tableau XXX, p. 134). Cependant, puisque l'hypothèse HS₁₁ sera également testée dans les prochaines sections, nous nous contenons pour l'instant d'accepter l'hypothèse HS₁₁ provisoirement.

3.2.2. L'obtention du statut d'inventeur en fonction du rôle, du sexe et de la réputation en recherche biomédicale

Nous analysons dans cette section le sous-ensemble de 450 auteurs de PBA en recherche biomédicale pour lesquels le rôle (principal, superviseur ou périphérique) a pu être identifié. Nous présentons d'abord quelques résultats descriptifs illustrant la proportion d'auteurs ayant obtenu le statut d'inventeur en fonction de leur rôle, de leur sexe et de leur réputation. Nous présentons ensuite les résultats de l'analyse de régression logistique effectuée afin de vérifier les hypothèses de recherche HS₇, HS₈, HS₁₀ et HS₁₁.

3.2.2.1. Résultats descriptifs

La Figure 12 (p. 137) présente la proportion d'auteurs de notre jeu de données ayant obtenu le statut d'inventeur en fonction du rôle, du sexe et de la réputation. Nous observons que, dans l'ensemble, les auteurs périphériques figurent rarement dans la liste des inventeurs, comparativement aux auteurs principaux et superviseurs. Pour les superviseurs, surtout, le statut d'inventeur semble fortement lié au sexe et à la réputation, les hommes étant plus souvent inventeurs que les femmes, et les auteurs hautement cités obtenant beaucoup plus souvent le statut d'inventeur que ceux dont le nombre de citations est plus faible. Selon le graphique combinant le sexe et la réputation, le lien entre le statut d'inventeur et la réputation est particulièrement fort pour les hommes superviseurs et les femmes qui sont auteures principales.

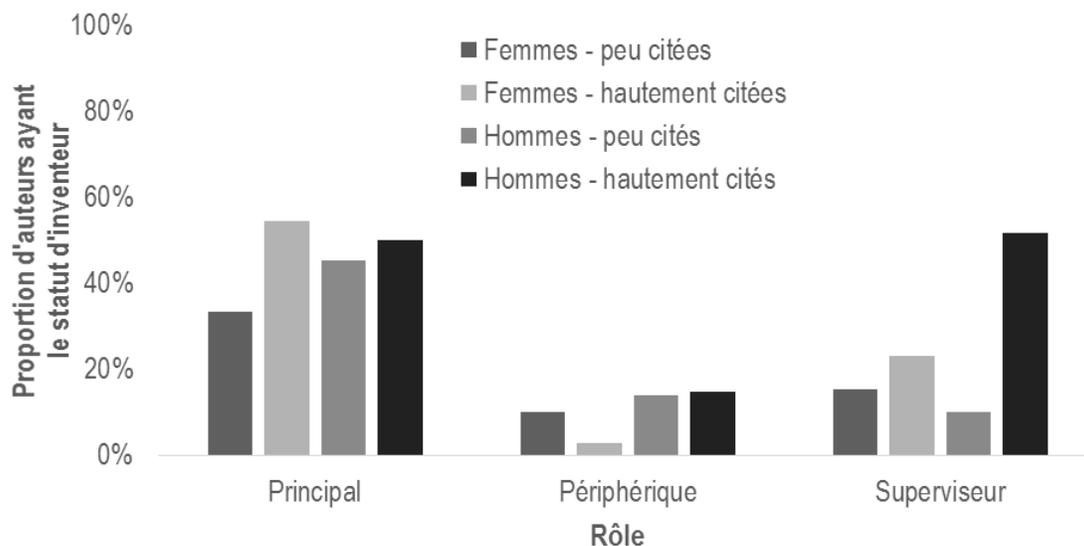


Figure 12. Proportion d’auteurs ayant le statut d’inventeur en fonction du rôle, du sexe et de la réputation.

3.2.2.2. Régression logistique

Nous utilisons un modèle de régression logistique pour tester les relations entre le rôle, le sexe et la réputation et l’obtention du statut d’inventeur, tout en contrôlant pour la position dans la liste des auteurs. Cependant, afin d’éviter de répartir les données dans un trop grand nombre de catégories comportant chacune un nombre d’observations trop faible pour obtenir des résultats statistiquement significatifs, la variable « position » n’a que trois modalités (« premier », « dernier » et « autres ») dans ce cas-ci.

3.2.2.2.1. Vérification des postulats de la régression logistique

Tout comme dans les modèles de régression de la section 3.2.1.2., les deux premiers postulats de la régression logistique (linéarité et indépendance des observations) sont respectés grâce à l’utilisation de variables nominales seulement ainsi qu’à la conception de notre étude. Aucune valeur extrême (résidu de Student supprimé supérieur à ± 3 écarts types) n’a été identifiée. Le modèle ne présente pas de problème de multicollinéarité puisque les coefficients de corrélation de Pearson présenté à l’Annexe 6 (p. xxxi) sont tous inférieurs à 0,8.

3.2.2.2.2. Qualité de l'ajustement et capacité prédictive du modèle

Notre modèle de départ inclut uniquement les trois modalités de la variable « position », la modalité « autres » étant ici choisie comme modalité de référence. Ce choix repose sur la démonstration faite précédemment (section 3.2.1.) que les premiers et derniers auteurs ont une probabilité plus grande d'être inventeurs que les autres auteurs. Nous y ajoutons successivement le rôle, le sexe et la réputation afin de déterminer l'apport respectif de chaque variable à la qualité de l'ajustement et à la capacité prédictive du modèle. Les auteurs périphériques sont la modalité de référence pour le rôle puisqu'en théorie ce sont ceux qui ont effectué les moins importantes contributions et ont donc une probabilité moins grande d'obtenir le statut d'inventeur. Le

Tableau XXXI (p. 138) illustre que les différences de khi carré (χ^2) sont statistiquement significatives ($p < 0,05$) lorsque le rôle et le sexe sont ajoutés au modèle. Ces deux variables augmentent respectivement le pseudo- R^2 de Nagelkerke (R_N^2) de 0,017 et 0,012. Cependant, l'inclusion de la variable « hautement_cité » n'améliore pas significativement la qualité de l'ajustement du modèle et sa capacité prédictive.

Tableau XXXI. Qualité de l'ajustement et capacité prédictive du modèle de régression logistique hiérarchique visant à prédire l'obtention du statut d'inventeur en fonction du rôle, du sexe et de la réputation.

Modèle	Qualité de l'ajustement			Différence avec modèle précédent			Capacité prédictive du modèle	
	χ^2	ddl	Sig.	$\Delta \chi^2$	Δ ddl	Sig. $\Delta \chi^2$	-2LV	R_N^2
1 ^a	93,483	1	0,000	93,483	1	0,000	413,715	0,277
2 ^b	99,645	2	0,000	6,162	1	0,013	407,553	0,294
3 ^c	104,347	3	0,000	4,702	1	0,030	402,850	0,306
4 ^d	104,396	4	0,000	0,049	1	0,825	402,802	0,306

a. Prédicteurs : (Constante), premier, dernier

b. Prédicteurs : (Constante), premier, dernier, principal, superviseur

c. Prédicteurs : (Constante), premier, dernier, principal, superviseur, homme

d. Prédicteurs : (Constante), premier, dernier, principal, superviseur, homme, hautement_cité

3.2.2.2.3. Coefficients et rapports de cotes des prédicteurs

Le Tableau XXXII (p. 139) présente les coefficients des prédicteurs et des variables de contrôle du modèle. Nous observons que la plupart des coefficients sont statistiquement significatifs ($p < 0,05$), le rôle de supervision (RC = 1,346; $p = 0,447$) et le fait d'être hautement cité (RC = 1,363; $p = 0,243$) faisant exception. Le principal prédicteur demeure la position dans la liste des auteurs, les derniers auteurs (RC = 9,826; $p < 0,001$) et les premiers auteurs (RC = 6,545; $p < 0,001$) étant plus souvent inventeurs que les autres auteurs. Au-delà de ces deux pôles, le fait d'être un auteur principal constitue aussi un prédicteur significatif mais plus faible de l'obtention du statut d'inventeur (RC = 2,737; $p < 0,01$). Finalement, le sexe est le prédicteur le plus faible, mais néanmoins significatif, les hommes ayant une probabilité plus élevée que les femmes d'obtenir le statut d'inventeur (RC = 1,791; $p = 0,049$). Dans l'ensemble la première et la dernière position dans la liste des auteurs demeurent les meilleurs prédicteurs de l'obtention du statut d'auteur, avec un RC d'un ordre de grandeur plus élevé que celui des autres prédicteurs.

Tableau XXXII. Coefficient du modèle de régression logistique visant à prédire le statut d'inventeur en fonction du rôle, du sexe et de la réputation.

Variables	B	ES	Wald	Ddl	Sig.	RC	Intervalle de confiance à 95 % pour RC	
							Inférieur	Supérieur
premier	1,879	0,411	39,349	1	0,000	6,545	2,923	14,658
dernier	2,285	0,498	21,025	1	0,000	9,826	3,700	26,094
principal	1,007	0,348	8,375	1	0,004	2,737	1,384	5,412
superviseur	0,297	0,391	0,579	1	0,447	1,346	0,626	2,898
homme	0,583	0,295	3,891	1	0,049	1,791	1,004	3,196
hautement_cité	0,310	0,266	1,361	1	0,243	1,363	0,810	2,294
Constante	-2,569	0,316	65,978	1	0,000	0,077		

3.2.2.2.4. Vérification des hypothèses

Le modèle de régression logistique utilisé dans cette partie de l'analyse permet de tester les hypothèses suivantes pour la recherche biomédicale seulement :

HS7 Les auteurs avec un rôle de supervision ont une probabilité plus grande d'être inventeurs que les auteurs principaux et les auteurs périphériques.

Les résultats présentés dans le Tableau XXXII (p. 139) montrent que, pris dans leur ensemble, ce sont les auteurs principaux et non les auteurs superviseurs qui sont nommés inventeurs le plus souvent. Cela peut sembler contradictoire avec le fait que les derniers auteurs (qui sont tous des auteurs superviseurs) ont une plus grande probabilité d'être inventeurs que tout autre auteur. Cependant, notre modèle de régression logistique montre que pour les auteurs superviseurs n'occupant pas la dernière position, la probabilité d'être inventeur n'est pas significativement différente que pour les auteurs périphériques. L'hypothèse HS₇ est donc partiellement acceptée : seulement les auteurs superviseurs occupant la dernière position ont plus de chance d'être inventeurs en recherche biomédicale.

HS₈ Les auteurs principaux ont une probabilité plus grande d'être inventeurs que les auteurs périphériques.

Le modèle de régression logistique (Tableau XXXII, p. 139) montre que les auteurs principaux ont plus de chance d'obtenir le statut d'inventeur que les auteurs périphériques. Les résultats permettent donc d'accepter l'hypothèse HS₇ pour la recherche biomédicale.

HS₁₀ Les hommes ont une probabilité plus grande que les femmes d'être inventeurs.

Selon les coefficients présentés dans le Tableau XXXII (p. 139), les hommes auraient plus de chance que les femmes d'être inventeur lorsque la position dans la liste des auteurs (premier, milieu, dernier), le rôle et la réputation sont contrôlées. La confirmation provisoire de l'hypothèse HS₁₀ (voir section 3.2.1.2.) pour la recherche biomédicale est donc maintenue.

HS₁₁ Plus le nombre de citations d'un auteur est élevé, plus la probabilité qu'il soit inventeur est grande.

Les données analysées n'ont pas permis d'observer une relation statistiquement significative entre le statut d'inventeur et le fait d'être hautement cité (Tableau XXXII, p. 139). Les résultats ne permettent donc pas d'accepter l'hypothèse HS₁₁.

3.2.3. L'obtention du statut d'inventeur en fonction de la contribution, du sexe et de la réputation en recherche biomédicale

Nous analysons dans cette section le sous-ensemble de 895 auteurs des 121 PBA en recherche biomédicale dont les articles comprenaient une déclaration des contributions

respectives de chaque auteur. Nous présentons d'abord quelques résultats descriptifs illustrant la proportion d'auteurs ayant obtenu le statut d'inventeur en fonction de leur contribution, de leur sexe et de leur réputation tout en contrôlant pour leur position dans la liste des auteurs. Nous présentons ensuite les résultats de l'analyse de régression logistique effectuée afin de vérifier les hypothèses de recherche HS₉ à HS₁₁.

3.2.3.1. Résultats descriptifs

La Figure 13 (p. 142) illustre la division des tâches entre les auteurs des articles en lien avec leur position dans la liste des auteurs. Comme pour l'analyse précédente (section 3.2.2), nous divisons les auteurs en trois groupes (« premier », « dernier », « autres »). Nous observons que, dans l'ensemble, la réalisation de la recherche est la contribution la plus fréquente, suivie de la rédaction, de l'analyse, de la conception, de la fourniture de matériel et des autres contributions. Nous remarquons également que la proportion d'auteurs ayant effectué une tâche donnée diffère selon leur position. En effet, pour les derniers auteurs, les contributions les plus fréquentes sont la rédaction, la conception et l'analyse, alors qu'ils participent très rarement à la réalisation de l'étude. Les premiers auteurs participent le plus souvent à la rédaction, tâche suivie de la réalisation, de l'analyse et de la conception. Finalement, les autres auteurs participent surtout à la réalisation, à la rédaction et à l'analyse. En faisant la somme des proportions pour chaque groupe d'auteurs, nous observons que les premiers auteurs effectuent en moyenne un plus grand nombre de tâches différentes (3,5) que les derniers auteurs (2,7) et les autres auteurs (1,9).

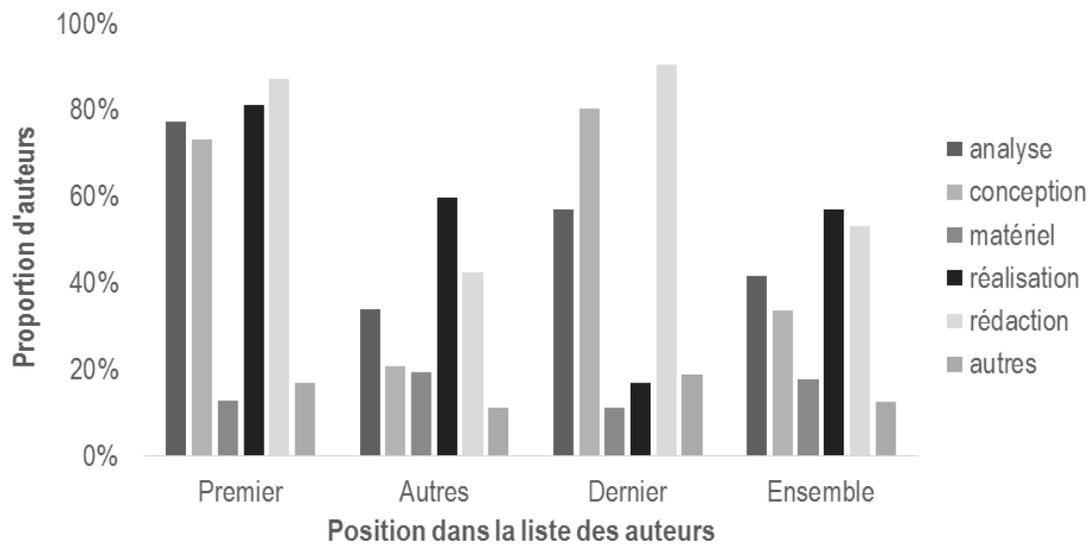


Figure 13. Proportion d’auteurs ayant effectué un type de contribution en fonction de la position dans la liste des auteurs.

La Figure 14 (p. 143) présente la proportion d’auteurs ayant obtenu le statut d’inventeur en fonction de la contribution effectuée et de la position dans la liste des auteurs. Conformément aux analyses présentées aux sections 3.2.1 et 3.2.2, les derniers auteurs sont plus souvent inventeurs que les premiers auteurs pour chaque type de contribution, à l’exception de la fourniture de matériel. Les auteurs n’occupant ni la première ni la dernière position obtiennent beaucoup moins souvent le statut d’inventeur, même lorsqu’ils ont contribué aux tâches qui, dans l’ensemble, semblent mener le plus souvent au statut d’inventeur, soit la conception, l’analyse et à la rédaction.

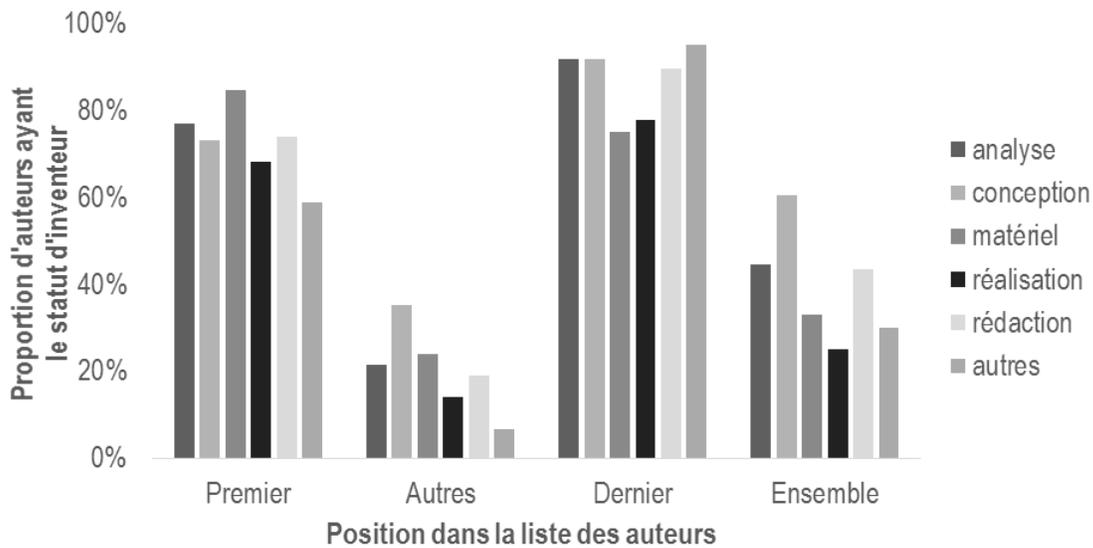


Figure 14. Proportion d’auteurs ayant le statut d’inventeur selon la contribution effectuée et la position dans la liste des auteurs.

La Figure 15 (p. 144) illustre les différences dans la proportion d’auteurs obtenant le statut d’inventeur en fonction du sexe, de la réputation et de la contribution. Nous observons que dans l’ensemble, la conception est la contribution menant le plus souvent au statut d’inventeurs, suivi de l’analyse et de la fourniture de matériel. La réalisation est la contribution qui mène le moins souvent au statut d’inventeurs. De plus, les hommes sont plus souvent inventeurs que les femmes indépendamment de leur niveau de citation et de la contribution, les femmes peu citées ayant participé à la rédaction sont la seule exception. Finalement, les auteurs hautement cités sont plus souvent inventeurs que les auteurs peu cités indépendamment du sexe et de la contribution. Ici, les exceptions sont les femmes ayant contribué à la fourniture de matériel, à la réalisation et à la rédaction, ainsi que les hommes ayant participé à la réalisation de la recherche.

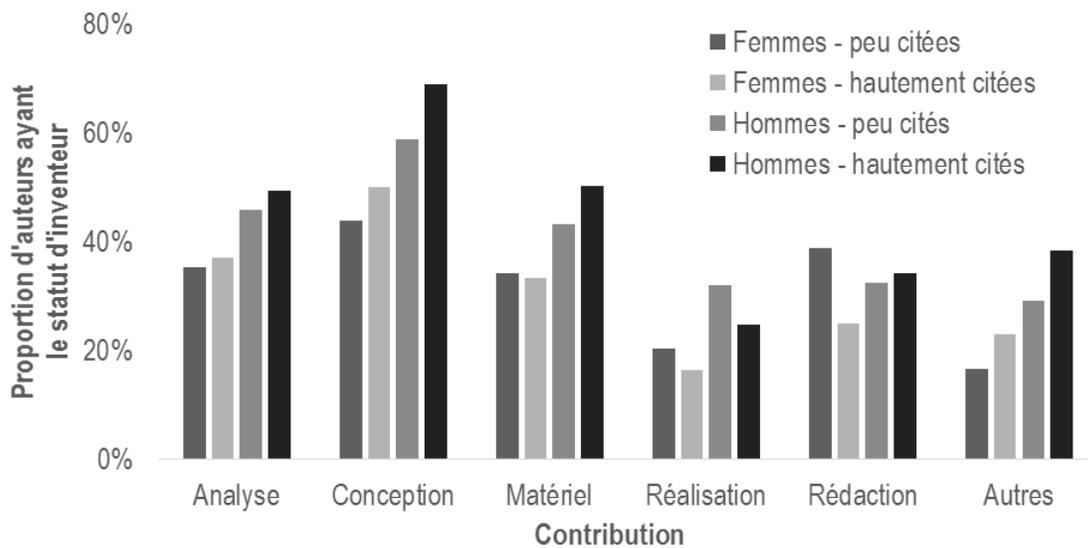


Figure 15. Proportion d'auteurs ayant le statut d'inventeur en fonction du sexe, de la réputation et de la contribution.

