



HAL
open science

Méthode quantitative utilisée dans les thèses portant sur les soins de santé primaires : étude rétrospective en Aquitaine sur 10 ans

Benjamin Ballon

► To cite this version:

Benjamin Ballon. Méthode quantitative utilisée dans les thèses portant sur les soins de santé primaires : étude rétrospective en Aquitaine sur 10 ans . Médecine humaine et pathologie. 2017. dumas-01551846

HAL Id: dumas-01551846

<https://dumas.ccsd.cnrs.fr/dumas-01551846>

Submitted on 30 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Bordeaux
U.F.R. DES SCIENCES MÉDICALES

Année 2017

Thèse n°104

Thèse pour l'obtention du
DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

Présentée et soutenue publiquement
le 26.06.2017

Par Benjamin BALLON
né le 21 juin 1984 à Bordeaux

**Méthode quantitative utilisée dans les thèses
portant sur les soins de santé primaires.**

Étude rétrospective en Aquitaine sur 10 ans.

Directeur de thèse Professeur Jean-Philippe JOSEPH

Membres du Jury :

Pr Dominique Dallay	Président
Pr Jean Saric	Juge
Pr Bernard Gay	Juge
Dr Albert Roche	Juge
Pr Jean-Philippe Joseph	Directeur
Dr Laurent Magot	Rapporteur

Document créé avec Pages® version 5.6.2 le 28 avril 2017

par

Benjamin BALLON



Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

<http://creativecommons.org/licenses/by/4.0/>

REMERCIEMENTS.

À mon Jury de thèse :

Je remercie monsieur le Professeur Dominique Dallay, PUPH gynécologue obstétricien, d'avoir participé à ma soutenance, comme président du jury. Vous m'avez fait un immense honneur.

Je remercie monsieur le Professeur Jean Saric, PUPH chirurgien viscéral, d'avoir participé à ma soutenance, comme membre du jury. Vous m'avez fait un immense honneur.

Je remercie monsieur le Professeur Bernard Gay, professeur émérite de médecine générale, d'avoir participé à ma soutenance, comme membre du jury. Vous m'avez fait un immense honneur.

Je remercie monsieur le Docteur Albert Roche, président du conseil département de l'ordre des médecins de la Gironde, médecin généraliste, d'avoir participé à ma soutenance, et de m'avoir aidé à composer mon jury. Vous m'avez fait un immense honneur.

Enfin, je remercie monsieur le Docteur Laurent Magot, d'avoir été mon rapporteur, je remercie tout particulièrement monsieur le Professeur Jean-Philippe JOSEPH, pour m'avoir guidé, tout d'abord durant le DES de médecine générale, puis tout au long de ce travail de thèse, jusqu'à la soutenance. Vous m'avez fait un immense honneur. Travailler avec vous fut une expérience extraordinaire.

Au DMG, à l'université, à tous ceux et celles qui m'ont permis d'aller aussi loin.

À mon père, qui nous a quitté trop tôt.

À ma famille qui m'a toujours soutenu, à ma mère, mon frère, à mes cousins et cousines.

À mes amis, à mes camarades, mes co-internes et à mes séniors durant mon internat.

À Catherine, qui m'a pris sous son aile pour mes premiers remplacements, et pour m'aider à me lancer seul dans la médecine de ville.

À Davina, l'amour de ma vie.

À tous ceux dont j'ai lu le travail.

À tous ceux qui liront le mien.

Table des matières

Préambule	1
INTRODUCTION	2
Question de recherche	3
Hypothèses	4
Objectif principal	4
Objectifs secondaires	4
MÉTHODES	5
Type d'étude	5
Population de l'étude	5
Technique de recueil des données	5
Mode de recueil initial des données exhaustives	6
Premier tri des données : déterminer le rapport aux soins de santé primaires à partir du titre	7
Deuxième tri des données : déterminer l'apport du titre comme déterminant le type de méthode de recherche employée	8
Troisième tri des données : détermination de la méthode basée sur le résumé, conclusion du type de méthode, relevé des variables des critères secondaires	10
Quatrième tri des données : exploration de la méthode des textes intégraux des thèses	18
Mode d'analyse des données	20
Aspects réglementaires et éthiques	21
RÉSULTATS	22
Population totale obtenue à partir des données d'APOGÉE	22
Diagramme de flux de l'étude	22
Population incluse « soins de santé primaires » à partir de la population totale avec répartition par année	23
Recours au directeur de thèse pour déterminer l'inclusion des thèses « douteuses »	23
Première étape : tri des titres des thèses par mots-clés	24
Deuxième étape : tri des thèses par lecture systématique du titre	
Comparaison aux résultats du tri par mots-clés	24
Troisième étape : tri des thèses par lecture du résumé avec conclusion sur la méthode de la thèse	25
Comparaison de la performance de chaque étape de tri	25
Détermination du nombre annuel des thèses classées « soins de santé primaires » de méthode quantitative ou mixte	26
Comparaison du nombre annuels de thèses classées « soins de santé primaires » de méthode quantitative/mixte et de méthode qualitative	27
Typologie Des méthodes quantitatives ou mixtes déduite des résumés	
Méthode principale relevée si plus d'une méthode	
Corrélation avec la déclaration du thésard	28
Paramètres statistiques des thèses de méthode quantitative ou mixte apparaissant dans le résumé	
Présentation en fonction de la typologie	30
Paramètres statistiques issus de la méthode après lecture du texte intégral des thèses	31