3.2.3.2. Régression logistique

À l'aide, encore une fois, d'un modèle de régression logistique, nous évaluons l'ampleur et la significativité statistique des relations entre la contribution, le sexe, la réputation et l'obtention du statut d'inventeur en contrôlant pour la position dans la liste des auteurs. Comme pour l'analyse précédente (section 3.2.2), la variable « position » n'a que trois modalités (« premier », « dernier » et « autres »), afin d'éviter le surajustement du modèle.

3.2.3.2.1. Vérification des postulats de la régression logistique

Tout comme dans les modèles de régression des sections 3.2.1.2 et 3.2.2.2, le postulat de linéarité ne s'applique pas à notre modèle et l'indépendance des observations est assurée par la conception de notre étude. Aucune valeur extrême n'a été identifiée. Les coefficients de corrélation de Pearson présentés à l'Annexe 6 (p. xxxi), tous inférieurs à 0,8, confirment l'absence de multicolinéarité.

3.2.3.2.2. Qualité de l'ajustement et capacité prédictive du modèle

Comme dans l'analyse précédente (section 3.2.2.2), nous débutons le modèle avec la position dans la liste des auteurs (la modalité « autres » est encore une fois choisie comme

modalité de référence), puis nous ajoutons successivement les types de contribution, le sexe et la réputation. Le

Tableau XXXIII (p. 145) montre que les différences de khi carré (χ^2) sont statistiquement significatives lorsque les contributions ($p < 0,001$) et le sexe ($p = 0,040$) sont ajoutés au modèle. Ces deux variables augmentent respectivement le pseudo- R^2 de Nagelkerke (R_N^2) de 0,085 et 0,004. Cependant, l'inclusion de la variable « hautement_cité » n'améliore pas significativement la qualité de l'ajustement du modèle et la capacité prédictive du modèle ($p = 0,203$).

Tableau XXXIII. Qualité de l'ajustement et capacité prédictive du modèle de régression logistique hiérarchique visant à prédire l'obtention du statut d'inventeur en fonction de la contribution, du sexe et de la réputation, et contrôlant pour la position dans la liste des auteurs.

Modèle	Qualité de l'ajustement			Différence avec le modèle précédent			Capacité prédictive du modèle	
	χ^2	Ddl	Sig.	$\Delta \chi^2$	Δ ddl	Sig. $\Delta \chi^2$	-2LV	R_N^2
1 ^a	317,668	2	0,000	317,668	2	0,000	769,783	0,425
2 ^b	397,112	8	0,000	79,444	6	0,000	690,339	0,510
3 ^c	401,312	9	0,000	4,200	1	0,040	686,139	0,514
4 ^d	402,932	10	0,000	1,620	1	0,203	684,519	0,515

a. Prédicteurs : (Constante), premier, dernier

b. Prédicteurs : (Constante), premier, dernier, analyse, conception, matériel, réalisation, rédaction, autres

c. Prédicteurs : (Constante), premier, dernier, analyse, conception, matériel, réalisation, rédaction, autres, homme

d. Prédicteurs : (Constante), premier, dernier, analyse, conception, matériel, réalisation, rédaction, autres, homme, hautement_cité

3.2.3.2.3. Coefficients et rapports de cotes des prédicteurs

Le Tableau XXXIV (p. 146) présente les coefficients des variables prédictives du modèle global. Nous remarquons que, *ceteris paribus*, quatre types de contribution sont positivement et significativement liés à l'obtention du statut d'inventeur : l'analyse (RC = 1,905; $p = 0,002$), la conception (RC = 3,291; $p < 0,001$), la fourniture de matériel (RC = 2,423; $p = 0,001$) et la rédaction (RC = 1,648; $p = 0,023$). Nous observons une relation entre le sexe et l'obtention du statut d'inventeur semblable à celle observée dans les analyses précédentes, les hommes étant plus souvent inventeurs que les femmes (RC = 1,603; $p = 0,033$).

Le modèle ne permet pas d'observer une relation statistiquement significative entre le fait d'être hautement cité et l'obtention du statut d'inventeur (RC = 1,131; p = 0,670). Les ordres de grandeur des relations observées sont similaires aux autres modèles : la première et la dernière position dans la liste des auteurs sont les meilleurs prédicteurs du statut d'inventeur, avec un RC d'un ordre de grandeur plus élevé que celui des autres prédicteurs.

Tableau XXXIV. Coefficient du modèle de régression logistique hiérarchique visant à prédire l'obtention du statut d'inventeur en fonction de la contribution, du sexe et de la réputation, et contrôlant pour la position dans la liste des auteurs.

Variables	B	ES	Wald	ddl	Sig.	RC	Intervalle de confiance à 95 % pour RC	
							Inférieur	Supérieur
premier	1,789	0,287	38,849	1	0,000	5,986	3,410	10,507
dernier	3,348	0,369	82,387	1	0,000	28,450	13,807	58,623
analyse	0,644	0,211	9,306	1	0,002	1,905	1,259	2,882
conception	1,191	0,212	31,605	1	0,000	3,291	2,173	4,985
matériel	0,885	0,257	11,863	1	0,001	2,423	1,464	4,008
réalisation	0,051	0,222	0,052	1	0,820	1,052	0,681	1,626
rédaction	0,500	0,220	5,149	1	0,023	1,648	1,070	2,537
autres	-0,242	0,318	0,578	1	0,447	0,785	0,421	1,465
homme	0,472	0,222	4,536	1	0,033	1,603	1,038	2,476
hautement_cité	-0,278	0,219	1,611	1	0,204	0,757	0,493	1,163
Constante	-3,021	0,319	89,772	1	0,000	0,049		

3.2.3.2.4. Vérification des hypothèses

Le modèle de régression logistique utilisé dans cette partie de l'analyse permet de tester les hypothèses suivantes pour la recherche biomédicale seulement :

HS₉ Les auteurs ayant contribué à la conception de l'étude ont une probabilité plus grande d'être inventeurs.

Les résultats présentés dans le Tableau XXXIV (p. 146) permettent de confirmer que la contribution à la conception d'une étude constitue le principal prédicteur de l'obtention du statut d'inventeur. Cependant, nous observons que trois autres types de contribution constituent également des prédicteurs significatifs, mais plus faibles, de l'obtention du statut d'inventeur :

l'analyse, la fourniture de matériel et la rédaction. Nous sommes donc en mesure d'accepter l'hypothèse de recherche HS₉ pour la recherche biomédicale, la conception étant le principal prédicteur du statut d'inventeur, mais notons au passage la relation positive et significative entre le statut d'auteur et la fourniture de matériel, l'analyse et la rédaction.

HS₁₀ Les hommes ont une probabilité plus grande que les femmes d'être inventeurs.

Les résultats de notre analyse (Tableau XXXIV, p. 146) montrent que les hommes ont significativement plus de chance que les femmes d'être inventeurs lorsque la position dans la liste des auteurs, la contribution et la réputation sont prises en compte. Cela permet de valider l'hypothèse HS₁₀ pour la recherche biomédicale.

Dans l'ensemble, les trois analyses de la section 3.2 permettent une confirmation partielle de l'hypothèse HS₁₀. En effet, les résultats montrent que les hommes ont une probabilité plus grande d'être inventeurs que les femmes en sciences naturelles (section 3.2.1) et en recherche biomédicale (sections 3.2.1 à 3.2.3), mais pas en génie (section 3.2.1).

HS₁₁ Plus le nombre de citations d'un auteur est élevé, plus la probabilité qu'il soit inventeur est grande.

Les résultats présentés dans le Tableau XXXIV (p. 146) ne permettent pas d'accepter l'hypothèse HS₁₁. Lorsque la position dans la liste des auteurs, la contribution et le sexe sont pris en compte, les auteurs hautement cités n'ont pas significativement plus ou moins de chances d'être nommés inventeurs que les autres auteurs en recherche biomédicale.

Dans l'ensemble, les trois analyses de la section 3.2 permettent d'accepter partiellement l'hypothèse HS₁₁. En effet, la probabilité d'obtenir le statut d'inventeur augmente significativement pour les auteurs de toutes les disciplines lorsque l'on tient compte de la position dans la liste des auteurs et du sexe (section 3.2.1). Les coefficients de régression ne sont cependant pas significatifs en recherche biomédicale lorsque nous contrôlons en plus pour le rôle (section 3.2.2) ou la contribution (section 3.2.3).

3.2.4. En résumé

Nous avons, dans cette seconde section du Chapitre 3, présenté un ensemble de résultats descriptifs ainsi que trois analyses utilisant la régression logistique pour prédire l'obtention du

statut d'inventeur à partir d'ensembles de variables prédictives pouvant être divisées en deux groupes : les prédicteurs liés à la nature ou à l'ampleur du travail effectué par les auteurs (position dans la liste des auteurs, rôle et contribution) et les prédicteurs liés aux caractéristiques individuelles de ces derniers (le sexe et la réputation).

La première analyse a montré que la position dans la liste des auteurs est un prédicteur fiable du statut d'inventeur, et que plus l'on s'éloigne de la première ou de la dernière position, plus la probabilité d'être inventeur diminue. Cette analyse a aussi permis d'observer que, alors que les derniers auteurs obtiennent plus souvent le statut d'inventeur que les premiers auteurs en recherche biomédicale, la tendance inverse prévaut en génie et en sciences naturelles, où les premiers auteurs sont plus souvent inventeurs que les derniers auteurs. Le deuxième ensemble de résultats a montré que parmi les auteurs n'occupant pas la première ou dernière position en recherche biomédicale, la probabilité d'être inventeur est moindre pour les auteurs faisant partie d'une séquence d'auteurs en ordre alphabétique (les auteurs périphériques). La troisième analyse a montré qu'au-delà de la position dans la liste des auteurs, quatre types de contribution sont liés positivement et significativement au statut d'inventeur en recherche biomédicale : la conception, l'analyse, la fourniture de matériel et la rédaction.

Dans chacune de ces trois analyses, nous avons également testé la relation entre le sexe, la réputation et le statut d'inventeur. Les trois analyses convergent en ce qui concerne l'effet du sexe : les hommes sont en moyenne plus souvent inventeurs que les femmes lorsque nous contrôlons pour la position dans la liste des auteurs, le rôle, la contribution et la réputation. Seule exception, les résultats ne permettent pas de conclure qu'il existe un lien significatif entre le sexe et le statut d'inventeur en génie. En ce qui concerne la réputation comme prédicteur du statut d'inventeur, les résultats sont moins concluants. Bien qu'être hautement cité soit positivement et significativement lié au statut d'inventeur selon les modèles de la section 3.2.1, le seuil de significativité statistique n'a pas été atteint dans les deux autres analyses portant sur la recherche biomédicale seulement (section 3.2.2 et 3.2.3).

Conclusion

Dans la première section de ce chapitre, nous avons modélisé les liens entre certaines caractéristiques du travail de recherche (discipline et facteur d'impact de la revue de

publication), certaines caractéristiques de l'équipe de recherche (nombre d'auteurs, nombre d'institutions impliquées) et le ratio inventeurs/auteurs à partir d'un ensemble de 6 301 PBA. Cette analyse, liée aux deux premiers objectifs de la thèse, a permis d'apporter des éléments de réponses aux questions de recherche QR1 à QR5 qui seront davantage développés dans le prochain chapitre. Notamment, nous avons établi l'existence de différences disciplinaires dans la relation entre le statut d'auteur et le statut d'inventeur. Nous avons également observé que, dans l'ensemble, la relation entre le nombre d'auteurs et d'inventeurs d'une PBA est influencée par certaines caractéristiques de la recherche (la collaboration interinstitutionnelle) ainsi que par le facteur d'impact de la revue de publication, bien que ces relations diffèrent également d'une discipline à l'autre.

Les auteurs des articles formant les PBA ont été l'unité d'analyse de la seconde section de ce chapitre. Nous y avons présenté des modèles de régression logistique visant à prédire l'attribution du statut d'inventeur aux auteurs en fonction de leur position dans la liste des auteurs, de leur rôle, de leur contribution, de leur sexe et de leur réputation. Les résultats de cette analyse ont également mis en lumière certaines différences disciplinaires dans la relation entre les statuts d'auteur et d'inventeur. Bien que nous observions un lien fort entre le fait d'être premier ou dernier auteur d'un article et l'obtention du statut d'inventeur sur le brevet correspondant, nos résultats démontrent que l'octroi du statut d'inventeur est également guidé par d'autres facteurs, en particulier le rôle et le type de contribution effectué. Nous observons également que le sexe et, dans une moindre mesure, la réputation sont aussi liés au statut d'inventeur.

Dans le Tableau XXXV (p. 150) nous présentons les hypothèses de recherche et indiquons si les résultats obtenus ont permis de les accepter (complètement ou partiellement) ou non.

Tableau XXXV. Tableau récapitulatif des résultats des hypothèses.

	Hypothèse	Résultats obtenus
HS₁	Le ratio inventeurs/auteurs diffère entre les disciplines.	Acceptée
HS₂	Plus le facteur d'impact de la revue de publication de l'article formant une PBA est élevé, plus le ratio inventeurs/auteurs sera élevé.	Partiellement acceptée
HS₃	Plus le nombre d'auteurs d'une PBA est élevé, plus le ratio inventeurs/auteurs sera faible.	Acceptée
HS₄	Les PBA issues d'une collaboration interinstitutionnelle auront un ratio inventeurs/auteurs plus faible que les PBA dont tous les auteurs sont affiliés à la même institution.	Partiellement acceptée
HS₅	Plus un auteur est près de la première ou dernière position dans la liste des auteurs de l'article, plus la probabilité qu'il soit inventeur est grande.	Acceptée
HS₆	La relation entre la position dans la liste des auteurs la probabilité d'obtenir le statut d'inventeur diffère en fonction de la discipline.	Acceptée
HS₇	Les auteurs avec un rôle de supervision ont une probabilité plus grande d'être inventeurs que les auteurs principaux et les auteurs périphériques.	Partiellement acceptée
HS₈	Les auteurs principaux ont une probabilité plus grande d'être inventeurs que les auteurs périphériques.	Acceptée
HS₉	Les auteurs ayant contribué à la conception de l'étude ont une probabilité plus grande d'être inventeurs.	Acceptée
HS₁₀	Plus le nombre de citations d'un auteur est élevé, plus la probabilité qu'il soit inventeur est grande.	Partiellement acceptée
HS₁₁	Les hommes ont une probabilité plus grande que les femmes d'être inventeurs.	Partiellement acceptée

Chapitre 4. Discussion

Introduction

Dans ce chapitre, nous discutons des résultats présentés au Chapitre 3 en faisant des liens avec notre cadre de référence (et notre revue de la littérature) présenté au Chapitre 1. Le chapitre comporte trois sections. La première porte sur la relation entre le nombre d'auteurs et d'inventeurs des PBA et couvre ainsi les trois premiers niveaux du cadre de référence que nous avons présentés dans le Tableau IV (p. 66), soit la discipline, la recherche, et l'équipe de recherche. La deuxième section du présent chapitre porte sur le quatrième niveau (le niveau individuel). Nous présentons les limites de la recherche dans la troisième section.

4.1. Le ratio inventeurs/auteurs dans les équipes de recherche

Dans cette première section, nous discutons des résultats obtenus lors de l'analyse de la relation entre le ratio inventeurs/auteurs des PBA et la discipline, la recherche et l'équipe de recherche.

4.1.1. Enjeux, structure, culture et normes de la discipline

Nous avons montré dans le Chapitre 3 que la relation entre le nombre d'auteurs et le nombre d'inventeurs des PBA diffère d'une discipline à l'autre. En effet, la proportion des auteurs qui sont aussi inventeurs est la plus élevée en génie et la plus faible en recherche biomédicale. Nous pouvons dégager trois hypothèses explicatives de notre cadre de référence : 1) les pratiques d'attribution du statut d'auteur diffèrent entre les disciplines; 2) les pratiques d'attribution du statut d'inventeur diffèrent entre les disciplines; 3) les pratiques de collaboration et de division du travail diffèrent entre les disciplines.

La première hypothèse est qu'il existe des différences disciplinaires dans les pratiques d'attribution du statut d'auteur faisant en sorte qu'à taille d'équipe et division du travail égales, les chercheurs de différentes disciplines auront tendance à attribuer le statut d'auteur différemment. Par exemple, le fait que le ratio inventeurs/auteurs moyen soit plus faible en

recherche biomédicale pourrait indiquer que les individus ayant effectué des tâches de moindre importance sont plus souvent nommés auteurs dans cette discipline. Cette hypothèse est appuyée par les études selon lesquelles les critères d'attribution du statut d'auteurs varient en fonction de la discipline (Bozeman et Youtie, 2016; Cronin, 2001; Knorr Cetina, 1999; Paul-Hus et al. 2017; Pontille 2004, 2006). Paul-Hus et al. (2017) ont illustré ces différences disciplinaires en montrant que certaines disciplines sont caractérisées par un faible nombre moyen d'auteurs et un grand nombre moyen de remerciés. Par contre, selon leur étude, ce n'est pas le cas dans les disciplines considérées dans notre étude, qui ont en moyenne moins d'un remercié par article. Ainsi, dans la mesure où l'on peut s'attendre à ce que les individus ayant contribué à une recherche sans obtenir le statut d'auteur aient tendance à être remerciés dans l'article, on peut douter que les différences disciplinaires observées dans la proportion d'auteurs qui obtiennent le statut d'inventeur soient causées uniquement par des différences disciplinaires au niveau de l'attribution du statut d'auteur.

La deuxième hypothèse avancée est qu'il existe des différences dans les pratiques d'attribution du statut d'inventeur. Par exemple, le ratio inventeurs/auteurs moins élevé en recherche biomédicale qu'en génie indiquerait qu'à contribution égale, les auteurs d'un article seraient plus souvent exclus de la liste des inventeurs du brevet dans la première discipline que dans la seconde. Cette deuxième hypothèse semble cependant moins plausible que la première étant donné les règles plus strictes qui contraignent l'attribution du statut d'inventeur^{43, 44, 45} et font en sorte que le statut d'inventeur revient de droit à certains individus et non à d'autres (Sheiness et Canady, 2006). Ainsi, dans la mesure où la prémisse que le statut d'inventeur est attribué de façon aussi stricte dans chaque discipline est valide, nous pouvons mettre en doute l'importance de la variation dans les pratiques d'attribution du statut d'inventeur en tant que principale cause des différences disciplinaires observées dans le ratio inventeurs/auteurs.

La troisième hypothèse est que les objets de recherche caractérisant les disciplines mènent au développement de cultures épistémiques distinctes, favorisant des modes d'organisation du travail particuliers (Knorr Cetina, 1999). L'exemple de la physique des

⁴³ *Hardee, In re*, 223 USPQ 1122 (Comm'r Pat 1984).

⁴⁴ *Smernoff, Ex parte*, 215 USPQ 545 (Bd App 1982).

⁴⁵ *Burroughs Wellcome Co v Barr Labs Inc*, 40 F.3d 1223, 32 USPQ2d 1915 (Fed Cir 1994).

particules, bien que cette discipline ne soit pas considérée dans notre étude, constitue un exemple éloquent de cette interprétation. La recherche dans ce domaine nécessite l'accès à des équipements extrêmement sophistiqués et coûteux (Katz et Martin, 1997) qui sont partagés souvent par des centaines, voire des milliers, de chercheurs dont le nom apparaîtra sur les articles issus des expérimentations (Biagioli, 2003; Birnholtz, 2006; Cronin, 2001; Pontille, 2016). À l'opposé, dans les sciences humaines et sociales, la recherche ne nécessite généralement pas d'équipements coûteux et les chercheurs peuvent mener leurs recherches individuellement s'ils le désirent (Cronin, McKenzie et Rubio, 1993). Dans cette perspective, nos résultats suggèrent que les objets d'étude en génie favoriseraient un mode collégial de collaboration où les contributions menant à l'obtention du statut d'inventeur sont partagées par un plus grand nombre d'individus—que Laudel (2002) appelle les collaborations avec division du travail. À l'inverse, les objets d'étude en recherche biomédicale favoriseraient davantage les grandes équipes, les collaborations de service (Laudel, 2002) et un mode d'organisation bureaucratique du travail (Walsh et Lee, 2015) diminuant la proportion d'auteurs qui sont inventeurs. Selon la perspective mertonienne, ces modes de travail seraient guidés par l'objectif de la discipline et d'un ensemble de normes qui en découlent (Merton, 1942). Les chercheurs choisiraient donc de former des équipes d'une certaine taille et de diviser le travail d'une certaine façon parce que cela favorise l'accomplissement des objectifs de l'institution scientifique.