Recueil et effectifs des tests statistiques de variables quantitatives (continues) déclarés par l'auteur dans la méthode	32
Recueil et effectifs des test statistiques de variables qualitatives (discrétisées, booléennes) déclarés par l'auteur dans la méthode	33
Recueil et effectifs des logiciels d'analyse statistique utilisés, comparaison au recours à un tiers hors du DMG pour l'analyse statistique	34
Recherche d'une corrélation entre la catégorie du directeur de thèse et la déclaration du recours à un tiers pour l'analyse statistique par le thésard	35
Recherche d'une corrélation entre la catégorie du directeur de thèse et la présence de paramètres statistiques ou évocateurs d'analyse statistique dans le résumé	36
Autre calculs	37
DISCUSSION	38
Résultats principaux	38
Comparaison des résultats principaux avec les travaux antérieurs	42
Limites du travail de recherche	44
Biais	46
Perspectives	47
CONCLUSION	48
RÉFÉRENCES	49
TABLES ET INDEX	51
Table des tableaux	51
Table des illustrations et schémas	53
Table des représentations graphiques	53
Liste des abréviations	54
ANNEXES	55
Annexe 1. Algorithmes utilisés	56
Annexe 2. Aspects de SUDOC	59
Annexe 3. Classification des méthodes des thèses selon les critères de Paul FRAPPÉ	62
Annexe 4. Diagrammes de flux théoriques de saisie des données.	63
Annexe 5. Résultats complémentaires	66
Annexe 6. Comparaison avec les travaux de Maylis LABERNÈDE et Alix De CUSSY-BOUVET	75
SERMENT MÉDICAL	78

Préambule

Le travail de thèse est l'étape inévitable et indispensable qui vient conclure le parcours de l'étudiant en médecine (1), il s'agit d'un travail de recherche avec une méthode spécifique qui nécessite donc une formation en amont des étudiants, orientée par leurs besoins futurs.

En nous référant aux cours reçus durant notre formation de médecine générale, deux méthodes principales se distinguent, la recherche quantitative et la recherche qualitative, associées à deux visions totalement différentes du déroulement du travail de recherche.

Un premier état des lieux exhaustif a récemment recensé les thèses de médecine générale soutenues à l'université de Bordeaux entre 2007 et 2014 (2), et a déterminé la proportion de thèses dont la thématique portait sur les soins de santé primaires.

La recherche en médecine est en expansion de manière constante (3), et le secteur spécifique des soins de santé primaires ne déroge pas à cette règle. Le développement de la filière universitaire de médecine générale, l'augmentation du nombre de postes en médecine générale (4, 5) et de celui de chefs de cliniques en médecine générale (6) justifient des travaux de thèses plus nombreux portant sur les soins de santé primaires.

Cependant, il existe peu d'information concernant la méthode utilisée par les étudiants, alors que cette information serait très utile pour aider les enseignants à choisir quelles méthodes de recherche aborder en cours, d'autant plus que des difficultés rencontrées par les étudiants futurs thésards ont déjà été rapportées par certaines études (7, 8).

INTRODUCTION

Cette thèse a pris substance lors d'une séance de cours de DES de médecine générale animée par le Professeur Jean-Philippe JOSEPH, qui me proposa un travail de recherche concernant les méthodes des thèses. Son postulat était qu'il y avait davantage de thèses de recherche à méthode quantitative, que certains étudiants étaient amenés à réaliser eux-même l'analyse statistique dans le cadre de leur recherche mais ne disposaient pas de la formation nécessaire. Cependant, aucune donnée n'était actuellement disponible ou suffisamment solide pour confirmer ces hypothèses.

Toute thèse étant un travail de recherche scientifique, elle repose sur un design consensuel logique. Partant d'une problématique, des faits connus et inconnus, puis d'une question de recherche et d'hypothèses, les objectifs principaux et secondaires permettent de choisir la méthode la plus adaptée pour répondre à la question de recherche soulevée par la problématique.

Nous avons choisi de réaliser l'inventaire de ce que les étudiants produisaient dans leurs travaux de thèses, de déterminer au sein des productions si certaines méthodes étaient plus souvent employées, quelles étaient les typologies de méthodes les plus fréquentes, avec pour objectif d'orienter et d'optimiser la formation des étudiants en donnant des arguments pour les enseignants.