Dans un cadre bourdieusien, ces mêmes pratiques ne découleraient pas d'objectifs communs et de normes techniques et morales partagées par tous, mais de la structure des champs disciplinaires et des intérêts individuels des dominants (Bourdieu, 1976). Les théories de Merton et Bourdieu sont similaires en ce qu'elles nous permettent de positionner le champ (ou l'institution) comme force structurante des actions. Comme nous l'avons vu à la section 1.1.3, leurs divergences proviennent entre autres du fait que pour Bourdieu, les intérêts du champ, ce sont les intérêts des dominants. Cela dit, l'habitus fait en sorte qu'il existe une certaine régularité dans les actions des agents du champ (Bourdieu, 1986). Il existe ainsi possiblement des régularités différentes d'un champ disciplinaire à l'autre, ce qui vient appuyer notre troisième hypothèse. Bourdieu insiste également sur le pragmatisme des agents, dont les actions sont aussi guidées par leur sens pratique : « Il faut se garder de chercher dans les productions de l'habitus plus de logique qu'il n'y en a : la logique de la pratique, c'est d'être logique jusqu'au point où

être logique cesserait d'être pratique. » (Bourdieu, 1986, p. 41) Ainsi, les pratiques observées sont le fruit de la structure du champ, de l'intérêt des dominants, et du sens pratique des chercheurs en quête de capital symbolique.

Que l'on adopte une perspective de Merton ou de Bourdieu, nos résultats montrent que la proportion d'auteurs obtenant le statut d'inventeur dans les PBA diffère d'une discipline à l'autre. Notre cadre théorique et notre revue de la littérature suggèrent que ces résultats reflètent en partie des différences disciplinaires dans les pratiques d'attribution du statut d'auteur, mais aussi dans les pratiques de collaboration et de division du travail. En montrant que le travail inventif est le plus partagé en génie et le plus concentré en recherche biomédicale, nous contribuons à mettre en lumière des normes, des intérêts et des stratégies « qui n'existaient qu'à l'état incorporé, sous forme de dispositions, de schèmes classificatoires dont les produits sont cohérents, mais d'une cohérence partielle. » (Bourdieu, 1986, p. 41)

4.1.2. Impact potentiel de la recherche

Nous avons également considéré la relation entre le facteur d'impact de la revue et le ratio inventeurs/auteurs, suivant les travaux de Jabbehdari et Walsh (2017); ces derniers ayant montré que les articles publiés dans des revues à haut facteur d'impact ont tendance à inclure moins d'auteurs ayant effectué une seule tâche. Les résultats de notre étude montrent que la proportion d'auteurs qui sont aussi inventeurs tend à être plus élevée lorsque l'article est publié dans une revue à haut facteur d'impact. Nous proposons deux hypothèses explicatives. La première hypothèse est que les auteurs dont les contributions ont été les moins importantes sont plus souvent exclus de la liste des auteurs lorsque le gain potentiel en capital symbolique est plus élevé. La seconde hypothèse est que les articles publiés dans les revues à haut facteur d'impact ont tendance à être le fruit de collaboration basée sur une plus grande division du travail inventif.

La première hypothèse est en accord avec l'interprétation qu'ont fait Jabbehdari et Walsh (2017) de leurs résultats. En effet, selon eux, le gain en capital symbolique potentiellement plus élevé pour ces articles réduirait l'inclinaison des chercheurs à attribuer le statut d'auteur aussi « généreusement » que lorsqu'il s'agit d'articles publiés dans des revues ayant une moins grande visibilité. Les auteurs se retrouvant exclus seront en toute probabilité ceux dont les contributions

ont été relativement mineures. Jabbehdari et Walsh (2017) appellent ces derniers des « auteurs spécialistes », mais un parallèle peut être fait avec les techniciens (Hagstrom, 1964), les auteurs soldats (Perneger et al., 2017), contributeurs (Baerlocher et al., 2007) ou périphériques (Mongeon et al., 2017), ou encore les individus participant à la recherche dans le cadre d'une collaboration de service (Laudel, 2002). Le gain potentiel en capital symbolique n'aurait cependant pas le même effet sur l'attribution du statut d'inventeur, qui suit en principe des règles plus strictes et qui n'ont rien à voir avec l'impact scientifique potentiel de la recherche.

Dans notre thèse, les articles dans le domaine du génie font exception; les résultats ne montrant aucune relation significative entre le fait qu'un article soit publié dans une revue à haut facteur d'impact et le ratio inventeurs/auteurs de la PBA. Cette particularité des PBA en génie est possiblement due au fait que, tel que mentionné à la section 4.1.1, les PBA y sont caractérisées par un ratio inventeurs/auteurs plus élevé que dans les autres disciplines. Avec en moyenne 78,7 % des auteurs des articles en génie ayant obtenu le statut d'inventeur, les possibilités d'exclure de la liste d'auteurs des contributeurs dont la contribution a été purement technique ou peu significative sont plus limitées que dans une discipline comme la recherche biomédicale où nos résultats suggèrent qu'en moyenne moins de la moitié des auteurs (49 %) ont obtenu le statut d'inventeur.

Nous émettons l'hypothèse alternative que ce sont les collaborations caractérisées par un plus grand partage du travail inventif (et donc une proportion plus élevée d'auteurs qui sont aussi inventeurs) qui mènent à des découvertes plus importantes et pour lesquelles les gains potentiels en capital symbolique sont donc plus élevés. En d'autres termes, les PBA dont l'article est publié dans une revue à haut facteur d'impact auraient en moyenne un plus grand nombre d'inventeurs et non un moins grand nombre d'auteurs. Notre revue de la littérature ne nous a cependant pas permis de repérer des études antérieures pouvant appuyer ou contredire cette hypothèse. En effet, les études antérieures s'étant intéressées à la relation entre la collaboration et l'impact ou la qualité des articles publiés partagent la même limite : elles mesurent la relation entre la taille des équipes telle que mesurée par le nombre d'auteurs avec divers indicateurs d'impact ou de qualité sans tenir compte de la division du travail.

Si nous adoptons le cadre mertonien selon lequel les chercheurs sont avant tout guidés par les normes institutionnelles (Merton, 1942), la première hypothèse semble moins plausible

que la seconde. En effet, une stricte adhésion des chercheurs aux normes rendrait peu probable que la nomination des auteurs soit influencée par les caractéristiques de l'article. De plus, l'objectif de la science étant l'avancement des connaissances, nos résultats pourraient être interprétés comme une démonstration que la collaboration impliquant une division du travail créatif donne effectivement lieu à des articles publiés dans des revues à plus haut FI, plus cités, et donc possiblement à des contributions à l'avancement des connaissances qui sont de plus grande importance (Katz et Martin, 1997).

Les deux hypothèses émises sont plausibles dans la perspective bourdieusienne. En effet, si les individus cherchent à maximiser leurs gains en capital symbolique et à améliorer leur position dans la structure du champ, ils peuvent aussi bien le faire en limitant le nombre d'individus avec lesquels ils partageront les gains en capital symbolique (hypothèse 1) ou en employant une stratégie de collaboration et d'organisation du travail qui favorisera la production de biens symboliques d'une plus grande « valeur » (hypothèse 2).

Les études antérieures ne permettent pas de préférer une hypothèse plutôt qu'une autre. Il est probable que les deux phénomènes aient lieu dans la réalité, mais d'autres études seront requises afin de déterminer leur poids respectif dans les résultats de notre étude.

4.1.3. Les caractéristiques de l'équipe et la concentration du statut d'inventeur

Nous avons proposé à la section 4.1.1 que les équipes de recherche tendent à être plus collégiales en génie (ratio inventeurs/auteurs élevé) et plus hiérarchiques en recherche biomédicale (ratio inventeurs/auteurs faible), les sciences naturelles se situant entre les deux autres disciplines à ce niveau. Dans la présente section, nous discutons non pas des différences disciplinaires, mais plutôt de la relation entre le nombre d'auteurs des articles et le ratio inventeurs/auteurs des PBA d'une même discipline.

Nos résultats montrent que plus le nombre d'auteurs augmente, plus la proportion d'auteurs qui sont aussi inventeurs diminue. Cela suggère que plus les équipes de recherche sont grandes, telles que mesurées par le nombre d'auteurs, plus elles ont tendance à adopter un mode d'organisation bureaucratique du travail caractérisé par une division et une standardisation du

travail, la présence de relations hiérarchiques, et une centralisation des processus décisionnels (Walsh et Lee, 2015). Les grandes équipes de recherche telles que mesurées par le nombre d'auteurs seraient donc plus susceptibles d'inclure des techniciens professionnels (Hagstrom, 1964), des auteurs soldats (Perneger et al., 2017), contributeurs (Baerlocher et al. 2007), ou périphériques (Mongeon et al. 2017). Nos résultats peuvent aussi être interprétés au regard de la typologie de collaboration de Laudel (2002), la structure bureaucratique favorisant les collaborations de service où de nombreux individus sont au service du directeur du laboratoire siégeant au sommet de la hiérarchie.

Nous n'avons trouvé dans la littérature aucune étude portant sur la relation entre le nombre d'auteurs et le nombre ou la proportion des auteurs qui sont aussi inventeurs. Il peut cependant être intéressant de comparer nos résultats à ceux obtenus par Mongeon et al. (2017). En effet, les auteurs ont mesuré la proportion d'auteurs périphériques des articles en recherche biomédicale ayant sept auteurs ou plus. Ils considèrent qu'un auteur est périphérique lorsqu'il apparaît dans une séquence d'auteurs en ordre alphabétique au milieu de l'article. Selon leurs résultats la proportion moyenne d'auteurs périphériques semble indépendante du nombre d'auteurs de l'article, avec en moyenne environ 60 % d'auteurs périphériques. Un peu plus de 20 % sont en moyenne des auteurs principaux apparaissant au début de la liste et un peu moins de 20 % sont des auteurs superviseurs apparaissant à la fin de la liste. À la lumière de ces résultats, nous aurions pu nous attendre à ce que la proportion des auteurs obtenant le statut d'inventeur soit aussi relativement indépendante de la taille de l'équipe, les auteurs principaux et superviseurs ayant possiblement effectué des contributions assez importantes pour obtenir le statut d'inventeur sur le brevet. Cependant, les résultats de Mongeon et al. (2017) contrastent plutôt avec les nôtres : nous observons une décroissance de la proportion d'auteurs qui sont aussi inventeurs lorsque le nombre d'auteurs augmente. Il nous semble plausible que cette divergence illustre la non-équivalence de l'importance d'une contribution individuelle à une recherche et de l'importance d'une contribution individuelle dans le processus d'invention. Cette idée est similaire à celle de Packer et Webster (1996) selon laquelle l'originalité en science n'est pas équivalente au concept d'originalité dans le domaine de la propriété intellectuelle. Il importerait donc de souligner que le fait d'avoir fait une contribution importante à un projet de recherche ne mène pas nécessairement à l'obtention du statut d'inventeur. En effet, comme nous

l'avons vu au Chapitre 1, c'est seulement la conception de l'invention qui, en principe, mène au statut d'inventeur⁴⁶.

En plus de cette spécificité du type de contribution menant au statut d'inventeur, notre revue de la littérature a fait ressortir l'importance du contrôle du processus créatif dans l'attribution du statut d'inventeur. En effet, l'inventeur peut considérer et adopter les idées d'autrui, tant qu'il conserve le contrôle du processus intellectuel de sélection et de rejet des idées ainsi que de leur mise en application pour la résolution du problème⁴⁷. Or, certains membres d'une équipe (p. ex. le directeur de laboratoire) peuvent occuper une position d'autorité qui leur confère un pouvoir à la fois sur la division du travail et sur l'attribution des statuts d'auteur et d'inventeur (Armstrong et Murphy, 2012; Seymore, 2006). Ainsi, il est possible que nos résultats ne reflètent pas seulement la division du travail, mais également une certaine division du contrôle de la recherche. De plus, le potentiel gain en capital économique qu'un brevet peut générer pour un inventeur est réduit par chaque ajout d'un inventeur additionnel. Dans la perspective bourdieusienne où les pratiques sont guidées par les rapports de force et les intérêts individuels, il semble plausible que la taille des équipes augmente sans que le nombre d'individus exerçant un contrôle sur le projet augmente proportionnellement. Nous reviendrons sur le rôle des rapports de force dans la section 4.2.1.

Notre analyse des facteurs pouvant influencer l'attribution des statuts d'auteur et d'inventeur au niveau de l'équipe de recherche a aussi montré que, dans les sciences naturelles seulement, les PBA issues d'une collaboration interinstitutionnelle ont tendance à avoir un ratio inventeurs/auteurs plus faible que les PBA dont tous les auteurs sont affiliés à la même institution. Bien que cette relation soit plutôt faible, elle est néanmoins cohérente avec les résultats de Jabbedhari et Walsh (2017) qui ont montré que les articles issus d'une collaboration interinstitutionnelle ont en moyenne moins de contributeurs exclus de la liste des auteurs. Ce résultat peut toutefois sembler contredire les propos de l'un des participants interviewés par Pontille (2006), selon lequel lorsque deux laboratoires collaborent, le deuxième auteur a souvent fait une contribution importante, et l'avant-dernier est souvent le directeur du deuxième laboratoire. Dans un tel cas, on pourrait effectivement s'attendre à ce que le ratio

⁴⁶ *Hardee, In re*, 223 USPQ 1122 (Comm'r Pat 1984).

⁴⁷ *Morse v Porter*, 155 USPQ 280 (Bd Pat Inter 1965).

inventeurs/auteurs augmente au lieu de diminuer pour les PBA en collaboration interinstitutionnelle. Par contre, il est aussi possible de s'imaginer qu'une collaboration entre deux laboratoires, bien qu'elle puisse mener à un partage du travail de conception entre les directeurs de laboratoires et quelques doctorants (ce qui se traduira par un nombre plus élevé d'inventeurs), mène également à la juxtaposition de deux structures bureaucratiques distinctes (Walsh et Lee, 2015) comprenant chacune une majorité d'auteurs périphériques (contributeurs, techniciens). L'augmentation du nombre d'auteurs qui obtiendront également le statut d'inventeur sera dans ce cas accompagnée d'une augmentation encore plus importante du nombre d'auteurs qui ne l'obtiendront pas, ce qui réduira le ratio inventeurs/auteurs de la PBA. Il se peut également qu'une équipe de recherche fasse appel à un chercheur d'une autre institution dans le cadre d'une collaboration de service (Laudel, 2002) et que le chercheur en question obtienne le statut d'auteur (mais non le statut d'inventeur) en échange pour ce service rendu (Pontille, 2016; Shibayama, Walsh et Baba, 2012). Dans un tel cas, la collaboration interinstitutionnelle sera associée à un plus grand nombre d'auteurs seulement, ce qui aura pour effet une diminution du ratio inventeurs/auteurs de la PBA.

4.2. Qui est inventeur?

Cette section porte sur le quatrième niveau de notre cadre de référence, soit le niveau individuel. Nous discutons de la relation entre le statut d'inventeur et 1) les facteurs liés au type de contribution effectuée, 2) à la réputation et 3) au sexe.

4.2.1. Entre contribution et rapports de force

Nous avons analysé au Chapitre 3 la relation entre l'obtention du statut d'inventeur et 1) la position dans la liste des auteurs, 2) le rôle et 3) la nature de la contribution effectuée. Nous reprenons ici chacun de ces prédicteurs et discutons des résultats obtenus et de leurs liens avec les études antérieures.

Nos résultats montrent que les inventeurs se retrouvent généralement au début et à la fin de la liste des auteurs des articles, apportant ainsi de nouvelles preuves empiriques de l'importance de la première et de la dernière position dans les domaines de recherche où la culture épistémique privilégie le laboratoire comme mode d'organisation du travail de recherche

(Pontille, 2004, 2016). Mongeon et al. (2017) ont utilisé l'ordre alphabétique partiel pour distinguer les auteurs périphériques (aussi appelés auteurs contributeurs, soldats ou techniciens) des auteurs principaux et superviseurs. Nous avons utilisé cette méthode pour distinguer dans notre jeu de données les auteurs principaux et superviseurs afin de vérifier l'hypothèse selon laquelle, au-delà de la position dans la liste des auteurs, ceux qui apparaissent à l'extérieur de la séquence alphabétique ont une probabilité plus élevée d'obtenir le statut d'inventeur. Cela s'est avéré le cas uniquement pour les auteurs apparaissant au début de la liste, mais non pour les auteurs occupant les positions près de la fin de liste. En effet, à l'exception des derniers auteurs, la probabilité d'obtenir le statut d'inventeur n'est pas significativement plus élevée pour les auteurs superviseurs que pour les auteurs périphériques. Une explication possible est que la dernière position est celle généralement occupée par le directeur du laboratoire (la personne au sommet de la hiérarchie) alors que les premiers auteurs sont des chercheurs juniors (typiquement des doctorants ou des chercheurs postdoctoraux) qui sont responsables du travail expérimental (Pontille, 2004, 2006), et que les listes d'auteurs reflètent ainsi une hiérarchie où une personne dirige et une ou plusieurs personnes exécutent. Cette hypothèse trouve aussi un appui dans les résultats de Mongeon et al. (2017) selon lesquels les articles ont en moyenne plus d'auteurs principaux (au début de la liste) que d'auteurs superviseurs (à la fin de la liste).

Les résultats relatifs à la relation entre la position dans la liste des auteurs et l'obtention du statut d'inventeur reflètent les différences disciplinaires au niveau de la division du travail dont nous avons discuté à la section 4.1.1. En effet, le domaine biomédical semble être plus hiérarchique, le dernier auteur et le premier auteur monopolisant le travail inventif. À l'opposé, le travail inventif en génie semble être partagé plus équitablement parmi les membres de l'équipe de recherche, comme en témoigne la répartition plus uniforme des inventeurs parmi les auteurs. L'analyse de la relation entre la position dans la liste des auteurs et l'obtention du statut d'inventeur montre aussi qu'en recherche biomédicale, ce sont les derniers auteurs qui obtiennent le plus souvent le statut d'inventeur, suivi des premiers auteurs et des deuxièmes auteurs. Ces résultats concordent avec ceux de Ducor (2000) et Lissoni et al. (2013). En effet, les derniers auteurs des 40 PBA en biologie moléculaire analysées par Ducor (2000) étaient plus nombreux (37/40) à avoir obtenu le statut d'inventeur que les premiers auteurs (26/40), alors que 49 % des premiers auteurs et 57 % des derniers auteurs des PBA analysées par

Lissoni et al. (2013) avaient obtenu le statut d'inventeur. Nous observons cependant la tendance inverse en génie où ce sont les premiers auteurs qui obtiennent le plus souvent le statut d'inventeur, suivis des deuxièmes et des derniers auteurs. Plusieurs hypothèses peuvent être mises de l'avant pour expliquer cette particularité des PBA en génie. Par exemple, il est possible que la recherche en génie soit effectuée dans des conditions différentes que la recherche en sciences naturelles ou en recherche biomédicale. Par exemple, la recherche en génie logiciel peut plus facilement être effectuée par un seul individu ou un petit groupe d'individus, sans nécessiter l'accès à un appareillage expérimental complexe comme cela peut être le cas en sciences naturelles ou en recherche biomédicale. Il est aussi possible que la norme voulant que la personne au sommet de la hiérarchie occupe la dernière position de la liste des auteurs ne soit pas aussi courante en génie que dans les autres disciplines.

Selon notre analyse, la conception est la contribution la plus fortement liée à l'obtention du statut d'inventeur. Cependant, ce lien n'est pas systématique. En effet, parmi les auteurs ayant effectué une contribution à la conception selon la déclaration de contribution, moins de 80 % des premiers auteurs, près de 90 % des derniers auteurs et moins de 40 % des autres auteurs ont obtenu le statut d'inventeur. De plus, d'autres contributions (l'analyse, la fourniture de matériel et la rédaction) sont aussi liées à l'obtention du statut d'inventeur. En fait, seules les contributions à la réalisation de la recherche et les contributions de types « autres » ne sont pas associées à une plus grande probabilité d'obtenir le statut d'inventeur. En ce qui concerne la relation entre la contribution à l'analyse ou à la rédaction et l'obtention du statut d'inventeur, ces résultats peuvent être expliqués par le fait que les principaux contributeurs à une recherche effectuent généralement plus d'un type de contribution (Larivière et al., 2016; Perneger et al., 2017). Bien que les coefficients de corrélation entre les variables incluses dans le modèle de régression logistique de la section 3.2.3 présentés à l'Annexe 7 (p. xxxi) ne posent pas de problème de multicollinéarité, nous observons une certaine corrélation entre les variables analyse, conception, et rédaction. Cependant, la fourniture de matériel est faiblement corrélée aux autres types de contribution (voir l'Annexe 7, p. xxxi), ce qui suggère que la fourniture de matériel est une contribution qui, par elle-même, est associée à l'obtention du statut d'inventeur.