Afin d'affiner la recherche et d'optimiser le temps nécessaire pour la saisie des données, il paraissait judicieux de créer une liste exhaustive des différentes entités représentant les typologies de méthodes d'études disponibles, mais également les éléments de forme qui apparaissaient dans les différents travaux. Nous avons pour postulat que toutes les thèses à inclure, dans leur rédaction, n'étaient pas égales entre elles quand à leur clarté et leur facilité de lecture (que nous qualifierons de paramètres de forme), indépendamment du sujet abordé, de la méthode ou des résultats (que nous qualifierons de paramètres de fond). Cette analyse suivrait un cheminement systématique des éléments les plus accessibles aux moins accessible, à savoir le titre, puis le résumé, et enfin le texte intégral.

Il était impossible de lire la totalité des textes intégraux (ce qui représentait pourtant un design Gold-Standard pour ce travail) ou de s'appuyer sur un système informatisé programmable pour le faire. De même, il était très compliqué de réaliser une recherche systématique s'appuyant exclusivement sur les mots-clés des différents éléments analysés, et cette performance aurait été moindre qu'une lecture systématique pouvant interpréter le sens des phrases dans leur ensemble, chose impossible avec les mots-clés seuls. De plus, des relectures seraient indispensables, ne serait-ce que pour vérifier les résultats.

Suivant les conseils du directeur de thèse, notre analyse devrait dans un premier temps s'orienter vers la recherche quantitative, qui était selon lui la plus représentée et la plus complexe à enseigner.

Il fallait déterminer le degré de complexité de l'analyse au sein de chaque thèse, chose qui n'était pas systématiquement explicite dans le titre ni dans le résumé, afin de constituer un échantillon représentatif, sans émettre le moindre jugement ni remettre en question la validité du moindre résultat ni de la moindre conclusion.

Une autre question semblait essentielle : supposant que certaines thèses faisaient appel de leur méthode à des analyses statistiques sophistiquées, dépassant le cadre de l'enseignement du DES de médecine générale, les étudiants étaient susceptibles d'avoir recours à une aide méthodologique, pour les assister dans la réalisation de l'analyse statistique de leur thèse. Or cette donnée n'apparaissait dans aucune publication à ce jour, la bibliographie en ce sens était introuvable. Cette variable serait originale, et utile comparée à d'autres paramètres en rapport avec la réalisation des calculs statistiques, nous permettant de déterminer si cette démarche était fréquente, et si elle avait une relation avec le déroulement de l'analyse dans leurs travaux.

QUESTION DE RECHERCHE

Quelles sont les méthodes de recherche quantitative utilisées au cours des thèses de médecine générale soutenues au cours des dix dernières années en Aquitaine ?

HYPOTHÈSES

- Il y a davantage de thèses de recherche quantitative que de thèses de recherche qualitative.
- Il existe une grande variété de méthodes utilisées dans les thèses portant sur les soins de santé primaires, mais peu utilisent des outils statistiques complexes car ils ne sont pas ou peu abordés au cours du DES de médecine générale.
- Les étudiants ont souvent recours à une aide méthodologique pour la réalisation de l'analyse statistique, et cette aide serait extérieure au DMG.
- La réalisation du travail de l'étudiant varie en fonction du directeur de thèse, selon qu'il soit médecin généraliste ou spécialiste, qu'il appartienne ou non au CHU, ou qu'il soit ou ne soit pas enseignant au DMG de Bordeaux.

OBJECTIF PRINCIPAL

Déterminer la typologie et la répartition des différentes méthodes de recherche quantitative utilisées dans les thèses de médecine générale soutenues à l'université de Bordeaux au cours des dix dernières années.

OBJECTIFS SECONDAIRES

- Déterminer la proportion de thèses ayant une méthode de recherche quantitative et évolution sur les dix dernières années.
- Déterminer les tests statistiques utilisés.
- Déterminer les logiciels utilisés pour l'analyse statistique.
- Détailler les méthodes statistiques : détermination du nombre de sujet nécessaire, calcul du risque α , et calcul du risque β .
- Déterminer le recours à une aide méthodologique externe au DMG et recherche l'existence d'un lien entre le recours à une aide méthodologique hors du DMG et la méthode statistique.
- Déterminer la catégorie du directeur de thèse, et recherche l'existence d'un lien entre la catégorie du directeur de thèse et la méthode statistique.

MÉTHODES

TYPE D'ÉTUDE

Etude descriptive transversale.

POPULATION DE L'ÉTUDE

Ensemble des thèses de médecine générale soutenues à l'université de Bordeaux, ayant pour champ les soins de santé primaires, entre juin 2006 et juin 2016.

TECHNIQUE DE RECUEIL DES DONNÉES

Nous avons relevé les données pour chaque thèse et enregistré dans un tableur Numbers© version 3.6.2 (tableur sous Mac OS X®, équivalent Excel).