Une première explication du fait que divers types de contributions sont liés à l'obtention du statut d'inventeur est que, comme l'ont montré Ilakovac et al. (2007), les contributions

autodéclarées sont un reflet plus ou moins fiable de la contribution effectivement effectuée par les auteurs. Cela suggère que nos résultats pourraient en partie être expliqués par un certain nombre de cas où il est indiqué dans la déclaration des contributions qu'un auteur a participé à la conception alors que ce n'est pas le cas en réalité. Une deuxième explication réside dans l'idée qu'une contribution considérée comme originale dans le champ scientifique n'est pas nécessairement une contribution inventive au sens de la loi sur les brevets (Packer et Webster, 1996).

Une autre interprétation possible est que, comme pour l'attribution du statut d'auteur, c'est souvent le directeur du laboratoire qui décide de l'attribution du statut d'inventeur (Armstrong et Murphy, 2012), et que celui-ci peut utiliser les idées des autres membres de l'équipe tant qu'il conserve le contrôle sur les choix qui sont faits⁴⁸. Les rapports de force pourraient faire en sorte que le premier auteur soit exclu de la liste des inventeurs même s'il a effectué une contribution inventive. Il pourrait donc s'agir, dans ce cas, d'une manifestation de l'effet Saint-Matthieu (Merton, 1968) causée par la concentration du pouvoir décisionnel où le directeur de recherche tend à surestimer sa propre contribution et à sous-estimer celle des autres. Il est aussi possible que sans sous-estimer le travail de ses collaborateurs, le chercheur jouissant d'une position d'autorité choisisse de se servir de cette autorité pour s'approprier une plus grande part du capital symbolique et potentiellement économique généré par le brevet. Nous discutons plus en détail de l'effet Saint-Matthieu à la lumière de nos résultats dans la section suivante.

4.2.2. Un effet Saint-Matthieu dans l'attribution du statut d'inventeur?

Dans l'ensemble, les analyses présentées au Chapitre 3 sont peu concluantes en ce qui concerne la relation entre le fait d'être hautement cité et l'obtention du statut d'inventeur. En effet, il semble exister une relation positive et statistiquement significative, mais toutefois relativement faible, entre la réputation et le statut d'inventeur lorsque nous contrôlons pour le sexe et la position dans la liste des auteurs (section 3.2.1). Ces résultats concordent avec ceux de Lissoni et al. (2013) selon lesquels les auteurs peu expérimentés ont une probabilité plus élevée d'être exclus de la liste des inventeurs. Cependant, le seuil de significativité statistique

⁴⁸ *Morse v Porter*, 155 USPQ 280 (Bd Pat Inter 1965).

n'est pas atteint lorsque nous réduisons la taille de l'échantillon et que le rôle (section 3.2.2) ou la contribution (3.2.3) sont contrôlés. Les résultats de ces deux analyses concordent davantage avec ceux de Haeussler et Sauermann (2013) qui n'ont observé aucune relation statistiquement significative entre la réputation et la probabilité d'obtenir le statut d'inventeur. Une première hypothèse explicative est que la réputation d'un auteur agit indirectement sur l'obtention du statut d'inventeur via la position dans la liste des auteurs. En effet, si le chercheur le plus réputé est aussi celui qui détient l'autorité, il est non seulement plus enclin à être le dernier auteur d'un article, mais aussi plus enclin à occuper un rôle principal ou de supervision dans l'équipe de recherche et à effectuer les tâches menant au statut d'inventeur. Cela réduit potentiellement l'effet marginal de la variable « hautement_cité » sur l'obtention du statut d'inventeur, qui devient éventuellement non significatif. Nous pouvons aussi émettre l'hypothèse que contrairement aux brevets, les articles donnent plus de place à l'exercice des rapports de force et facilitent le « don » du statut d'auteur (auteurs honorifiques) à des chercheurs réputés. En effet, l'attribution du statut d'auteur ne fait pas l'objet de règles et de contrôles très stricts, et l'ajout d'un auteur ne réduit à peu près pas le gain potentiel en capital symbolique des autres auteurs (Nudelman et Landers, 1972). À l'inverse, les règles sur l'attribution du statut d'inventeur sont plus strictes, et l'ajout d'un inventeur réduit le gain potentiel en capital économique des autres inventeurs. Il n'y aurait donc que peu d'incitatifs pour les chercheurs à attribuer le statut d'inventeur à des individus autres que ceux ayant réellement effectué une contribution inventive à la recherche. En plus de cette absence d'incitatif, l'attribution du statut d'inventeur est en principe basée seulement sur la nature de la contribution et ne devrait donc pas être influencée par la réputation des chercheurs impliqués dans la recherche (Seymore, 2006). Conjointement, ces deux mécanismes auront pour effet d'augmenter le nombre de chercheurs réputés parmi les auteurs, mais non parmi les inventeurs, affaiblissant ainsi le possible lien entre la réputation et l'obtention du statut d'inventeur.

Dans la mesure où ils suggèrent l'existence d'une relation, sous certaines conditions, entre la réputation d'un chercheur et l'accès au statut d'inventeur, nos résultats appuient l'idée selon laquelle le capital scientifique peut être accumulé et éventuellement reconverti en une autre forme de capital (Bourdieu, 1975). Nos résultats suggèrent que ce mécanisme passe en partie par la structure stratifiée de la science qui détermine l'accès aux ressources et aux

opportunités (Merton, 1968). L'opportunité dans ce cas-ci est l'obtention du statut d'inventeur et le gain potentiel en capital symbolique et économique qui l'accompagne.

4.2.3. Inégalités femmes-hommes dans l'attribution du statut d'inventeur

L'effet du sexe est plus constant dans nos résultats; ceux-ci montrant que la probabilité d'obtenir le statut d'inventeur est généralement plus faible pour les femmes que pour les hommes. Nos analyses ont en effet permis d'observer une relation statistiquement significative entre le sexe et la probabilité d'obtenir le statut d'inventeur en sciences naturelles et en recherche biomédicale, les femmes qui figurent sur la liste des auteurs des articles obtenant moins souvent le statut d'inventeur que les hommes. Ces résultats sont en accord avec ceux de Lissoni et al. (2013), mais contrastent avec ceux de Haeussler et Sauermann (2013) qui n'ont pas observé de relation significative entre le sexe et l'obtention du statut d'inventeur.

Lorsque nous contrôlons seulement la position dans la liste des auteurs et la réputation (section 3.2.1), nos résultats supportent l'hypothèse selon laquelle les contributions des femmes en science tendent à être sous-estimées comparativement à celles des hommes (Rossiter, 1993). Une relation similaire, mais de moindre ampleur, entre le sexe et l'obtention du statut d'inventeur a été observée dans les modèles tenant compte du rôle (section 3.2.2) et des contributions autodéclarées (section 3.2.3) des auteurs. Cela suggère que la relation entre le sexe et l'obtention du statut d'inventeur observée dans notre étude et dans celle de Lissoni et al. (2013) est peut-être en partie expliquée par le fait que les femmes ont tendance à effectuer des tâches différentes de celles effectuées par les hommes. En effet, Macaluso et al. (2016) ont montré que les femmes contribuent moins souvent que les hommes à la conception, à l'analyse, à la rédaction et à la fourniture de matériel que les hommes, alors qu'elles sont largement surreprésentées parmi les auteurs ayant participé à la réalisation de l'étude. Or, nos résultats montrent que la réalisation de la recherche ne mène généralement pas à l'obtention du statut d'inventeur.

Notre devis corrélationnel prédictif ne nous permet pas de nous prononcer sur l'existence d'une relation de cause à effet et de conclure que ce que nous observons est une discrimination envers les femmes. Ils renforcent néanmoins l'hypothèse de l'existence d'un effet Matilda en ce qui concerne l'attribution du statut d'inventeur.

4.3. Limites de la recherche

La principale limite de notre recherche est intrinsèque aux méthodes quantitatives. Bien qu'un devis quantitatif permette l'analyse d'un grand nombre d'observations, et donc l'identification de tendances potentiellement généralisables à l'ensemble d'une population, il ne permet pas d'analyser un phénomène avec autant de profondeur qu'un devis qualitatif pour ce qui est de comprendre, en contexte, comment le phénomène se produit. Par exemple, nous considérons dans notre étude que l'ensemble des articles d'une même discipline, qui ont le même nombre d'auteurs et qui sont publiés dans des revues dont le facteur est semblable constituent un ensemble d'observations comparables pour l'analyse des pratiques de division du travail et de la relation entre le statut d'auteur et d'inventeurs. Cependant, l'analyse qualitative d'un petit nombre d'articles incluant des entrevues avec les chercheurs impliqués, par exemple, permettrait de mettre en lumière la particularité des contextes de travail, des dynamiques d'équipe et des prises de décisions concernant la division des tâches et l'attribution des statuts d'auteurs et d'inventeur. La généralisabilité des résultats de notre analyse quantitative se fait donc au détriment d'une analyse approfondie des contextes particuliers dans lesquels les phénomènes observés se produisent.

En plus du fait qu'en mettant l'accent sur les grandes tendances, nous simplifions les phénomènes observés en leur soutirant une part de leurs particularités, l'usage de données bibliométriques fait en sorte que les phénomènes sont observés de façon indirecte. Ainsi, dans notre étude, nous n'observons pas directement les processus de collaboration, de division du travail, et d'attribution des statuts d'auteur et d'inventeur, mais plutôt le reflet plus ou moins précis de ces processus que nous donnent les articles et les brevets.

Dans le même ordre d'idées, les données utilisées peuvent mener à une opérationnalisation imparfaite de certains concepts étudiés, notamment la taille de l'équipe de recherche, ainsi que le rôle, la contribution, la réputation et le sexe des chercheurs. Par exemple, nous considérons que la liste des auteurs d'un article est représentative de l'ensemble des individus ayant contribué à la recherche qui y est exposée. Or, il est possible que dans certains cas des individus ayant contribué à la recherche n'obtiennent pas le statut d'auteur, mais soient plutôt remerciés (Cronin et Overfelt, 1994). Comme il existe des différences disciplinaires en

ce qui concerne le rapport entre le nombre moyen d'auteurs et d'individus remerciés dans les articles (Paul-Hus et al., 2017), cela pourrait légèrement avoir un impact sur les différences disciplinaires observées. De plus, la réputation est mesurée par les citations qui sont une mesure imparfaite du capital scientifique que détient un chercheur. Finalement, une contribution originale à la recherche n'est pas nécessairement équivalente à une contribution inventive telle qu'elle est définie dans le domaine de la propriété intellectuelle (Packer et Webster, 1996). Ainsi, le statut d'inventeur est aussi, comme le statut d'auteur, un indicateur imparfait de la nature et de l'ampleur des contributions effectuées par les chercheurs. Cela peut faire en sorte que le ratio inventeurs/auteurs sous-estime la réelle division du travail créatif dans les équipes de recherche.

Le choix d'utiliser l'ordre alphabétique partiel pour distinguer les auteurs principaux et superviseurs des auteurs périphériques s'appuie sur l'idée que les auteurs qui ne font pas partie de la séquence alphabétique sont de ce fait mis en valeur (Zuckerman, 1968; Mongeon et al., 2017). Cependant, aucune étude empirique ne s'est depuis intéressée à la relation entre la division du travail dans les équipes de recherche et l'usage de l'ordre alphabétique partiel. D'autres travaux seront nécessaires afin de mieux comprendre la relation entre la position dans une liste d'auteurs en ordre alphabétique partiel et le type de contribution, et l'ampleur de la contribution et le statut hiérarchique des chercheurs. Aussi, nous déterminons les contributions respectives des auteurs d'un article à partir des déclarations de contribution parfois incluses dans les articles, ces dernières étant un reflet plus ou moins précis du travail ayant effectivement été effectué (Ilakovac et al., 2007). De plus, les déclarations de contribution mentionnent typiquement la nature des contributions effectuées, mais ne permettent pas d'en connaître l'ampleur et l'importance.

Finalement, les données sur le rôle, la contribution et le sexe ne sont pas disponibles pour l'ensemble des observations. Cela a réduit la taille des échantillons utilisés pour les analyses visant à répondre aux questions QR7 et QR8. Ces échantillons réduits ont eu pour effet de limiter ces analyses à une seule discipline (la recherche biomédicale) et donc d'interdire la généralisation des résultats aux autres disciplines.

Conclusion

La discussion a permis de comparer nos résultats au cadre théorique de notre étude ainsi qu'aux études antérieures présentées dans notre revue de la littérature (Chapitre 1), et plus particulièrement les études portant sur la collaboration et la division du travail, les pratiques d'attribution des statuts d'auteur et d'inventeur ainsi que les facteurs associés à une plus ou moins grande probabilité d'obtenir le statut d'inventeur pour les individus ayant participé à une recherche.

Nous avons dans un premier temps interprété et mis en relation avec la littérature les résultats de la première partie de notre étude portant sur les prédicteurs du ratio inventeurs/auteurs des PBA, soit la discipline, le nombre d'auteurs, la collaboration interinstitutionnelle et le facteur d'impact de la revue de publication. Nous avons vu que les résultats obtenus mettent en lumière les différences disciplinaires au niveau des pratiques de collaboration et de division du travail.

La discussion a aussi permis de décrire les variables qui sont associées à une plus ou moins grande probabilité d'obtenir le statut d'inventeur. Au-delà de la position dans la liste des auteurs (qui est le plus important prédicteur de l'obtention du statut d'inventeur), les autres prédicteurs sont le rôle et la contribution, la réputation et le sexe. Les résultats obtenus ont notamment permis de confirmer que les auteurs au début et à la fin de la liste sont typiquement ceux ayant effectué les plus importantes contributions et que les contributions à la conception de la recherche sont les plus fortement associées à l'obtention du statut d'auteur. Nos résultats appuient également l'idée selon laquelle il existe un effet Saint-Matthieu et un effet Mathilda dans les pratiques d'attribution du statut d'inventeur.

Finalement, nous avons discuté des limites de notre recherche. Celles-ci sont principalement liées à l'information plus ou moins précise que fournissent l'ordre des auteurs, l'ordre alphabétique partiel et les déclarations de contribution sur les rôles et les contributions réelles des auteurs d'un article. De plus, les données utilisées ne permettent pas de tenir compte du contexte particulier de chacun des cas analysés. Enfin, une dernière limite découle du fait qu'une contribution originale en science n'est pas équivalente au concept de nouveauté en

propriété intellectuelle (Packer et Webster, 1996), ce qui limite la précision du statut d'inventeur comme indicateur de l'importance d'une contribution.

Conclusion générale

Notre étude a examiné les PBA et plus particulièrement leurs listes d'auteurs et d'inventeurs afin de déterminer, d'une part, la relation entre divers facteurs et le ratio inventeurs/auteurs et, d'autre part, la relation entre divers facteurs associés à l'obtention du statut d'auteur et d'inventeur. Cet objectif général se subdivise en quatre objectifs spécifiques, auxquels sont associées au total dix questions de recherche :

Objectif 1. Déterminer la relation entre les caractéristiques de la recherche et le nombre d'auteurs et d'inventeurs.

- QR1. Quelle est la relation entre la discipline de recherche et le ratio inventeurs/auteurs des PBA?
- QR2. Quelle est la relation entre l'impact scientifique potentiel de la recherche et le ratio inventeurs/auteurs des PBA?

Objectif 2. Déterminer la relation entre les caractéristiques de l'équipe de recherche et le nombre d'auteurs et d'inventeurs.

- QR3. Quelle est la relation entre la taille de l'équipe de recherche et le ratio inventeurs/auteurs des PBA?
- QR4. Quelle est la relation entre la collaboration interinstitutionnelle et le ratio inventeurs/auteurs des PBA?

Objectif 3. Déterminer la relation entre les contributions respectives des auteurs et l'obtention du statut d'inventeur.

- QR5. Quelle est la relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur?
- QR6. La relation entre la position dans la liste des auteurs d'un article et l'obtention du statut d'inventeur est-elle liée à la discipline?
- QR7. Quelle est la relation entre le rôle dans l'équipe de recherche et l'obtention du statut d'inventeur?

QR8. Quelle est la relation entre le type de contribution à la recherche et l'obtention du statut d'inventeur?

Objectif 4. Déterminer la relation entre des caractéristiques individuelles des auteurs et l'obtention du statut d'inventeur.

QR9. Quelle est la relation entre la réputation d'un auteur et l'obtention du statut d'inventeur?

QR10. Quelle est la relation entre le sexe d'un auteur et l'obtention du statut d'inventeur?

Notre thèse a adopté une approche quantitative utilisant des méthodes bibliométriques pour l'analyse de PBA formées à partir des métadonnées des brevets du USPTO dont le dépôt a été effectué entre 1991 et 2015, et des articles indexés dans le WoS et publiés entre 1991 et 2016. À partir de ces bases de données, nous avons d'abord formé un ensemble de 6 312 PBA que nous avons ensuite pu caractériser à l'aide de diverses variables : 1) le nombre d'auteurs, 2) le nombre d'articles, 3) la discipline, 4) le facteur d'impact de la revue et 5) le caractère multi-institutionnel de la collaboration. Ensuite, nous avons recueilli un ensemble d'informations sur chacun des auteurs des articles formant les PBA : 1) la position dans la liste des auteurs, 2) le rôle, 3) la contribution, 4) la réputation et 5) le sexe.

La revue de la littérature a permis de formuler onze hypothèses de recherche, testées à l'aide de modèles de régression linéaire multiple (QR1 à QR4) et de modèles de régression logistique (QR5 à QR10). Notre étude a permis d'en accepter six complètement et cinq partiellement. Nous avons ainsi déterminé que le ratio inventeurs/auteurs des PBA varie selon la discipline, l'impact potentiel de la recherche (Facteur d'Impact de la revue), le mode de collaboration (interinstitutionnelle vs uni-institutionnelle) ainsi que la taille de l'équipe (nombre d'auteurs). Au niveau des chercheurs, nous avons observé des relations entre l'obtention du statut d'inventeur et la position dans la liste des auteurs, le rôle, et la contribution, la réputation et le sexe, la taille et la significativité des effets observés variant selon la discipline.

Dans l'ensemble, nos résultats montrent une plus grande division du travail inventif dans la recherche en génie, alors qu'en sciences naturelles et, surtout, en recherche biomédicale, le travail inventif est typiquement effectué par quelques auteurs, la majorité ayant effectué une

contribution (possiblement technique ou routinière) ne menant pas à l'obtention du statut d'inventeur. Notre étude a également montré que les auteurs occupant la première et la dernière position dans les listes d'auteurs sont ceux qui obtiennent plus souvent le statut d'inventeur et que les hommes ont aussi une probabilité plus grande d'obtenir le statut d'inventeur. Nous présentons dans les sections suivantes les implications théoriques, méthodologiques et pratiques de la recherche.

Implications théoriques

Notre étude démontre que la particularité des cultures épistémiques (Knorr Cetina, 1999) peut s'observer dans les pratiques de collaboration et de division du travail dans les équipes de recherche. Elle permet, en effet, d'illustrer comment la production des connaissances dans un domaine comme le génie est caractérisée par de petites collaborations où la plupart des individus contribuent au travail intellectuel de conception, alors que les équipes en recherche biomédicale et sciences naturelles semblent avoir tendance à adopter des modes d'organisation du travail plus fragmentés et hiérarchiques, donnant lieu à de plus grandes équipes de travail où la plupart des individus ne participent pas à la conception de l'étude et réalisent un travail plus technique.

La relation observée en recherche biomédicale et en sciences naturelles entre le ratio inventeurs/auteurs des PBA et l'impact potentiel de la recherche illustre le caractère relativement subjectif de l'attribution du statut d'auteur. Cela renforce l'idée selon laquelle celle-ci est susceptible d'être influencée par les rapports de force entre les membres d'une équipe, par des considérations stratégiques ou par les préférences personnelles des individus, par exemple.

L'existence d'un effet Saint-Matthieu (Merton, 1968) et d'un mécanisme d'accumulation des avantages et des désavantages (Zuckerman, 1998; Bourdieu, 1976) est moins clairement illustrée par les résultats de notre étude, l'effet marginal de la réputation ne s'étant pas avéré significatif dans plusieurs de nos analyses. Cela dit, comme nous l'avons vu dans le Chapitre 4, il est possible que cet effet soit réduit parce que l'effet Saint-Matthieu est plus prévalent dans l'attribution du statut d'auteur que dans l'attribution du statut d'inventeur. Il est aussi possible que nos modèles de régression logistique permettent difficilement d'identifier l'effet Saint-Matthieu parce que les chercheurs réputés ont tendance à occuper la

position d'autorité dans l'équipe de recherche et donc à contrôler la répartition des tâches et l'attribution des statuts d'auteur et d'inventeur. Ainsi, en bénéficiant davantage aux chercheurs dominants dans la structure stratifiée de la science (Cole et Cole, 1973), la commercialisation de la recherche pourrait contribuer à accroître l'écart entre les dominants et les dominés.

Notre étude apporte aussi de nouvelles preuves empiriques de l'existence d'un effet Matilda en science selon lesquelles à contribution égale, les femmes obtiennent moins de reconnaissances que les hommes (Rossiter, 1993). Les résultats obtenus suggèrent que les inégalités entre les sexes dans l'obtention du statut d'inventeur ne sont pas seulement dues au fait que les femmes jouent le plus souvent un rôle de soutien dans les équipes de recherche où elles effectuent surtout des tâches techniques (Macaluso et al., 2016). En effet, à position dans la liste des auteurs, rôle, contribution et réputation égaux, les femmes ont une probabilité plus faible que les hommes d'obtenir le statut d'inventeur.

Implications méthodologiques

Les implications méthodologiques de notre étude portent sur l'usage des PBA pour l'étude de la division du travail dans les équipes de recherche, ainsi que sur la méthode utilisée pour l'appariement des PBA.

Notre étude est à notre connaissance la première à analyser des PBA pour mieux comprendre les modes de collaboration et de division du travail dans les équipes de recherche, les études antérieures s'étant surtout intéressées aux facteurs déterminant l'accès aux statuts d'auteur et d'inventeur. Les résultats obtenus et les conclusions que nous avons pu en tirer montrent que cette méthode peut permettre l'usage de jeux de données volumineux pour observer les grandes tendances concernant la collaboration et la division du travail en recherche et de comparer différents groupes (p. ex. différentes disciplines).

La méthode utilisée étant très similaire à celle de Magerman et al. (2015), elle ne constitue pas une importante contribution méthodologique. La contribution de notre recherche à ce niveau découle plutôt de notre validation manuelle de l'ensemble des PBA formées ayant un score de similarité de 0,30. Les résultats de cette vérification permettent de constater que suivant notre méthode, les PBA avec un score de similarité au-delà de 0,50 sont généralement

valides. Ce seuil pourrait donc être potentiellement utilisé par des processus d'appariement des articles et de brevets sans nécessiter une vérification manuelle.

Implications pratiques

Les résultats obtenus dans le cadre de cette recherche permettent de mettre en lumière certaines limites de la bibliométrie pour l'évaluation de la recherche au niveau individuel. En effet, la plupart des indicateurs bibliométriques généralement utilisés pour comparer les chercheurs entre eux ne tiennent pas compte de la position dans la liste des auteurs. Notre étude démontre que les articles scientifiques signés par plusieurs auteurs sont les fruits de pratiques de collaboration et de division du travail très diverses, et que les listes d'auteurs fournissent peu d'information sur les contributions respectives de chacun. Nos résultats peuvent ainsi servir de mise en garde contre l'utilisation peu critique d'indicateurs bibliométriques qui ne tiennent pas compte des différences dans la nature et l'ampleur des contributions respectives des auteurs, et qui peuvent ainsi en favoriser certains au détriment des autres, et rendent banales les comparaisons entre les disciplines.

De plus, bien que notre recherche ait permis d'observer des tendances générales particulières à chacune des disciplines étudiées, s'y fier pour déduire l'ampleur ou la contribution d'un individu à un article particulier serait commettre une erreur écologique (Robinson, 1950). Ce type d'erreur survient lorsque l'on interprète une observation unique en fonction des tendances observées à partir de données agrégées.

Les résultats de notre étude appuient l'idée que les femmes et les chercheurs peu cités sont plus souvent exclus de la liste des inventeurs d'un brevet que leurs collègues masculins s'étant bâti une certaine réputation. Empêcher certains individus d'accéder à ce capital symbolique et potentiellement économique contribue potentiellement à une accumulation injustifiée des avantages pour certains et des désavantages pour d'autres (effet Saint-Matthieu). Or, comme nous l'avons vu au Chapitre 1, cela constitue une violation de la norme d'universalisme (Merton, 1942) pouvant nuire à la performance de l'institution scientifique. Nos résultats, sans apporter de solutions concrètes, peuvent néanmoins servir à appuyer le développement de politiques visant à réduire les inégalités en science, et plus particulièrement les inégalités dans l'attribution des statuts d'auteur et d'inventeur.

Recherches futures

Notre étude n'a pas tenu compte des pays et institutions d'affiliation des auteurs et inventeurs. Or, il est possible que le milieu (industrie, académie, gouvernement) d'un individu guide son intérêt pour un type de capital plutôt qu'un autre. Par exemple, il peut être plus important pour un chercheur provenant de l'industrie (et pour son employeur) d'obtenir le statut d'inventeur, puisque cela permettra à son employeur de commercialiser l'invention et d'en tirer un profit. Pour cet individu, l'enjeu est davantage l'innovation technologique à des fins commerciales que l'avancement des connaissances afin de construire sa réputation dans le champ scientifique. Il sera donc intéressant de tenir compte de cette variable dans de futures recherches. Aussi, il est possible que des différences nationales et régionales soient observées, comme celles observées par Pontille (2004) dans le cas de la sociologie.

L'interdisciplinarité des équipes de recherche serait aussi intéressante à considérer dans de futures recherches. En effet, une collaboration impliquant des chercheurs de différentes disciplines peut constituer un choc des cultures épistémiques pouvant se traduire par des formes particulières de division du travail et d'attribution du statut d'auteur ou d'inventeur.

Une autre avenue de recherche future est l'utilisation de méthodes qualitatives, comme l'étude de cas. En effet, une des principales limites de notre étude est la difficulté à illustrer la diversité des pratiques et décrire comment les décisions concernant l'attribution du statut d'auteur sont prises par les équipes de recherche (ou par les directeurs de laboratoire). Des études qualitatives au moyen d'entrevues, par exemple, permettraient de mieux comprendre et exemplifier l'aspect stratégique et pragmatique de l'attribution des statuts d'auteur et d'inventeur.

Finalement, il serait intéressant d'élargir le concept d'équipe de travail pour y inclure tous les individus ayant participé à la recherche, incluant ceux qui ne figurent pas dans la liste des auteurs. Cela pourrait être fait en comptant les individus qui sont remerciés dans l'article (Paul-Hus et al., 2017) ou encore en demandant aux participants d'un sondage ou d'entrevues combien d'individus ont participé à la recherche sans obtenir le statut d'auteur. Cette recherche future permettrait notamment de déterminer dans quelle mesure les différences disciplinaires en termes de ratio inventeurs/auteurs que nous avons observées sont causées par différentes

pratiques de collaboration et de division du travail, ou plutôt par différentes pratiques d'attribution du statut d'auteur.

Bibliographie

- Anderson, M. H., Cislo, D., Saavedra, J. et Cameron, K. (2013). Why international inventors might want to consider filing their first patent application at the United States Patent Office & the convergence of patent harmonization and e-commerce. *Santa Clara High Technology Law Journal*, 30(4), 555–579.
- Archambault, A., Mongeon, P. et Larivière, V. (2017). On the effects of the reunification on German researchers' publication patterns. *Scientometrics*, 111(1), 337–347. <https://doi.org/10.1007/s11192-017-2254-9>
- Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V. et Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329–342. <https://doi.org/10.1007/s11192-006-0115-z>
- Armstrong, M. et Murphy, G. M. (2012). Inventorship and ownership considerations and pitfalls with collaborative research. *ACS Medicinal Chemistry Letters*, 3(5), 349–351. <https://doi.org/10.1021/ml300084e>
- Azoulay, P., Bonatti, A. et Krieger, J. (2015). The career effects of scandal: Evidence from scientific retractions. *NBER Working Paper*. <https://doi.org/10.3386/w21146>
- Azoulay, P., Furman, J. L., Krieger, J. et Murray, F. (2015). Retractions. *Review of Economics and Statistics*, 97(5), 1118–1136. https://doi.org/10.1162/REST_a_00469
- Baker, S. L. et Lancaster, F. W. (1991). *The measurement and evaluation of library services* (2^e éd.). Arlington, VA : Information Resources Press.
- Bassecoulard, E. et Zitt, M. (2005). Patents and publications. Dans H. F. Moed, W. Glanzel et U. Schmoch (dir.), *Handbook of quantitative science and technology research* (p. 665–694). Londres : Kluwer Academic Publishers.
- Bates, T., Anić, A., Marušić, M. et Marušić, A. (2004). Authorship criteria and disclosure of contributions. *JAMA*, 292(1), 86–88. <https://doi.org/10.1001/jama.292.1.86>
- Beaver, D. de B. et Rosen, R. (1979). Studies in scientific collaboration. Part II: Scientific co-authorship, research productivity and visibility in the French scientific elite, 1799–1830. *Scientometrics*, 1(2), 133–149. <https://doi.org/10.1007/BF02016966>
- Bebeau, M. et Monson, V. (2011). Authorship and publication practices in the social sciences: historical reflections on current practices. *Science and Engineering Ethics*, 17(2), 365–388. <https://doi.org/10.1007/s11948-011-9280-4>

- Bérubé, N., Ghiasi, G. et Larivière V. (en préparation). Assigning gender to first names using Wikipedia.
- Bhopal, R., Rankin, J., McColl, E., Thomas, L., Kaner, E., Stacy, R., ... Rodgers, H. (1997). The vexed question of authorship: Views of researchers in a British medical faculty. *BMJ*, 314(7086), 1009. <https://doi.org/10.1136/bmj.314.7086.1009>
- Biagioli, M. (2003). Rights or rewards? Changing frameworks of scientific authorship. Dans M. Biagioli et P. Galison (dir.), *Scientific Authorship : Credit and Intellectual Property in Science* (p. 253–279). New York, NY : Routledge.
- Birnholtz, J. P. (2006). What does it mean to be an author? The intersection of credit, contribution, and collaboration in science. *Journal of the American Society for Information Science and Technology*, 57(13), 1758–1770. <https://doi.org/10.1002/asi.20380>
- Bonnewitz, P. (2002). *Pierre Bourdieu : vie, œuvres, concepts*. Paris : Ellipses.
- Bošnjak, L. et Marušić, A. (2012). Prescribed practices of authorship: Review of codes of ethics from professional bodies and journal guidelines across disciplines. *Scientometrics*, 93(3), 751–763. <https://doi.org/10.1007/s11192-012-0773-y>
- Bourdieu, P. (1971). Le marché des biens symboliques. *L'Année sociologique*, 3(22), 49–126. <https://doi.org/10.2307/27887912>
- Bourdieu, P. (1975). La spécificité du champ scientifique et les conditions sociales du progrès de la raison. *Sociologie et sociétés*, 7(1), 91–118. <https://doi.org/10.7202/001089ar>
- Bourdieu, P. (1976). Le champ scientifique. *Actes de la recherche en sciences sociales*, 2(2), 88–104. <https://doi.org/10.3406/arss.1976.3454>
- Bourdieu, P. (1979). Les trois états du capital culturel. *Actes de la recherche en sciences sociales*, 30(1), 3–6. <https://doi.org/10.3406/arss.1979.2654>
- Bourdieu, P. (1980). Le capital social. *Actes de la recherche en sciences sociales*, 31(1), 2–3. <https://doi.org/10.3406/arss.1980.2069>
- Bourdieu, P. (1986a). Habitus, code et codification. *Actes de la recherche en sciences sociales*, 64(1), 40–44. <https://doi.org/10.3406/arss.1986.2335>
- Bourdieu, P. (1986b). The forms of capital. Dans J. G. Richardson (dir.), *Handbook of theory and research for the sociology of education* (p. 241–258). New York, NY : Greenwood.
- Bourdieu, P. (1987). *Choses dites*. Paris : Éditions de minuit.
- Bourdieu, P. (1989). Social space and symbolic power. *Sociological Theory*, 7(1), 14–15. <https://doi.org/10.2307/202060>

- Bourdieu, P. (1993). Esprits d'État : Genèse et structure du champ bureaucratique. *Actes de la recherche en sciences sociales*, 96(1), 49–62. <https://doi.org/10.3406/arss.1993.3040>
- Bourdieu, P. (1997). *Les usages sociaux de la science : pour une sociologie clinique du champ scientifique*. Paris : Institut national de la recherche agronomique.
- Bourdieu, P. (2001). *Science de la science et réflexivité*. Paris : Raisons d'agir.
- Boyack, K. W. et Klavans, R. (2008). Measuring science–technology interaction using rare inventor–author names. *Journal of Informetrics*, 2(3), 173–182. <https://doi.org/10.1016/j.joi.2008.03.001>
- Bozeman, B. et Youtie, J. (2016). Trouble in paradise: Problems in academic research co-authoring. *Science and Engineering Ethics*, 22(6), 1717–1743. <https://doi.org/10.1007/s11948-015-9722-5>
- Bradford, S. (1934). Sources of information on specific subjects. *Engineering: An Illustrated Weekly Journal*, 137, 85–86.
- Breschi, S. et Catalini, C. (2010). Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. *Research Policy*, 39(1), 14–26. <https://doi.org/10.1016/j.respol.2009.11.004>
- Cabanac, G.. (2015). Unconventional academic writing. figshare. <https://doi.org/10.6084/m9.figshare.1306561.v1>
- Caron, E. et van Eck, N. J. (2014). Large scale author name disambiguation using rule-based scoring and clustering. Dans E. Noyons (dir.), *Proceedings of the 19th International Conference on Science and Technology Indicators* (p. 79–86). Leiden : CWTS-Leiden University.
- Cassiman, B., Glenisson, P. et Van Looy, B. (2007). Measuring industry-science links through inventor-author relations: A profiling methodology. *Scientometrics*, 70(2), 379–391. <https://doi.org/10.1007/s11192-007-0208-3>
- Ceci, S. J. et Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, 108(8), 3157–3162. <https://doi.org/10.1073/pnas.1014871108>
- Claxton, L. D. (2005). Scientific authorship: Part 2. History, recurring issues, practices, and guidelines. *Mutation Research/Reviews in Mutation Research*, 589(1), 31–45. <https://doi.org/10.1016/j.mrrev.2004.07.002>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.

- Cole, J. R. et Cole, S. (1973). *Social stratification in science*. Chicago, IL : University of Chicago Press.
- Cole, S. et Cole, J. R. (1967). Scientific output and recognition: A study in the operation of the reward system in science. *American Sociological Review*, 32(3), 377–390. <https://doi.org/10.2307/2091085>
- Cook, T. D. et Campbell, D. T. (1979). *Quasi-experimentation: design & analysis issues for field settings*. Chicago, IL : Rand McNally.
- Council of Science Editors. (2012). *CSE's white paper on promoting integrity in scientific journal publications*. Repéré à <https://www.councilscienceeditors.org/resource-library/editorial-policies/white-paper-on-publication-ethics/>
- Couture, M. et Malissard, P. (2010). La réglementation de la propriété intellectuelle dans les universités canadiennes. Dans M. Couture, M. Dubé et P. Malissard (dir.), *Propriété intellectuelle et université : entre la libre circulation des idées et la privatisation des savoirs* (p. 257–304). Québec : Presses de l'Université du Québec.
- Coward, H. R. et Franklin, J. J. (1989). Identifying the science-technology interface: Matching patent data to a bibliometric model. *Science, Technology & Human Values*, 14(1), 50–77. <https://doi.org/10.1177/016224398901400106>
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks : Sage Publications.
- Cronin, B., Mckenzie, G. et Rubio, L. (1993). The norms of acknowledgment in four humanities and social sciences disciplines. *Journal of Documentation*, 49(1), 29–43. <https://doi.org/10.1108/eb026909>
- Cronin, B. (2001). Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *Journal of the American Society for Information Science and Technology*, 52(7), 558–569. <https://doi.org/10.1002/asi.1097>
- Cronin, B. (2005). *The hand of science: Academic writing and its rewards*. Lanham, MD : Scarecrow Press.
- Cronin, B., McKenzie, G. et Rubio, L. (1993). The norms of acknowledgment in four humanities and social sciences disciplines. *Journal of Documentation*, 49(1), 29–43. <https://doi.org/10.1108/eb026909>
- Cronin, B. et Overfelt, K. (1994). The scholar's courtesy: A survey of acknowledgement behaviour. *Journal of Documentation*, 50(3), 165–196. <https://doi.org/10.1108/eb026929>

- Czarnitzki, D., Glänzel, W. et Hussinger, K. (2007). Patent and publication activities of German professors: An empirical assessment of their co-activity. *Research Evaluation*, 16(4), 311–319. <https://doi.org/10.3152/095820207X254439>
- Czarnitzki, D., Hussinger, K. et Schneider, C. (2011). Commercializing academic research: The quality of faculty patenting. *Industrial and Corporate Change*, 20(5), 1403–1437. <https://doi.org/10.1093/icc/dtr034>
- Dasgupta, P. et David, P. A. (1994). Toward a new economics of science. *Research Policy*, 23(5), 487–521. [https://doi.org/10.1016/0048-7333\(94\)01002-1](https://doi.org/10.1016/0048-7333(94)01002-1)
- de Melo-Martín, I. (2013). Patenting and the gender gap: Should women be encouraged to patent more? *Science and Engineering Ethics*, 19(2), 491–504. <https://doi.org/10.1007/s11948-011-9344-5>
- Desrochers, N. et Larivière, V. (2016). Je veux bien, mais me citerez-vous? On publication language strategies in an anglicized research landscape. Dans I. Rafols, J. Molas-Gallart, E. Castro-Martinez et R. Woolley (dir.), *Proceedings of the 21st International Conference on Science and Technology Indicators* (p. 887–892). Valence : Université polytechnique de Valence.
- Ding, W. W., Murray, F. E. et Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *Science*, 313(5787), 665–667. <https://doi.org/10.1126/science.1124832>
- Dornbusch, F., Schmoch, U., Schulze, N. et Bethke, N. (2013). Identification of university-based patents: A new large-scale approach. *Research Evaluation*, 22(1), 52–63. <https://doi.org/10.1093/reseval/rvs033>
- Dubé, M. (2010a). La propriété intellectuelle : nature et portée juridique. Dans M. Couture, M. Dubé et P. Malissard (dir.), *Propriété intellectuelle et université : entre la libre circulation des idées et la privatisation des savoirs* (p. 11–24). Québec : Presses de l'Université du Québec.
- Dubé, M. (2010b). La titularité de la propriété intellectuelle. Dans M. Couture, M. Dubé et P. Malissard (dir.), *Propriété intellectuelle et université : entre la libre circulation des idées et la privatisation des savoirs* (p. 55–77). Québec : Presses de l'Université du Québec.
- Ducor, P. (2000). Intellectual property: Coauthorship and coinventorship. *Science*, 289(5481), 873–875. <https://doi.org/10.1126/science.289.5481.873>
- Durand, P. (s. d.). Capital symbolique, dans A. Glinoyer et D. Saint-Amand (dir.), *Le lexique socius*, Repéré à <http://ressources-socius.info/index.php/lexique/21-lexique/39-capital-symbolique>

- Etzkowitz, H. (1998). The norms of entrepreneurial science: Cognitive effects of the new university–industry linkages. *Research Policy*, 27(8), 823–833. [https://doi.org/10.1016/S0048-7333\(98\)00093-6](https://doi.org/10.1016/S0048-7333(98)00093-6)
- Etzkowitz, H. et Leydesdorff, L. (2000). The dynamics of innovation: From National Systems and “Mode 2” to a Triple Helix of university–industry–government relations. *Research Policy*, 29(2), 109–123. [https://doi.org/10.1016/S0048-7333\(99\)00055-4](https://doi.org/10.1016/S0048-7333(99)00055-4)
- Fang, F. F. C., Steen, R. G. et Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, 109(42), 17028–17033. <https://doi.org/10.1073/pnas.1212247109>
- Fehder, D. C., Murray, F. et Stern, S. (2014). Intellectual property rights and the evolution of scientific journals as knowledge platforms. *International Journal of Industrial Organization*, 36, 83–94. <https://doi.org/10.1016/j.ijindorg.2014.08.002>
- Field, A. P. (2009). *Discovering statistics using SPSS: And sex and drugs and rock ‘n’ roll* (3^e éd). Los Angeles, CA : SAGE Publications.
- Flanagin, A., Carey, L. A., Fontanarosa, P. B., Phillips, S. G., Pace, B. P., Lundberg, G. D. et Rennie, D. (1998). Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *JAMA*, 280(3), 222–224. <https://doi.org/10.1001/jama.280.3.222>
- Fortin, M.-F. et Gagnon, J. (2016). *Fondements et étapes du processus de recherche : méthodes quantitatives et qualitatives* (3^e éd.). Montréal : Chenelière Éducation.
- Fox, J. (2016). *Applied regression analysis and generalized linear models* (3^e éd.). Los Angeles, CA : Sage Publications.
- Frandsen, T. T. F. et Nicolaisen, J. (2010). What is in a name? Credit assignment practices in different disciplines. *Journal of Informetrics*, 4(4), 608–617. <https://doi.org/10.1016/j.joi.2010.06.010>
- Gans, J. S., Murray, F. E. et Stern, S. (2017). Contracting over the disclosure of scientific knowledge: Intellectual property and academic publication. *Research Policy*, 46(4), 820–835. <https://doi.org/10.1016/j.respol.2017.02.005>
- Garfield, E. (1983). What’s in a name? The eponymic route to immortality. *Current contents*, (47), 5–16.
- Garfield, E. (1990). How ISI selects journals for coverage: Quantitative and qualitative considerations. *Current Contents*, (22), 5–13.
- Gelman, A. et Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, NY : Cambridge University Press.