Plusieurs copies du document Numbers furent enregistrées de manière synchronisée à titre de prévention d'une éventuelle panne ou perte des données, sur supports physiques et dématérialisés.

MODE DE RECUEIL INITIAL DES DONNÉES EXHAUSTIVES

Inclusion systématique à partir des données d'APOGÉE¹. Exclusions des thèses de spécialités et des thèses hors soins de santé primaires. La période ciblée couvrait la totalité des années disponibles.

Les premières données furent récupérées au format PDF sur le site du département de médecine générale de l'université de Bordeaux², après accord des responsables (Professeur JP. Joseph, Madame C. Bonnin).

Ces données comportaient sous forme de colonnes, de gauche à droite, les variables suivantes pour chaque étudiant :

N° étudiant, Nom, Prénom	Date de lère inscription	Titre de la thèse, N° thèse, année	Directeur Co-directeur Second co- directeur	Jury Qualité	Rôle (des membres du jury)	Date de soutenance, résultat, mention
--------------------------------	--------------------------------	---------------------------------------	--	-----------------	----------------------------------	--

Nous avons relevé le nom et le prénom de chaque étudiant, sa mention, le sujet de la thèse.

Par la suite, le sexe de l'étudiant a été déduit de son prénom, avec recours pour les prénoms inhabituels ou étrangers à une recherche de son identité sur société.com (pour les libéraux, ce site présentant l'avantage de désigner les personnes par Madame ou Monsieur). En cas de litige ou si l'information était indisponible, ou pour des prénoms androgynes (comme Dominique par exemple), nous avons contacté directement la personne ou son secrétariat, à partir de son numéro de téléphone trouvé sur les pages jaunes, pour vérifier qu'elle était bien le thésard ou la thésarde supposé(e), et le cas échéant, récupérer l'information.

Au terme de cette étape, chaque thèse comportait donc à titre de variables :

- Un titre.
- L'identité du thésard.
- Son sexe.
- La date de soutenance.
- La mention obtenue.

¹ APOGÉE est une application informatique de l'Agence de Mutualisation des Universités et des Établissements, qui rassemble au format PDF, par année universitaire, l'ensemble des thèses soutenues. Plus d'informations ici : <http://www.cri.u-bordeaux2.fr/index.php/apogee/53-presentation/61-gestion-des-enseignements-et-des-etudiants>

² <https://sante.u-bordeaux.fr/Vos-etudes/DES-Medecine-generale/These-d-exercice-de-Medecine-Generale>

PREMIER TRI DES DONNÉES : DÉTERMINER LE RAPPORT AUX SOINS DE SANTÉ PRIMAIRES À PARTIR DU TITRE

Ce tri fut réalisé après avoir récupéré l'ensemble des titres des thèses de médecine générale pour la période prévue. Chaque titre fut lu puis coté :

1. Inclus.
2. Douteux.
3. Non-inclus.

Cette lecture reposait sur une grille issue des critères WONCA définissant les soins de santé primaires en médecine générale (9, 10) :

- Le thème est-il dans le **champ de la médecine générale** ou de son **enseignement** ?
- Les conclusions attendues sont-elles d'**intérêt pour la pratique** du médecin généraliste?
- Le problème étudié est-il **suffisamment fréquent** pour être d'intérêt pour la pratique du médecin généraliste ?
- Le mode de recrutement permet-il d'**extrapoler les résultats** à une **population proche** de celle rencontrée en médecine générale ?
- L'étude entre dans le cadre d'une **recherche initiée par le DMG** ?
- L'étude est réalisée en coopération avec une autre spécialité, ou une autre profession, mais intéresse les soins de santé primaires (soins de santé primaires) ?

Cette grille était réduite à un nombre restreint de questions à réponse dichotomiques, en cas de doute une réponse positive était retenue, en fonction de l'avis de l'auteur.

Pour chaque titre coté « 2. douteux », une relecture eut lieu avec le directeur de thèse dans le but de trancher.

Au moins un critère présent était nécessaire pour qu'une thèse soit incluse.

Au terme de cette étape, chaque thèse comportait, en complément des étapes précédentes, une variable déterminant son inclusion ou sa non-inclusion en tant que soins de santé primaires.

L'analyse des thèses de soins non-primaires prenait fin à cette étape.

DEUXIÈME TRI DES DONNÉES : DÉTERMINER L'APPORT DU **TITRE** COMME DÉTERMINANT LE TYPE DE MÉTHODE DE RECHERCHE EMPLOYÉE

Première approche basée sur les mots-clés

Cette étape intéressait les thèses triées « soins de santé primaires », les autres thèses ayant été exclues de l'étude.

Pour chaque titre coté « inclus soins de santé primaires », une recherche de mots-clés a été réalisée de manière systématique. Cette étape avait pour but de démontrer l'effet d'une recherche de mots-clés sur des titres dont l'écriture était libre.