- Ghiasi, G., Larivière, V. et Sugimoto, C. R. (2015). On the compliance of women engineers with a gendered scientific system. *PLoS ONE*, 10(12), e0145931. <https://doi.org/10.1371/journal.pone.0145931>
- Gingras, Y. et Mosbah-Natanson, S. (2010). Les sciences sociales françaises entre ancrage local et visibilité internationale. *Archives européennes de sociologie*, 51(02), 305- 321. <https://doi.org/10.1017/S0003975610000147>
- Gingras, Y. (2014). *Les dérives de l'évaluation de la recherche : du bon usage de la bibliométrie*. Paris : Raisons d'agir.
- Glänzel, W. et Meyer, M. (2003). Patents cited in the scientific literature: An exploratory study of “reverse” citation relations. *Scientometrics*, 58(2), 415–428. <https://doi.org/10.1023/A:1026248929668>
- Gläser, B. G. (1963). Variations in the importance of recognition in scientists' careers. *Social Problems*, 10(3), 268–276. <https://doi.org/10.2307/799436>
- Godlee, F. (1996). Definition of “authorship” may be changed. *BMJ (Clinical Research Ed.)*, 312(7045), 1501–1502. <https://doi.org/10.2307/29731972>
- Grosset-Fournier, C. et Dacheux, A. (2012). *Le brevet d'invention : la cause des inventeurs : guide pratique en 101 questions*. Paris : Tec & doc.
- Guba, E. G. et Lincoln, Y. S. (1989). *Fourth generation evaluation*. Thousand Oaks, CA : Sage Publications.
- Haeussler, C. et Colyvas, J. A. (2011). Breaking the ivory tower: Academic entrepreneurship in the life sciences in UK and Germany. *Research Policy*, 40(1), 41–54. <https://doi.org/10.1016/j.respol.2010.09.012>
- Haeussler, C. et Sauermann, H. (2013). Credit where credit is due? The impact of project contributions and social factors on authorship and inventorship. *Research Policy*, 42(3), 688–703. <https://doi.org/10.1016/j.respol.2012.09.009>
- Haeussler, C., & Sauermann, H. (2015). The Anatomy of Teams: Division of Labor in Collaborative Knowledge Production. *Academy of Management Proceedings*, 2015(1), 11383-11383. <https://doi.org/10.5465/AMBPP.2015.11383abstract>
- Hagstrom, W. O. (1964). Traditional and modern forms of scientific teamwork. *Administrative Science Quarterly*, 9(3), 241. <https://doi.org/10.2307/2391440>
- Henriksen, D. (2016). The rise in co-authorship in the social sciences (1980–2013). *Scientometrics*, 107(2), 455–476. <https://doi.org/10.1007/s11192-016-1849-x>

- Ho, C. M. (2004). Who deserves the patent pot of gold? An inquiry into the proper inventorship of patient-based discoveries. *DePaul Journal of Health Care Law*, 7(2), 185–243. Repéré à <http://via.library.depaul.edu/jhcl/vol7/iss2/2>
- Holden, C. (2001). General contentment masks gender gap in first AAAS salary and job survey. *Science*, 294(5541), 396–411. <https://doi.org/10.1126/science.294.5541.396>
- Hong, W. (2008). Domination in a scientific field: Capital struggle in a Chinese isotope lab. *Social Studies of Science*, 38(4), 543–570. <https://doi.org/10.1177/0306312706092456>
- Hu, X. (2009). Loads of special authorship functions: Linear growth in the percentage of “equal first authors” and corresponding authors. *Journal of the American Society for Information Science and Technology*, 60(11), 2378–2381. <https://doi.org/10.1002/asi.21164>
- Huth, E. J. (1986a). Irresponsible authorship and wasteful publication. *Annals of Internal Medicine*, 104(2), 257–259. <https://doi.org/10.7326/0003-4819-104-2-257>
- Huth, E. J. (1986b). Guidelines on authorship of medical papers. *Annals of Internal Medicine*, 104(2), 269–274. <https://doi.org/10.7326/0003-4819-104-2-269>
- Ilakovac, V., Fister, K., Marusic, M. et Marusic, A. (2006). Reliability of disclosure forms of authors’ contributions. *Canadian Medical Association Journal*, 176(1), 41–46. <https://doi.org/10.1503/cmaj.060687>
- Jabbehdari, S. et Walsh, J. P. (2017). Authorship norms and project structures in science. *Science, Technology et Human Values*, 42(5), 872–900. <https://doi.org/10.1177/0162243917697192>
- Jourdain, A. et Naulin, S. (2011). *La théorie de Pierre Bourdieu et ses usages sociologiques*. Paris : Armand Colin.
- Kan, S. S. (2009a). Court standards on joint inventorship and authorship. *DePaul Journal of Art, Technology & Intellectual Property Law*, 19(Spring), 267–297. Repéré à <http://via.library.depaul.edu/jatip/vol19/iss2/4>
- Kan, S. S. (2009b). The efficient boundary of inventorship and authorship. *DePaul Journal of Art, Technology & Intellectual Property Law*, 19(Spring), 235–256. Repéré à <http://via.library.depaul.edu/jatip/vol19/iss2/3>
- Kassirer, J. P. et Angell, M. (1991). On authorship and acknowledgments. *New England Journal of Medicine*, 325(21), 1510–1512. <https://doi.org/10.1056/NEJM199111213252112>
- Katz, J. S. et Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18. [https://doi.org/10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1)

- Knorr Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA : Harvard University Press.
- Kosmulski, M. (2012). The order in the lists of authors in multi-author papers revisited. *Journal of Informetrics*, 6(4), 639–644. <https://doi.org/10.1016/j.joi.2012.06.006>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Larivière, V. et Costas, R. (2016). How Many Is Too Many? On the Relationship between Research Productivity and Impact. *PLOS ONE*, 11(9), e0162709. <https://doi.org/10.1371/journal.pone.0162709>
- Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A. et Sugimoto, C. R. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Science*, 46(3), 417–435. <https://doi.org/10.1177/0306312716650046>
- Larivière, V., Gingras, Y., Sugimoto, C. R. et Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323–1332. <https://doi.org/10.1002/asi.23266>
- Larivière, V., Ni, C., Gingras, Y., Cronin, B. et Sugimoto, C. R. (2013). Global gender disparities in science. *Nature*, 504(7479), 211–213. <https://doi.org/10.1038/504211a>
- Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation*, 11(1), 3–15. <https://doi.org/10.3152/147154402781776961>
- Leash, E. (1997). Is it time for a new approach to authorship? *Journal of Dental Research*, 76(3), 724–727. <https://doi.org/10.1177/00220345970760030101>
- Lebaron, F. (2008). *La sociologie de A à Z : 250 mots pour comprendre*. Paris : Dunod.
- Ley, T. J. et Hamilton, B. H. (2008). The gender gap in NIH grant applications. *Science*, 322(5907), 1472–1474. <https://doi.org/10.1126/science.1165878>
- Leydesdorff, L. et Meyer, M. (2003). The Triple Helix of university-industry-government relations. *Scientometrics*, 58(2), 191–203. <https://doi.org/10.1023/A:1026276308287>
- Lincoln, A. E., Pincus, S., Koster, J. B. et Leboy, P. S. (2012). The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science*, 42(2), 307–320. <https://doi.org/10.1177/0306312711435830>
- Lissoni, F., Montobbio, F. et Zirulia, L. (2013). Inventorship and authorship as attribution rights: An enquiry into the economics of scientific credit. *Journal of Economic Behavior & Organization*, 95, 49–69. <https://doi.org/10.1016/j.jebo.2013.08.016>

- Lu, S. F., Jin, G. Z., Uzzi, B. et Jones, B. (2013). The retraction penalty: Evidence from the Web of Science. *Scientific Reports*, 3(1), 3146–3150. <https://doi.org/10.1038/srep03146>
- Macaluso, B., Larivière, V., Sugimoto, T. et Sugimoto, C. R. (2016). Is science built on the shoulders of women? A study of gender differences in contributorship. *Academic Medicine*, 91(8), 1136–1142. <https://doi.org/10.1097/ACM.0000000000001261>
- MacRoberts, M. H. et MacRoberts, B. R. (2017). The mismeasure of science: Citation analysis. *Journal of the Association for Information Science and Technology*. Prépublication. doi:10.1002/asi.23970
- Magerman, T., Van Looy, B., Baesens, B. et Debackere, K. (2011). *Assessment of Latent Semantic Analysis (LSA) text mining algorithms for large scale mapping of patent and scientific publication documents*. FBE research report. Louvain : KU Leuven, Faculty of Business and Economics.
- Magerman, T., Van Looy, B. et Debackere, K. (2015). Does involvement in patenting jeopardize one's academic footprint? An analysis of patent-paper pairs in biotechnology. *Research Policy*, 44(9), 1702–1713. <https://doi.org/10.1016/j.respol.2015.06.005>
- Magerman, T., Van Looy, B. et Song, X. (2009). Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications. *Scientometrics*, 82(2), 289–306. <https://doi.org/10.1007/s11192-009-0046-6>
- Malissard, P. (2010). L'université et la commercialisation des innovations techniques. Dans M. Couture, M. Dubé et P. Malissard (dir.), *Propriété intellectuelle et université : entre la libre circulation des idées et la privatisation des savoirs* (p. 135–167). Québec : Presses de l'Université du Québec.
- Malissard, P., Gingras, Y. et Gemme, B. (2003). La commercialisation de la recherche. *Actes de la recherche en sciences sociales*, 148(1), 57–67. <https://doi.org/10.3406/arss.2003.3322>
- Mansfield, E. (1991). Academic research and industrial innovation. *Research Policy*, 20(1), 1–12. [https://doi.org/10.1016/0048-7333\(91\)90080-A](https://doi.org/10.1016/0048-7333(91)90080-A)
- Medawar, P. (1963). Is the scientific paper a fraud? *The Listener*, (70), 377–378.
- Melin, G. (2000). Pragmatism and self-organization. *Research Policy*, 29(1), 31–40. [https://doi.org/10.1016/S0048-7333\(99\)00031-1](https://doi.org/10.1016/S0048-7333(99)00031-1)
- Meng, Y. (2016). Collaboration patterns and patenting: Exploring gender distinctions. *Research Policy*, 45(1), 56–67. <https://doi.org/10.1016/j.respol.2015.07.004>
- Merton, R. K. (1942). Science and technology in a democratic order. *Journal of Legal and Political Sociology*, 1–2, 115–126. <https://doi.org/10.2307/39015008014428>

- Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22(6), 635. <https://doi.org/10.2307/2089193>
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56–63. <https://doi.org/10.2307/1723414>
- Merton, R. K. (1973). “Recognition” and “Excellence”: Instructive ambiguities. Dans R. K. Merton (dir.), *The sociology of science. Theoretical and empirical investigations* (p. 419–437). Chicago, IL : The University of Chicago Press.
- Merton, R. K. (1979). *The sociology of science: An episodic memoir*. Carbondale, IL : Southern Illinois University Press.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623. <https://doi.org/10.1086/354848>
- Merton, R. K. et Zuckerman, H. A. (1973). Institutionalized patterns of evaluation in science. Dans R. K. Merton (dir.), *The sociology of science : Theoretical and empirical investigations* (p. 460–496). Chicago, IL : The University of Chicago Press.
- Meyer, M. (2000). Does science push technology? Patents citing scientific literature. *Research Policy*, 29(3), 409–434. [https://doi.org/10.1016/S0048-7333\(99\)00040-2](https://doi.org/10.1016/S0048-7333(99)00040-2)
- Meyer, M. (2003). Academic patents as an indicator of useful research? A new approach to measure academic inventiveness. *Research Evaluation*, 12(1), 17–27. <https://doi.org/10.3152/147154403781776735>
- Meyer, M. (2006). Are patenting scientists the better scholars? *Research Policy*, 35(10), 1646–1662. <https://doi.org/10.1016/j.respol.2006.09.013>
- Mitroff, I. I. (1974). Norms and the counter-norms in a select group of the Appolo moon scientists: A case study of the ambivalence of scientists. *American Sociological Review*, 39(4), 579–595.
- Moed, H. F. (2005). Citation analysis in research evaluation. *Information Science and Knowledge Management*, 9. Dordrecht : Springer. <https://doi.org/10.1007/1-4020-3714-7>
- Mongeon, P. et Larivière, V. (2016). Costly collaborations: The impact of scientific fraud on co-authors’ careers. *Journal of the Association for Information Science and Technology*, 67(3), 535–542. <https://doi.org/10.1002/asi.23421>
- Mongeon, P. et Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>

- Mongeon, P., Smith, E., Joyal, B. et Larivière, V. (2017). The rise of the middle authors: Investigating collaboration and division of labor in biomedical research using partial alphabetical authorship. *PLoS ONE*, 12(9), e0184601. <https://doi.org/10.1371/journal.pone.0184601>
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J. et Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41), 16474–16479. <https://doi.org/10.1073/pnas.1211286109>
- Moulopoulos, S. D., Sideris, D. A. et Georgilis, K. A. (1983). For debate. . . Individual contributions to multiauthor papers. *BMJ*, 287(6405), 1608–1610. <https://doi.org/10.1136/bmj.287.6405.1608>
- Mowatt, G., Shirran, L., Grimshaw, J. M., Rennie, D., Flanagan, A., Yank, V., ... Bero, L. A. (2002). Prevalence of honorary and ghost authorship in Cochrane reviews. *JAMA*, 287(21), 2769–2771. <https://doi.org/10.1001/jama.287.21.2769>
- Murray, F. (2002). Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy*, 31(8–9), 1389–1403. [https://doi.org/10.1016/S0048-7333\(02\)00070-7](https://doi.org/10.1016/S0048-7333(02)00070-7)
- Murray, F. et Graham, L. (2007). Buying science and selling science: gender differences in the market for commercial science. *Industrial and Corporate Change*, 16(4), 657–689. <https://doi.org/10.1093/icc/dtm021>
- Murray, F. et Stern, S. (2007). Do formal intellectual property rights hinder the free flow of scientific knowledge? *Journal of Economic Behavior & Organization*, 63(4), 648–687. <https://doi.org/10.1016/j.jebo.2006.05.017>
- Narin, F., Hamilton, K. S. et Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy*, 26(3), 317–330. [https://doi.org/10.1016/S0048-7333\(97\)00013-9](https://doi.org/10.1016/S0048-7333(97)00013-9)
- Narin, F. et Olivastro, D. (1998). Linkage between patents and papers: An interim EPO/US comparison. *Scientometrics*, 41(1–2), 51–59. <https://doi.org/10.1007/BF02457966>
- Nudelman, A. E. et Landers, C. E. (1972). The failure of 100 divided by 3 to equal 33-1/3. *The American Sociologist*, 7(9), 9. Repéré à <http://www.jstor.org/stable/27702041>
- Office de la propriété intellectuelle du Canada (2017). Qu'est-ce qu'un brevet? Repéré à <http://www.ic.gc.ca/eic/site/cipointernet-internetopic.nsf/fra/wr03716.html>
- Osborne, J. et Holland, A. (2009). What is authorship, and what should it be? A survey of prominent guidelines for determining authorship in scientific publications. *Practical Assessment, Research & Evaluation*, 14(15). Repéré à <http://pareonline.net/getvn.asp?v=14&n=15>

- Owen-Smith, J. et Powell, W. W. (2003). The expanding role of university patenting in the life sciences: Assessing the importance of experience and connectivity. *Research Policy*, 32(9), 1695–1711. [https://doi.org/10.1016/S0048-7333\(03\)00045-3](https://doi.org/10.1016/S0048-7333(03)00045-3)
- Packer, K. et Webster, A. (1996). Patenting culture in science: Reinventing the scientific wheel of credibility. *Science, Technology & Human Values*, 21(4), 427–453. <https://doi.org/10.1177/016224399602100403>
- Paul-Hus, A., Mongeon, P., Sainte-Marie, M. et Larivière, V. (2017). The sum of it all: Revealing collaboration patterns by combining authorship and acknowledgements. *Journal of Informetrics*, 11(1), 80–87. <https://doi.org/10.1016/j.joi.2016.11.005>
- Peberdy, M. et Strowel, A. (2010). Employee’s rights to compensation for inventions – A European perspective. Dans *PLC Cross-border Life Sciences Handbook 2009/10* (p. 63–70). Londres : Practical Law Company.
- Perneger, T. V, Poncet, A., Carpentier, M., Agoritsas, T., Combescure, C. et Gayet-Ageron, A. (2017). Thinker, Soldier, Scribe: Cross-sectional study of researchers’ roles and author order in the Annals of Internal Medicine. *BMJ Open*, 7(6), e013898. <https://doi.org/10.1136/bmjopen-2016-013898>
- Pickard, A. J. (2007). *Research methods in information*. Londres : Facet publishing.
- Pinto, L. (s. d.). Champ, sociologie. Dans Universalis Éducation (en ligne), *Encyclopaedia Universalis*. Repéré à <http://www.universalis-edu.com/encyclopedie/champ-sociologie/>
- Pontille, D. (2004). *La signature scientifique : une sociologie pragmatique de l’attribution*. Paris : CNRS.
- Pontille, D. (2006). Qu’est-ce qu’un auteur scientifique? *Sciences de la société*, 67, 77–93. Repéré à <https://halshs.archives-ouvertes.fr/halshs-00261793/document>
- Pontille, D. (2016). *Signer ensemble : Contribution et évaluation en sciences*. Paris : Economica.
- Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation*, 25(4), 348–349.
- Rennie, D. et Flanagan, A. (1994). Authorship! Authorship! Guests, ghosts, grafters, and the two-sided coin. *JAMA*, 271(6), 469. <https://doi.org/10.1001/jama.1994.03510300075043>
- Rennie, D., Flanagan, A. et Yank, V. (2000). The contributions of authors. *JAMA*, 284(1), 89–91. <https://doi.org/10.1001/jama.284.1.89>

- Rennie, D. et Yank, V. (1998). If authors became contributors, everyone would gain, especially the reader. *American Journal of Public Health*, 88(5), 828–830. <https://doi.org/10.2105/AJPH.88.5.828>
- Rennie, D., Yank, V. et Emanuel, L. (1997). When authorship fails: A proposal to make contributors accountable. *JAMA*, 278(7), 579. <https://doi.org/10.1001/jama.1997.03550070071041>
- Resnik, D. B. et Master, Z. (2011). Authorship policies of bioethics journals. *Journal of Medical Ethics*, 37(7), 424–428. <https://doi.org/10.1136/jme.2010.040675>
- Rizzi, E. (2013). La régression logistique. Dans G. Masuy-Stroobant et R. Costa (Eds.), *Analyser les données en sciences sociales : De la préparation des données à l'analyse multivarié* (p. 253–278). New York, NY : Peter Lang.
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 351. <https://doi.org/10.2307/2087176>
- Rossiter, M. W. (1993). The Matthew Matilda Effect in science. *Social Studies of Science*, 23(2), 325–341. <https://doi.org/10.1177/030631293023002004>
- Sanberg, P. R., Gharib, M., Harker, P. T., Kaler, E. W., Marchase, R. B., Sands, T. D., ... Sarkar, S. (2014). Changing the academic culture: Valuing patents and commercialization toward tenure and career advancement. *Proceedings of the National Academy of Sciences*, 111(18), 6542–6547. <https://doi.org/10.1073/pnas.1404094111>
- Sauermann, H., & Haeussler, C. (2017). Authorship and contribution disclosures. *Science Advances*, 3(11), e1700404. <https://doi.org/10.1126/sciadv.1700404>
- Schoumaker, B. (2013). La régression linéaire multiple. Dans G. Masuy-Stroobant et R. Costa (Eds.), *Analyser les données en sciences sociales : De la préparation des données à l'analyse multivarié* (p. 227–252). New York, NY : Peter Lang.
- Seymore, S. B. (2006). My patent, your patent, or our patent? Inventorship disputes within academic research groups. *Albany Law Journal of Science and Technology*, 16, 125–167. Repéré à <https://ssrn.com/abstract=976397>
- Shadish, W. R., Cook, T. D. et Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA : Houghton Mifflin.
- Shapin, S. (1989). The invisible technician. *American Scientist*, 77(6), 554–563. <https://doi.org/10.2307/27856006>
- Sheiness, D. et Canady, K. (2006). The importance of getting inventorship right. *Nature Biotechnology*, 24(2), 153–154. <https://doi.org/10.1038/nbt0206-153>

- Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature*, 495(7439), 22–24. <https://doi.org/10.1038/495022a>
- Shibayama, S., Walsh, J. P. et Baba, Y. (2012). Academic entrepreneurship and exchange of scientific resources: Material transfer in life and materials sciences in Japanese universities. *American Sociological Review*, 77(5), 804–830. <https://doi.org/10.1177/0003122412452874>
- Silipigni Connaway, L. et Radford, M. L. (2017). *Research methods in library and information science* (6^e éd.). Santa Barbara, CA : Libraries Unlimited.
- Sismondo, S. (2009). Ghosts in the machine: Publication planning in the medical sciences. *Social Studies of Science*, 39(2), 171–198. <https://doi.org/10.1177/0306312708101047>
- Smith, E., et Williams-Jones, B. (2012). Authorship and responsibility in health sciences research: A review of procedures for fairly allocating authorship in multi-author studies. *Science and Engineering Ethics*, 18(2), 199–212. <https://doi.org/10.1007/s11948-011-9263-5>
- Sproull, N. L. (1988). *Handbook of research methods: A guide for practitioners and students in the social sciences*. Metuchen, NJ : Scarecrow Press.
- Steen, R. G. (2011). Misinformation in the medical literature: What role do error and fraud play? *Journal of Medical Ethics*, 37(8), 498–503. <https://doi.org/10.1136/jme.2010.041830>
- Steen, R. G., Casadevall, A. et Fang, F. C. (2013). Why has the number of scientific retractions increased? *PLoS ONE*, 8(7), e68397. <https://doi.org/10.1371/journal.pone.0068397>
- Stephan, P. E. et El-Ganainy, A. (2007). The entrepreneurial puzzle: Explaining the gender gap. *The Journal of Technology Transfer*, 32(5), 475–487. <https://doi.org/10.1007/s10961-007-9033-3>
- Stevens, A. J., Johnson, G. A. et Sanberg, P. R. (2011). The role of patents and commercialization in the tenure and promotion process. *Technology & Innovation*, 13(3), 241–248. <https://doi.org/10.3727/194982411X13189742259479>
- Subramanyam, K. (1983). Bibliometric studies of research collaboration: A review. *Journal of Information Science*, 6(1), 33–38. <https://doi.org/10.1177/016555158300600105>
- Sugimoto, C. R., Robinson-Garcia, N., Murray, D. S., Yegros-Yegros, A., Costas, R. et Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, 550(7674), 29–31. <https://doi.org/10.1038/550029a>
- Sugimoto, C. R., Ni, C., West, J. D. et Larivière, V. (2015). The academic advantage: Gender disparities in patenting. *PLoS ONE*, 10(5), e0128000. <https://doi.org/10.1371/journal.pone.0128000>

- Tabachnick, B. G. et Fidell, L. S. (2013). *Using multivariate statistics* (6^e éd.). Boston, MA : Pearson Education.
- Thursby, J., Fuller, A. W. et Thursby, M. (2009). US faculty patenting: Inside and outside the university. *Research Policy*, 38(1), 14–25. <https://doi.org/10.1016/j.respol.2008.09.004>
- van Eck, N. J. et Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, 7(3), 50–54.
- Verbeek, A., Debackere, K., Luwel, M., Andries, P., Zimmermann, E. et Deleus, F. (2002). Linking science to technology: Using bibliographic references in patents to build linkage schemes. *Scientometrics*, 54(3), 399–420. <https://doi.org/10.1023/A:1016034516731>
- von Proff, S. (2011). Patent and publication productivity of German professors: A life cycle view. *International Journal of Technology Transfer and Commercialisation*, 10(3/4), 392. <https://doi.org/10.1504/IJTTC.2011.040897>
- Wager, E. (2007a). Authors, ghosts, damned lies, and statisticians. *PLoS Medicine*, 4(1), e34. <https://doi.org/10.1371/journal.pmed.0040034>
- Wager, E. (2007b). Do medical journals provide clear and consistent guidelines on authorship? *MedGenMed: Medscape General Medicine*, 9(3), 16. Repéré dans Pubmed : <https://www.ncbi.nlm.nih.gov/pubmed/18092023>
- Wager, E. (2009). Recognition, reward and responsibility: Why the authorship of scientific papers matters. *Maturitas*, 62(2), 109–112. <https://doi.org/10.1016/j.maturitas.2008.12.001>
- Wager, E., Fiack, S., Graf, C., Robinson, A. et Rowlands, I. (2009). Science journal editors' views on publication ethics: Results of an international survey. *Journal of Medical Ethics*, 35(6), 348–353. <https://doi.org/10.1136/jme.2008.028324>
- Walsh, J. P. et Lee, Y.-N. (2015). The bureaucratization of science. *Research Policy*, 44(8), 1584–1600. <https://doi.org/10.1016/j.respol.2015.04.010>
- Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, 6(4), 700–711. <https://doi.org/10.1016/j.joi.2012.07.008>
- Weiner, C. (1987). Patenting and academic research: Historical case studies. *Science, Technology & Human Values*, 12(1), 50–62. <https://doi.org/10.1177/016224398701200105>
- West, J. D., Jacquet, J., King, M. M., Correll, S. J. et Bergstrom, C. T. (2013). The role of gender in scholarly authorship. *PLoS ONE*, 8(7), e66212. <https://doi.org/10.1371/journal.pone.0066212>

- Whitley, R. (2000). *The intellectual and social organization of the sciences* (2^e éd.). New York, NY : Oxford University Press.
- Williams, W. M. et Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, 112(17), 5360–5365. <https://doi.org/10.1073/pnas.1418878112>
- Wislar, J. S., Flanagin, A., Fontanarosa, P. B. et DeAngelis, C. D. (2011). Honorary and ghost authorship in high impact biomedical journals: A cross sectional survey. *BMJ*, 343, d6128–d6128. <https://doi.org/10.1136/bmj.d6128>
- Wong, P. K. et Singh, A. (2010). University patenting activities and their link to the quantity and quality of scientific publications. *Scientometrics*, 83(1), 271–294. <https://doi.org/10.1007/s11192-009-0003-4>
- Wuchty, S., Jones, B. B. F. et Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036–1039. <https://doi.org/10.1126/science.1136099>
- Zuckerman, H. (1967). Nobel laureates in science: Patterns of productivity, collaboration, and authorship. *American Sociological Review*, 32(3), 391–403. <https://doi.org/10.2307/2091086>
- Zuckerman, H. A. (1968). Patterns of name ordering among authors of scientific papers: A study of social symbolism and its ambiguity. *American Journal of Sociology*, 74(3), 276–291. <https://doi.org/10.2307/2775535>
- Zuckerman, H. A. (1970). Stratification in American science. *Sociological Inquiry*, 40(2), 235–257. <https://doi.org/10.1111/j.1475-682X.1970.tb01010.x>
- Zuckerman, H. A. (1977). *Scientific elite: Nobel laureates in the United States*. New York, NY : Free Press.
- Zuckerman, H. A. (1998). Accumulation of advantage and disadvantage: The theory and its intellectual biography. Dans C. Mongardini et S. Tabboni (dir.), *Robert K. Merton and contemporary sociology* (p. 139–161). New Brunswick, NJ : Transaction Publishers.

Bibliographie – documents légaux

Législation

Canada

Loi sur les brevets, LRC 1985, c. P-4, art 1.

États-Unis

Patents, 35 USC §§ 100, 116, 256.

US Const art I § 8.

France

Code de la propriété intellectuelle, art R611(15).

Royaume-Uni

Patents Act 1977 (R-U), c 37, art 3.

Jurisprudence

États-Unis

Burroughs Wellcome Co v Barr Labs Inc, 40 F.3d 1223, 32 USPQ2d 1915 (Fed Cir 1994).

DeBaun, In re, 687 F.2d 459, 214 USPQ 933 (CCPA 1982).

Eli Lilly & Co v Aradigm Corp, 376 F.3d 1352 (Fed Cir 2004).

Fritsch v Lin, 21 USPQ2d 1731 (Bd Pat App & Inter 1991).

Hardee, In re, 223 USPQ 1122 (Comm'r Pat 1984).

Katz, In re, 687 F.2d 450 (CCPA 1982).

Kimberly-Clark Corp v Procter & Gamble Distrib Co, 973 F.2d 911, 23 USPQ2d 1921 (Fed Cir 1992).

Moler v Purdy, 131 USPQ 276 (Bd Pat Inter 1960).

Smernoff, Ex parte, 215 USPQ 545 (Bd App 1982).

Stark v Advanced Magnetics Inc, 119 F.3d 1551, 1556 (Fed Cir 1997).

Royaume-Uni

Pope Appliance Corp v Spanish River Pulp and Paper Mills, [1929] AC 269.

Annexe 1. Extrait des recommandations de l'International Committee of Medical Journal Editors

Defining the Role of Authors and Contributors

1. Why Authorship Matters

Authorship confers credit and has important academic, social, and financial implications. Authorship also implies responsibility and accountability for published work. The following recommendations are intended to ensure that contributors who have made substantive intellectual contributions to a paper are given credit as authors, but also that contributors credited as authors understand their role in taking responsibility and being accountable for what is published.

Because authorship does not communicate what contributions qualified an individual to be an author, some journals now request and publish information about the contributions of each person named as having participated in a submitted study, at least for original research. Editors are strongly encouraged to develop and implement a contributorship policy, as well as a policy that identifies who is responsible for the integrity of the work as a whole. Such policies remove much of the ambiguity surrounding contributions, but leave unresolved the question of the quantity and quality of contribution that qualify an individual for authorship. The ICMJE has thus developed criteria for authorship that can be used by all journals, including those that distinguish authors from other contributors.

2. Who Is an Author?

The ICMJE recommends that authorship be based on the following 4 criteria:

- 1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
- 2) Drafting the work or revising it critically for important intellectual content; AND
- 3) Final approval of the version to be published; AND
- 4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

In addition to being accountable for the parts of the work he or she has done, an author should be able to identify which co-authors are responsible for specific other parts of the work. In addition, authors should have confidence in the integrity of the contributions of their co-authors.

All those designated as authors should meet all four criteria for authorship, and all who meet the four criteria should be identified as authors. Those who do not meet all four criteria should be

acknowledged—see Section II.A.3 below. These authorship criteria are intended to reserve the status of authorship for those who deserve credit and can take responsibility for the work. The criteria are not intended for use as a means to disqualify colleagues from authorship who otherwise meet authorship criteria by denying them the opportunity to meet criterion #s 2 or 3. Therefore, all individuals who meet the first criterion should have the opportunity to participate in the review, drafting, and final approval of the manuscript.

The individuals who conduct the work are responsible for identifying who meets these criteria and ideally should do so when planning the work, making modifications as appropriate as the work progresses. It is the collective responsibility of the authors, not the journal to which the work is submitted, to determine that all people named as authors meet all four criteria; it is not the role of journal editors to determine who qualifies or does not qualify for authorship or to arbitrate authorship conflicts. If agreement cannot be reached about who qualifies for authorship, the institution(s) where the work was performed, not the journal editor, should be asked to investigate. If authors request removal or addition of an author after manuscript submission or publication, journal editors should seek an explanation and signed statement of agreement for the requested change from all listed authors and from the author to be removed or added.

The corresponding author is the one individual who takes primary responsibility for communication with the journal during the manuscript submission, peer review, and publication process, and typically ensures that all the journal's administrative requirements, such as providing details of authorship, ethics committee approval, clinical trial registration documentation, and gathering conflict of interest forms and statements, are properly completed, although these duties may be delegated to one or more coauthors. The corresponding author should be available throughout the submission and peer review process to respond to editorial queries in a timely way, and should be available after publication to respond to critiques of the work and cooperate with any requests from the journal for data or additional information should questions about the paper arise after publication. Although the corresponding author has primary responsibility for correspondence with the journal, the ICMJE recommends that editors send copies of all correspondence to all listed authors.

When a large multi-author group has conducted the work, the group ideally should decide who will be an author before the work is started and confirm who is an author before submitting the manuscript for publication. All members of the group named as authors should meet all four criteria for authorship, including approval of the final manuscript, and they should be able to take public responsibility for the work and should have full confidence in the accuracy and integrity of the

work of other group authors. They will also be expected as individuals to complete conflict-of-interest disclosure forms.

Some large multi-author groups designate authorship by a group name, with or without the names of individuals. When submitting a manuscript authored by a group, the corresponding author should specify the group name if one exists, and clearly identify the group members who can take credit and responsibility for the work as authors. The byline of the article identifies who is directly responsible for the manuscript, and MEDLINE lists as authors whichever names appear on the byline. If the byline includes a group name, MEDLINE will list the names of individual group members who are authors or who are collaborators, sometimes called non-author contributors, if there is a note associated with the byline clearly stating that the individual names are elsewhere in the paper and whether those names are authors or collaborators.

3. Non-Author Contributors

Contributors who meet fewer than all 4 of the above criteria for authorship should not be listed as authors, but they should be acknowledged. Examples of activities that alone (without other contributions) do not qualify a contributor for authorship are acquisition of funding; general supervision of a research group or general administrative support; and writing assistance, technical editing, language editing, and proofreading. Those whose contributions do not justify authorship may be acknowledged individually or together as a group under a single heading (e.g. "Clinical Investigators" or "Participating Investigators"), and their contributions should be specified (e.g., "served as scientific advisors," "critically reviewed the study proposal," "collected data," "provided and cared for study patients", "participated in writing or technical editing of the manuscript").

Because acknowledgment may imply endorsement by acknowledged individuals of a study's data and conclusions, editors are advised to require that the corresponding author obtain written permission to be acknowledged from all acknowledged individuals.

Source : <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>

Annexe 2. Exemple de description des contributions individuelles des coauteurs d'un article publié dans le *Journal of the American Medical Association*

Article Information

[Back to top](#)

Corresponding Author: Matthew M. Churpek, MD, MPH, PhD, Section of Pulmonary and Critical Care Medicine, Department of Medicine, University of Chicago, 5841 S Maryland Ave, MC 6076, Chicago, IL 60637 (matthew.churpek@uchospitals.edu).

Author Contributions: Drs Volchenbom and Churpek had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Volchenbom, Edelson, Howell, Churpek.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Volchenbom, Mayampurath, Göksu-Gürsoy.

Critical revision of the manuscript for important intellectual content: Volchenbom, Mayampurath, Edelson, Howell, Churpek.

Statistical analysis: Mayampurath, Göksu-Gürsoy, Howell, Churpek.

Administrative, technical, or material support: Volchenbom, Edelson, Howell, Churpek.

Annexe 3. Exemple de brevet (première page seulement)



US008664095B2

(12) **United States Patent**
Javey et al.

(10) **Patent No.:** **US 8,664,095 B2**
(45) **Date of Patent:** **Mar. 4, 2014**

(54) **BLACK GE BASED ON
CRYSTALLINE/AMORPHOUS CORE/SHELL
NANONEEDLE ARRAYS**

(75) Inventors: **Ali Javey**, Emeryville, CA (US);
Yu-Lun Chueh, Hsinchu (TW);
Zhiyong Fan, Clear Water Bay (HK)

(73) Assignee: **The Regents of the University of
California**, Oakland, CA (US)

(*) Notice: Subject to any disclaimer, the term of this
patent is extended or adjusted under 35
U.S.C. 154(b) by 0 days.

(21) Appl. No.: **13/332,743**

(22) Filed: **Dec. 21, 2011**

(65) **Prior Publication Data**
US 2012/0161290 A1 Jun. 28, 2012

Related U.S. Application Data

(60) Provisional application No. 61/426,166, filed on Dec.
22, 2010.

(51) **Int. Cl.**
H01L 21/20 (2006.01)

(52) **U.S. Cl.**
USPC **438/479**; 257/613; 257/616

(58) **Field of Classification Search**
USPC 257/76, 616, E21.126, E21.132,
257/E33.068, E51.02, 613; 438/479
See application file for complete search history.

(56) **References Cited**

PUBLICATIONS

Yu-Lun Chueh, Zhiyong Fan, Kuniharu Takei, Hyunhyub Ko, Rehan
Kapadia, Asghar A. Rathore, Nate Miller, Kyoungsik Yu, Ming Wu,
E. E. Haller and Ali Javey. "Black Ge Based on Crystalline/Amor-
phous Core/Shell Nanoneedle Arrays," Nano Lett., 2010, 10 (2), pp.
520-523, Publication Date (Web): Dec. 30, 2009.*

* cited by examiner

Primary Examiner — Yu-Hsi D Sun

Assistant Examiner — Christina Sylvia

(74) *Attorney, Agent, or Firm* — Lawrence Berkeley
National Laboratory

(57) **ABSTRACT**

Direct growth of black Ge on low-temperature substrates,
including plastics and rubber is reported. The material is
based on highly dense, crystalline/amorphous core/shell Ge
nanoneedle arrays with ultrasharp tips (~4 nm) enabled by the
Ni catalyzed vapor-solid-solid growth process. Ge nanonee-
dle arrays exhibit remarkable optical properties. Specifi-
cally, minimal optical reflectance (<1%) is observed, even for
high angles of incidence (~75°) and for relatively short
nanoneedle lengths (~1 μm). Furthermore, the material
exhibits high optical absorption efficiency with an effective
band gap of ~1 eV. The reported black Ge can have important
practical implications for efficient photovoltaic and photode-
tector applications on nonconventional substrates.

15 Claims, 8 Drawing Sheets

Annexe 4. Nom, description et source des champs de la base de données SQL utilisée pour la recherche

Table	Champ	Description	Source*
Article	id_article	Identifiant unique de l'article	WoS
	titre_article	Titre de l'article	WoS
	résumé_article	Résumé de l'article	WoS
	année_pub	Année de publication	WoS
	nb_auteurs	Nombre d'auteurs de l'article	OST
	Revue	Revue de publication	WoS
	disc_revue	Discipline de la revue	OST
	gr_disc	Groupe disciplinaire de l'article (recherche biomédicale, sciences naturelles, génie)	CP
	type_alpha	Catégorie d'ordre alphabétique de l'article (complet, partiel, autre)	CP
	fi_2ans	Facteur d'impact de la revue calculé sur une période de 2 ans	OST
multi_institution	Variable dichotomique prenant la valeur 1 si l'article est une collaboration interinstitutionnelle, et 0 sinon	OST	
Brevet	id_brevet	Identifiant unique du brevet	PatFT
	titre_brevet	Titre du brevet	PatFT
	résumé_brevet	Résumé du brevet	PatFT
	année_dépôt	Année de dépôt de la demande de brevet	PatFT
	nb_inventeur	Nombre d'inventeurs sur le brevet	PatFT
Syntagmes nominaux articles	id_article	Identifiant unique de l'article	WoS
	syntagme_nominal	Syntagme nominal	CWTS
	nb_occurrences	Nombre d'occurrences du syntagme dans le titre et le résumé de l'article	CWTS
Syntagmes nominaux brevets	id_brevet	Identifiant unique du brevet	PatFT
	syntagme_nominal	Syntagme nominal	CWTS
	nb_occurrences	Nombre d'occurrences du syntagme dans le titre et le résumé du brevet	CWTS
Paires brevet-article	id_pba	Identifiant unique de la PBA	CP
	id_article	Identifiant unique de l'article	WoS
	id_brevet	Identifiant unique du brevet	PatFT
	similarité	Score de similarité	CP
	nb_auteurs_inventeurs	Nombre d'auteurs de l'article figurant parmi les inventeurs du brevet	CP
	ratio_inventeurs_auteurs	Proportion des auteurs de l'article figurant parmi les inventeurs du brevet	CP

* WoS = Champs disponibles tels quels dans les notices bibliographiques du WoS, rendus disponibles par l'OST.

PatFT = Champs disponibles tels quels dans les notices de la PatFT, rendus disponibles par l'OST.

OST = Champs créés et rendus disponibles par l'OST.

CWTS = Champs créés et rendus disponibles par le CWTS.

Nous = Champs créés à l'issue d'un processus de collecte effectué par le chercheur principal ou une assistante de recherche.