Au préalable, tous les textes des différents titres avaient été dépouillés de leurs accents, pour favoriser leur mise en évidence, car selon l'année le titre de la thèse pouvait être en lettres capitales et dépourvu d'accents.

Deux groupes de mots-clés (ou de troncatures) ont été utilisés, basés sur les critères proposés par Paul FRAPPÉ (11) :

- Un groupe de mots-clés évocateurs de recherche quantitative : epidemiologi, descripti, ecologi, transversal, analyti, prospecti, retrospecti, essai, entretien, temoin, cohorte, quantitati, etude.
- Un groupe de mots-clés évocateurs de recherche qualitative : litterature, entretien, focus, observation, qualitati.

Ces mots étaient volontairement tronqués pour couvrir le plus de synonymes possibles, les accents étaient absents. L'absence d'accents permettait de mieux identifier les mots clés.

Leur mise en évidence fut réalisée grâce à la fonction **surlignage conditionnel** du tableur, selon deux colonnes où les titres étaient recopiés. De ce surlignage fut tirée une cotation déduite comme suit :

1. Quantitatif	Au moins un mot-clé de recherche quantitative	Aucun mot-clé de recherche qualitative
2. Douteux	Au moins un mot-clé de recherche quantitative	Au moins un mot-clé de recherche qualitative
3. Qualitatif	Aucun mot-clé de recherche quantitative	Au moins un mot-clé de recherche qualitative
4. Indéterminable	Aucun mot-clé de recherche quantitative	Aucun mot-clé de recherche qualitative

Deuxième approche par relecture systématique

Dans l'hypothèse d'une faible performance de la recherche de mots-clés, une seconde lecture des titres correspondant aux thèses triées « soins de santé primaires » fut réalisée, binc qu'influencée par la cotation basée sur les mots-clés.

Cette nouvelle cotation suivait une règle inspirée de l'étape précédente, tout en introduisant la possibilité d'une méthode mixte : quantitative et qualitative. La nouvelle cotation était donc :

1. Quantitatif
2. Douteux
3. Qualitatif
4. Indéterminable
5. Quantitative et Qualitative

De plus, il fut noté grâce à un algorithme de déduction (annexe 1.1) si la relecture était cohérente avec la première cotation déduite par mots-clés, « inchangé », et si la nouvelle cotation était différente, « initialement X, changé pour Y » (où X était la cotation déduite par mots-clés et Y la cotation déduite par relecture systématique influencée par la cotation par mots-clés).

Synthèse du deuxième tri

Au final, cette étape permettrait de trier au sein des thèses de soins de santé primaires celles dont le titre était explicitement de méthode qualitative, dans le but de les exclure.

En cas de cotation discordante, était retenue pour la suite la cotation de relecture systématique, qui tenait compte du titre dans son ensemble, et avait plus de sens et de poids qu'une recherche quantitative de mots-clés.

Les thèses étaient donc réparties en 3 catégories :

- Le titre était contributif et évocateur de méthode quantitative, son résumé serait lu pour déterminer la typologie quantitative et les variables des critères secondaires.
- Le titre était peu contributif, la méthode était douteuse ou indéterminable, son résumé serait lu pour déterminer le type de méthode
- Le titre était contributif et évocateur de méthode qualitative, la thèse était exclue.

TROISIÈME TRI DES DONNÉES : DÉTERMINATION DE LA MÉTHODE BASÉE SUR LE **RÉSUMÉ**, CONCLUSION DU TYPE DE MÉTHODE, RELEVÉ DES VARIABLES DES CRITÈRES SECONDAIRES

Dans la but de travailler sur les résumés, nous avons réalisé une recherche systématique sur la catalogue SUDOC³ des différentes thèses non-exclues. Cette recherche utilisait comme mot-clé de manière récursive le(s) nom(s) et le(s) prénom(s) du thésard, puis le titre de la thèse, puis le nom du directeur de thèse jusqu'à obtenir la fiche de la-dîte thèse.

Le lien pérenne de cette fiche était enregistré pour une utilisation ultérieure (principalement relecture de la saisie du tableur), un algorithme (annexe 1.2) était utilisé pour vérifier que le lien avait été correctement enregistré et éviter les erreurs ultérieures.

Les données pouvaient être, selon l'année et l'auteur, à partir des données de SUDOC :

- Intégralement disponibles en ligne, résumé et texte intégral.
- Seul le résumé pouvait être disponible en ligne, le texte intégral serait consulté à la bibliothèque universitaire.
- Exclusivement disponible à la bibliothèque universitaire, aucune information n'était disponible en ligne.

Certaines thèses pourraient ne pas être trouvées sur SUDOC, une recherche élargie serait réalisée sur internet de manière plus large, incluant DUMAS⁴, utilisant la mêmes mots-clés (nom(s), prénom(s), titre de la thèse, directeur). Le cas échant, le lien permanent vers la thèse était enregistré.

En cas d'absence totale d'information disponible en ligne, la thèse était considérée comme étant exclusivement disponible en bibliothèque universitaire.