Annexe 4. Nom, description et source des champs de la base de données SQL utilisée pour la recherche (suite)

Table	Champ	Description	Source*
Auteur	id_auteur	Identifiant unique de l'auteur	CWTS
	id_article	Identifiant unique de l'article	WoS
	nom_initiales	Nom de famille et initiales de l'auteur	WoS
	prénom	Prénom de l'auteur	WoS
	sexe	Sexe de l'auteur	OST**
	nb_citations	Nombre de citations reçues par l'auteur en date de l'année de dépôt du brevet	OST/CWTS
	ordre_auteur	Position de l'auteur dans la liste des auteurs de l'article (variable numérique)	WoS
	position	Position de l'auteur dans la liste la liste des auteurs de l'article (premier, deuxième, troisième, milieu, antépénultième, pénultième, dernier)	CP
	rôle	Rôle de l'auteur dans la recherche (principal, périphérique, superviseur)	CP
	analyse	Variable dichotomique dont la valeur est 1 si l'auteur a contribué à l'analyse, et 0 sinon	AR
	conception	Variable dichotomique dont la valeur est 1 si l'auteur a contribué à la conception, et 0 sinon	AR
	matériel	Variable dichotomique dont la valeur est 1 si l'auteur a contribué à la fourniture de matériel, et 0 sinon	AR
	réalisation	Variable dichotomique dont la valeur est 1 si l'auteur a contribué à la réalisation, et 0 sinon	AR
	autres	Variable dichotomique dont la valeur est 1 si l'auteur a contribué à d'autres types de tâche, et 0 sinon	AR
inventeur	Variable dichotomique dont la valeur est 1 si l'auteur figure parmi les inventeurs du brevet, et 0 sinon	CP	
Inventeur	id_brevet	Identifiant unique du brevet	PatFT
	nom_prenom	Nom et prénom de l'inventeur	PatFT

* WoS = Champs disponibles tels quels dans les notices bibliographiques du WoS, rendus disponibles par l'OST.

PatFT = Champs disponibles tels quels dans les notices de la PatFT, rendus disponibles par l'OST.

OST = Champs créés et rendus disponibles par l'OST.

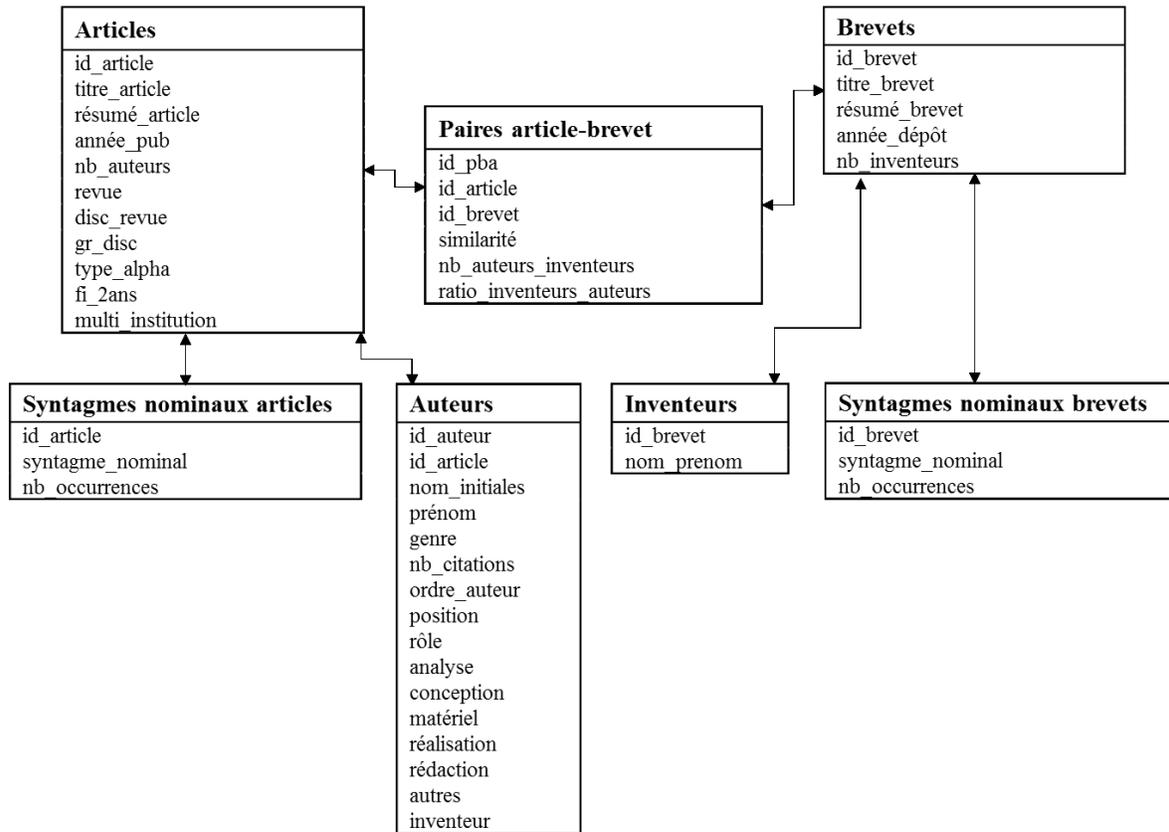
CWTS = Champs créés et rendus disponibles par le CWTS.

CP = Champs créés à l'issu d'un processus de collecte effectué par le chercheur principal.

AR = Champs créés à l'issu d'un processus de collecte effectué par une assistante de recherche.

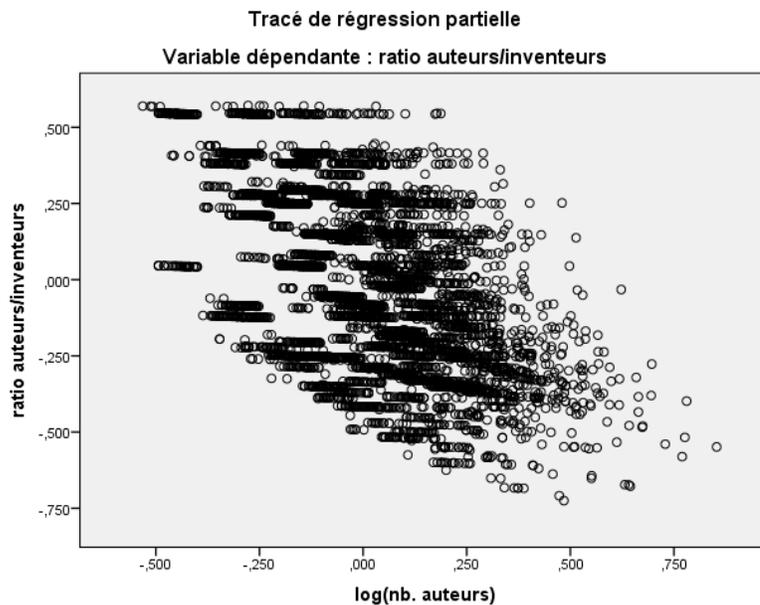
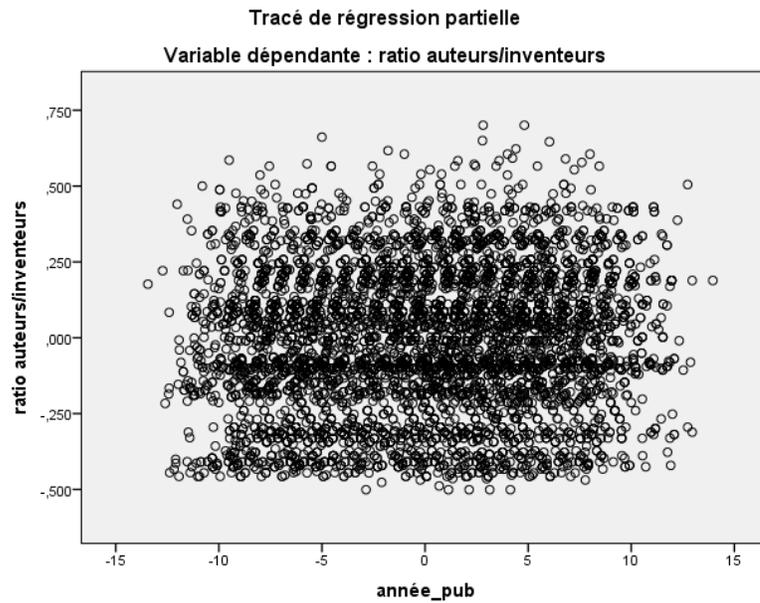
** Les données utilisées pour l'attribution du sexe ont été générées par l'algorithme de Bérubé et al. (en préparation) décrit au Chapitre 2 et ont été obtenues à partir de la base de données de l'OST.

Annexe 5. Schéma de la base de données relationnelle utilisée pour la recherche



Annexe 6. Tableaux et figures pour la vérification du respect des postulats de la régression linéaire

Linéarité de la relation entre la variable prédite et les variables prédictives



Absence de multicollinéarité

Coefficients^a

Modèle	B	ES	Bêta	t	Sig.	Statistiques de colinéarité	
						Tolérance	VIF
1 (Constante)	-5,837	1,043		-5,594	,000		
année_pub	,003	,001	,063	6,592	,000	,944	1,060
log(nb. auteurs)	-,778	,014	-,596	-53,679	,000	,709	1,410
haut_fis	,041	,010	,039	4,118	,000	,972	1,029
Génie	,103	,008	,134	12,160	,000	,719	1,391
sc_nat	,073	,007	,109	10,403	,000	,793	1,261
Multi_institutions	-,014	,007	-,023	-2,205	,027	,824	1,213

a. Variable dépendante : ratio auteurs/inventeurs

Corrélations du coefficient^a

Modèle		multi-institution	année_pub	haut_fi2	sc_nat	génie	log(nb. auteurs)	
1	Corrélations	multi-institution	1,000	-,063	-,036	,105	,064	-,331
		année_pub	-,063	1,000	-,049	-,105	-,201	-,130
		haut_fi2	-,036	-,049	1,000	-,043	-,051	-,125
		sc_nat	,105	-,105	-,043	1,000	,427	,242
		génie	,064	-,201	-,051	,427	1,000	,358
		log(nb. auteurs)	-,331	-,130	-,125	,242	,358	1,000

a. Variable dépendante : ratio auteurs/inventeurs

Indépendance des erreurs (statistique de Durbin-Watson)

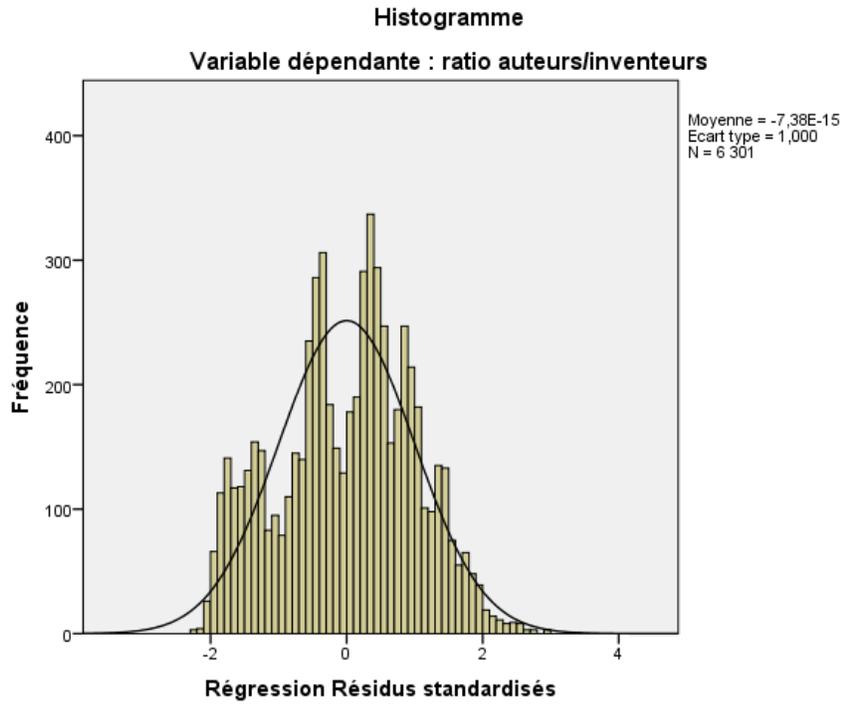
Récapitulatif des modèles^b

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation	Variation de R-deux	Modifier les statistiques			Sig. de F	Durbin-Watson
						Variation de F	ddl1	ddl2		
1	,670 ^a	,449	,449	,230320	,449	856,123	6	6294	,000	1,326

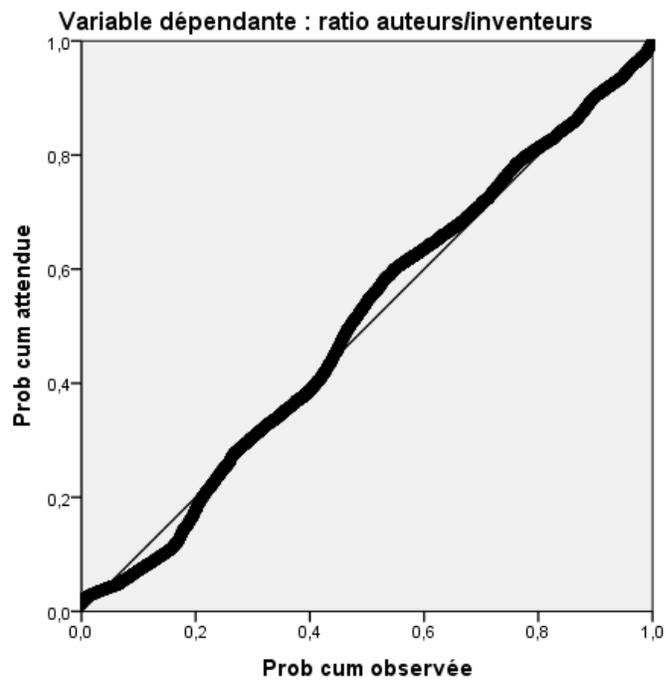
a. Prédicteurs : (Constante), multi-institution, année_pub, haut_fi2, sc_nat, génie, log(nb. auteurs)

b. Variable dépendante : ratio auteurs/inventeurs

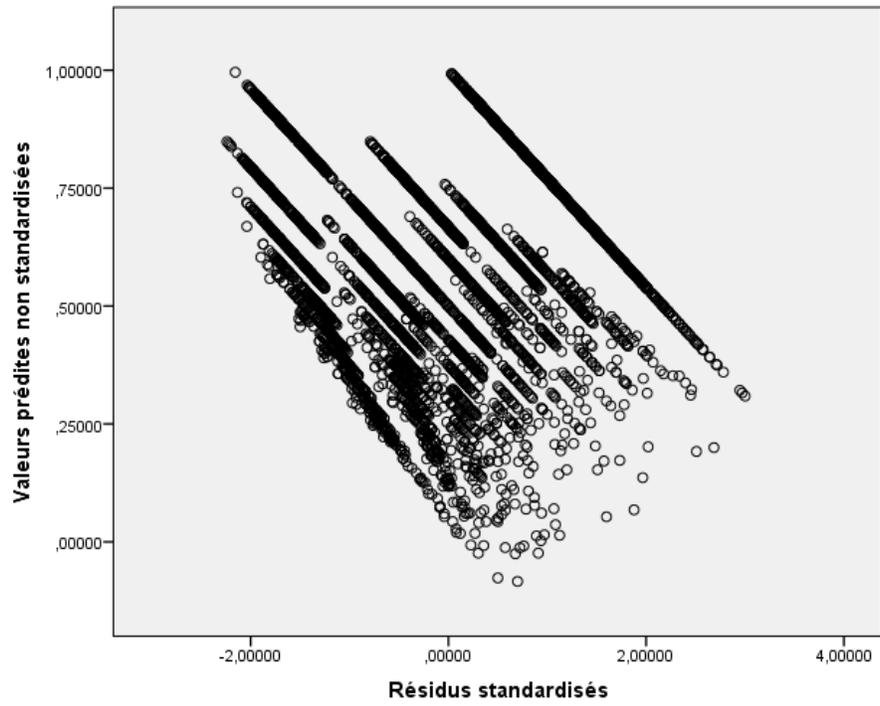
Distribution normale des résidus



Tracé P-P normal de régression Résidus standardisés



Homocédasticité



Annexe 7. Tableaux de corrélation pour la vérification des postulats de la régression logistique

Modèle de la section 3.2.1.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. inventeur	1,000	0,290	-0,013	-0,131	-0,151	-0,057	0,271	-0,344	0,184	-0,266	0,154	0,140	0,145
2. premier	0,290	1,000	-0,202	-0,140	-0,143	-0,203	-0,257	-0,227	0,106	-0,128	0,060	0,012	-0,130
3. deuxième	-0,013	-0,202	1,000	-0,111	-0,113	-0,161	-0,204	-0,180	0,015	-0,028	0,020	-0,035	-0,103
4. troisième	-0,131	-0,140	-0,111	1,000	-0,079	-0,112	-0,142	-0,125	-0,057	0,055	-0,018	-0,046	-0,061
5. antépénultième	-0,151	-0,143	-0,113	-0,079	1,000	-0,114	-0,144	-0,127	-0,057	0,063	-0,026	-0,004	0,006
6. pénultième	-0,057	-0,203	-0,161	-0,112	-0,114	1,000	-0,205	-0,181	0,004	-0,023	0,022	0,042	0,065
7. dernier	0,271	-0,257	-0,204	-0,142	-0,144	-0,205	1,000	-0,229	0,092	-0,113	0,055	0,107	0,240
8. milieu	-0,344	-0,227	-0,180	-0,125	-0,127	-0,181	-0,229	1,000	-0,153	0,225	-0,133	-0,101	-0,047
9. génie	0,184	0,106	0,015	-0,057	-0,057	0,004	0,092	-0,153	1,000	-0,486	-0,228	0,079	0,000
10. biomed	-0,266	-0,128	-0,028	0,055	0,063	-0,023	-0,113	0,225	-0,486	1,000	-0,740	-0,144	0,000
11. sc_nat	0,154	0,060	0,020	-0,018	-0,026	0,022	0,055	-0,133	-0,228	-0,740	1,000	0,100	0,000
12. homme	0,140	0,012	-0,035	-0,046	-0,004	0,042	0,107	-0,101	0,079	-0,144	0,100	1,000	0,142
13. hautement cité	0,145	-0,130	-0,103	-0,061	0,006	0,065	0,240	-0,047	0,000	0,000	0,000	0,142	1,000

Modèle de la section 3.2.2.

Variables	1	2	3	4	5	6	7	8	9
1. inventeur	1,000	0,357	0,291	-0,485	0,266	-0,334	0,127	0,146	0,118
2. premier	0,357	1,000	-0,102	-0,730	0,576	-0,375	-0,138	0,034	-0,021
3. dernier	0,291	-0,102	1,000	-0,606	-0,162	-0,312	0,536	0,116	0,224
4. autres	-0,485	-0,730	-0,606	1,000	-0,349	0,514	-0,258	-0,107	-0,137
5. principal	0,266	0,576	-0,162	-0,349	1,000	-0,600	-0,302	0,071	-0,068
6. périphérique	-0,334	-0,375	-0,312	0,514	-0,600	1,000	-0,581	-0,134	-0,102
7. superviseur	0,127	-0,138	0,536	-0,258	-0,302	-0,581	1,000	0,088	0,191
8. homme	0,146	0,034	0,116	-0,107	0,071	-0,134	0,088	1,000	0,156
9. hautement cité	0,118	-0,021	0,224	-0,137	-0,068	-0,102	0,191	0,156	1,000

Modèle de la section 3.2.3.

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. inventeur	1,000	0,310	-0,599	0,478	0,277	0,484	0,033	-0,114	0,324	0,004	0,158	0,086
2. premier	0,310	1,000	-0,648	-0,131	0,258	0,298	-0,045	0,174	0,243	0,045	0,013	-0,151
3. autres	-0,599	-0,648	1,000	-0,670	-0,282	-0,502	0,081	0,099	-0,394	-0,086	-0,121	-0,101
4. derniers	0,478	-0,131	-0,670	1,000	0,115	0,363	-0,062	-0,299	0,277	0,067	0,144	0,279
5. analyse	0,277	0,258	-0,282	0,115	1,000	0,309	-0,123	-0,110	0,237	-0,068	0,117	0,005
6. conception	0,484	0,298	-0,502	0,363	0,309	1,000	-0,014	-0,177	0,348	-0,058	0,130	0,119
7. matériel	0,033	-0,045	0,081	-0,062	-0,123	-0,014	1,000	-0,124	-0,124	-0,070	0,067	0,141
8. réalisation	-0,114	0,174	0,099	-0,299	-0,110	-0,177	-0,124	1,000	-0,137	-0,152	-0,204	-0,364
9. rédaction	0,324	0,243	-0,394	0,277	0,237	0,348	-0,124	-0,137	1,000	0,175	0,095	0,095
10. autres	0,004	0,045	-0,086	0,067	-0,068	-0,058	-0,070	-0,152	0,175	1,000	0,005	0,058
11. homme	0,158	0,013	-0,121	0,144	0,117	0,130	0,067	-0,204	0,095	0,005	1,000	0,196
12. hautement cité	0,086	-0,151	-0,101	0,279	0,005	0,119	0,141	-0,364	0,095	0,058	0,196	1,000