³ <http://www.sudoc.abes.fr/>

⁴ <https://dumas.ccsd.cnrs.fr/>

Détermination de la méthode basée sur la lecture du résumé

Chaque thèse dont le titre était peu contributif, dont la méthode était douteuse ou indéterminable, était coté « lecture du résumé nécessaire car recherche quantitative ».

Chaque thèse dont le titre était contributif, dont la méthode était quantitative, était coté « lecture du résumé nécessaire pour détermination des critères secondaires ».

Une lecture du résumé était réalisée, et sa méthode était déterminée. Six classes étaient définies dont classes d'erreur :

0. Obstacle technique
1. Quantitatif
2. Non-nécessaire
3. Qualitatif
4. Mixte
5. Autres méthodes

La classe « 0. Obstacle technique » regroupait les thèses dont les données étaient introuvables en bibliothèque universitaire, ou incomplètes.

La classe « 2. Non-nécessaire » était réservée aux thèses dont la méthode était déterminée quantitative par la titre seul, et avait pour seul but de confirmer la détermination antérieure.

La classe « 4. Mixte » remplaçait la classe Quantitative et Qualitative, le sens était le même mais la variable avait été raccourcie pour simplifier son intégration dans les algorithmes tout en conservant le même sens.

La classe « 5. Autres méthodes » était prévue pour des thèses qui ne concernaient pas la recherche à proprement parler, comme la rédaction de référentiels.

Recueil des variables du critère de jugement principal et des critères secondaires depuis le résumé

Les thèses étaient classées selon leur méthode.

Ont été exclues les thèses de méthode « 3. Qualitatif » et « 5. Autres méthodes », ainsi que les thèses « 0. Obstacle technique ».

Ont été conservées les thèses de méthode « 1. Quantitatif » et « 4. Mixte » pour recherche des différents critères secondaires.

Pour chacune de ces thèses, nous avons relevé les variables suivantes, toutes collectées à partir des résumés et des informations complémentaires de la fiche disponible sur SUDOC (exemples en annexe 2) :

Concernant la forme

- Quelle était la langue dans laquelle le résumé était disponible ? Français/Bilingue anglais/français.
- Le résumé était-il rédigé selon une structure IMRaD (Introduction, Matériel et Méthodes, Résultats et Discussion), et au moins 3 des mots-clés Introduction/Justification/Contexte ; Matériel et/ou Méthode(s) ; Résultats ; Discussion apparaissaient-ils ? Oui, Non.

Nous avons appliqué une tolérance concernant l'absence du premier mot-clé, considérant qu'il pouvait être tacite au début du résumé.

- Si le résumé était rédigé selon une structure IMRaD, la structure était-elle explicite, les mots étaient-ils démarqués des autres (italique, gras, sous-ligné, majuscules, lettres capitales) ? Oui, Non.
- Quel était le nombre de pages de la thèse ? Valeur numérique, une cotation NC était appliquée quand l'information était indisponible.

Concernant la disponibilité

- Le texte intégral était-il disponible en ligne ? Oui, Non.

Si le résumé était indisponible en ligne, sa cote en bibliothèque universitaire était déduite par un algorithme prédictif (annexe 1.3) du lien sur le site Babord+⁵.

- Le résumé était-il disponible en ligne ? Oui, Non.

Si le texte intégral était indisponible en ligne, sa cote en bibliothèque universitaire était déduite par un algorithme prédictif (annexe 1.3) du lien sur le site Babord+.

Concernant le directeur de thèse

- Quel était le directeur de thèse ? Étaient relevés son nom et son prénom, ainsi que son domaine d'activité :

1 : Médecin spécialiste extérieur au CHU (dont urgences, médecine du sport, gynécologie et pédiatrie).

2 : Médecin généraliste extérieur au DMG (dont médecine à exercice particulier, gériatrie, médecin militaire généraliste et médecins rattachés à l'enseignement des DOM-TOM).

3 : Médecin généraliste enseignant rattaché au DMG de Bordeaux.

4 : Médecin du CHU.

5 : Non-médecin.

- Le cas échéant, si un co-directeur de thèse était déclaré, étaient relevés les mêmes paramètres que ceux du directeur de thèse.

⁵ <https://babordplus.u-bordeaux.fr/>

Concernant les données scientifiques

- La taille de l'échantillon apparaissait-elle dans le résumé ? Oui/Non.
- La taille de l'échantillon était relevée. Des variables spécifiques étaient déterminées ; -1 était réservé à une donnée absente du résumé ; -2 était réservé à une donnée absente pour cause d'embargo.
- L'échantillon avait-il été tiré au sort ? Oui/Non.
- Combien de groupes constituaient l'échantillon ? Un groupe, Deux groupes, Plus de deux groupes, Non-relevant.

La dernière classe était réservée aux études où le nombre de groupe n'était pas relevant, en l'occurrence les méta-analyses.

- Type de variables utilisées supposées : basé sur les résultats affichés, nous avons relevé une variable prédisant si les variables utilisées étaient :

1. Continues.
2. Booléennes/Discretisées.
3. Mixtes, continues et/ou discretisées/booléennes. Cotation utilisée en cas de doute sur une variable unique apparaissant dans le résumé.
4. Inconnu, quand aucune variable n'apparaissait.

- Des paramètres statistiques apparaissaient-ils ? ***p values***, intervalles de confiance [IC], moyenne/médiane/percentiles/quartiles ? Oui/Non.

- La présence d'une analyse statistique pouvait-elle être déduite du résumé ?

1. Vraisemblable : ont été considérés comme déterminant « vraisemblable » la présence dans le résumé d'au moins un des mots-clés suivants : « significatif/significative » ; « statistique/statistiquement ».
2. Certaine : ont été considérées comme déterminant « certaine » la présence dans le résumé d'au moins un des mots-clés suivants : tout paramètre statistique ; nom d'un test statistique.
3. Absente.

Concernant la typologie de méthode

Cette étape reposait sur une évaluation de la méthode basée sur les données du résumé, analysées et triées via une classification objective proposée par Paul FRAPPÉ (6). Cette classification complète est disponible en annexe (annexe 3). Chaque résumé était lu, sa question de recherche identifiée (en fin d'introduction), et son sous-type d'étude déduit par le résumé.

Ci-dessous sont détaillés les critères d'orientation en fonction de la question de recherche, inspirés de P. Frappé (7). *Les lignes rouges correspondent aux typologies qualitatives exclues.*

Types d'études selon la classification proposée par Paul FRAPPÉ		Question de recherche										
		1.1 DÉCRIRE DES VARIABLES	1.2 DÉCRIRE DES FAITS	2.1 ANALYSER UNE VARIABLE : CHERCHER UNE ASSOCIATION	2.2 ANALYSER UNE VARIABLE : CHERCHER UNE CAUSALITÉ	3.1 EVALUER UN TRAITEMENT, UNE PRATIQUE, UN TEST	3.2 EVALUER UN PRONOSTIC	3.3 EVALUER UN IMPACT ECONOMIQUE	3.4 EVALUER DES PREFERENCES ECONOMIQUES	4 ACCORDER : ETABLIR UN CONSENSUS	5 EXPLORER LE SENS, LE VÉCU, LES PERCEPTIONS, LE COMPORTEMENT	6 ETUDE DE FAISABILITE
1 Quantitatif	1.1 : Epidémiologique, étude de cas / Série de cas											
	1.2 : Epidémiologique, étude écologique											
	1.3 : Epidémiologique, étude transversale											
	1.4 : Epidémiologique, Cas/témoins											
	1.5 : Epidémiologique, Cohorte											
	1.6 : Interventionnel, Essai clinique											
2 Evaluation test diagnostic	2.0 : Evaluation test diagnostic											
3 Medico-économique	3.1 : Médico-économique, minimisation des coûts											
	3.2 : Médico-économique, coût efficacité											
	3.3 : Médico-économique, coût-utilité											
	3.4 : Médico-économique, coût-bénéfice											
4 Semi-Quantitatif	4.1 : Semi-Quantitatif, processus Delphi											
	4.2 : Semi-Quantitatif, Groupe nominal											
5 Evaluation de la qualité des soins	5.1 : Evaluation de la Qualité des Soins, enregistrement systématique											
	5.2 : Evaluation de la Qualité des Soins, enquête de pratique											
	5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques											
	5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale											
	5.5 : Evaluation de la Qualité des Soins, audit clinique											
	5.6 : Evaluation de la Qualité des Soins, essai en population											
	5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs											
	5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après											
6 Qualitative	6.1 : Qualitative, entretien individuel structuré											
	6.2 : Qualitative, observation directe											
	6.3 : Qualitative, observation indirecte											
	6.4 : Qualitative, entretien individuel semi-structuré											
	6.4 : Qualitative, entretien individuel non structuré											
	6.6 : Qualitative, focus group											
7 Recherche secondaire	7.1 : Recherche secondaire, revue systématique											
	7.2 : Recherche secondaire, méta-analyse quantitative											
	7.3 : Recherche secondaire, méta-analyse qualitative											

Illustration M1. Extrait de Initiation à la recherche de Paul FRAPPÉ, page 27.

CRITÈRES DE CHOIX DU TYPE D'ÉTUDE.

Tableau de contingence des associations question de recherche / étude(s) de faisabilité et sous-méthodes de recherche possibles.

Les éléments suivants ont servi à déduire la typologie de sous-méthode quantitative employée par l'étudiant.

- De la méthode déclarée par l'étudiant. *Il était cependant possible que la méthode déclarée par l'étudiant soit sommaire, se limite à « étude » ou « étude prospective/retrospective ».*
- De la durée du relevé des variables (ponctuelle, répétée, suivie dans le temps).
- De la nature de la population étudiée (patients/malades, médecins).
- Du type de variable étudiée (prévalence, incidence, résultat d'un test, coût, pratique) le cas échéant.
- De l'existence d'une comparaison au sein de l'étude.
- Du caractère prospectif ou rétrospectif de l'étude.
- Du mode de recueil des données, tout particulièrement les questionnaires.
- Du(des) résultat(s).

La méthode principale fut relevée. Si une méthode secondaire était présente elle était également relevée. Une de ces méthodes était nécessairement qualitative pour les thèses de méthode cotée « 4. Mixte ».

Ont également été relevées :

- La méthode déclarée par l'étudiant dans sa thèse a été prise en compte. Il y avait plusieurs cas de figure :
 1. Consensuel : la méthode de l'auteur était déclarée et informative, et cohérente avec la méthode déduite selon les critères de Paul FRAPPÉ.
 2. Non-Consensuel : la méthode de l'auteur était déclarée, mais non-cohérente avec la méthode déduite selon les critères de Paul FRAPPÉ. Étaient inclus dans ce groupe les méthodes de déclaration incomplète ou peu informatives, le sous-type d'étude était attendu dans le résumé.
 3. Non-précisé par l'auteur : la méthode n'était pas déclarée par l'auteur dans le résumé.

- La source des données de l'étude fut relevée, comme suit :

1. Bases de données médicales hospitalières.
2. Bases de données médicales non-hospitalières.
3. Bases de données de santé publique (INSERM, ISPED, CPAM, autres caisses primaires de sécurité sociale).
4. Dossier médical du cabinet du médecin généraliste/du spécialiste.
5. Questionnaire conçu pour l'étude/issue d'une autre étude.
6. Multiples, ne comportant pas de questionnaires.
7. Multiples, comportant un questionnaire.
8. Données issues d'une autre étude.
9. Données issues de la faculté, contenu d'autres thèses.

- Enfin, fut relevée la date de fin d'un éventuel embargo.

Synthèse du troisième tri

Étaient récupérés :

- Présence du résumé en ligne (sur SUDOC, sur une autre plate-forme).
- Lien vers le résumé si disponible.
- Langue du résumé.
- Structure du résumé, présence d'une structure IMRaD, caractère explicite de la structure IMRaD.
- Présence du texte intégral en ligne.
- Si texte intégral et/ou résumé indisponible en ligne, relevé de la coté babord+.
- Conclusion du critère de jugement principal (quantitatif/mixte/qualitatif/aucun des 2).
- Identification directeur de thèse et fonction.
- Identification co-directeur et fonction si relevant.
- Méthodes de recherche quantitative ou mixte selon critères de Paul FRAPPÉ.
- Détermination du mode de recueil des données.
- Cohérence déclaration de l'auteur/relecture du résumé pour déterminer la méthode.
- Déclaration de la taille de l'échantillon dans le résumé, si relevant la taille était relevée.
- Echantillon tiré au sort ou non.
- Nombre de groupes de l'échantillon.
- Type de variables supposées déclarées apparaissant dans le résumé.
- Paramètres significatifs (p values, [IC], moyennes, médianes, quartiles, percentiles) apparaissant dans le résumé.
- Analyse statistique vraisemblable, possible, exclue.
- Date de fin d'embargo, si relevant.

QUATRIÈME TRI DES DONNÉES : EXPLORATION DE LA MÉTHODE DES TEXTES INTÉGRAUX DES THÈSES

Nous avons sélectionné des thèses dont le résumé comportait des éléments évocateurs d'analyse statistique, comme évoqué en fin du paragraphe « concernant les données scientifiques » du troisième tri.

Toute thèse dont l'analyse statistique était supposée « vraisemblable » ou « certaine » était lue dans son texte intégral, et la partie méthode était analysée pour relever de données pour les critères de jugement secondaire.

Les textes intégraux étaient :

- Disponibles en ligne, le lien était trouvé sur la page de SUDOC (annexe 2.2) et relevé. Ce lien provenait des archives DUMAS. Le numéro de page était relevé pour faciliter la relecture.
- Indisponible en ligne, la cote Babord+ était relevée, et le texte intégral était lu en bibliothèque universitaire.
- Sous embargo, si cela n'avait pas été noté lors de l'étape précédente, la date de fin était relevée, et l'analyse prenait fin.

Pour chaque thèse comportant une analyse statistique « vraisemblable » ou « certaine », ont été relevés :

- Le calcul du NSN (Nombre de Sujets Nécessaires) était déclaré par l'auteur dans la méthode ?
Oui
Non
- La détermination du risque α apparaissait dans la méthode ?
Oui, $\alpha=0,05$ ou $\alpha=5\%$
Oui, $\alpha<0,05$ ou $\alpha<5\%$
Oui, $\alpha>0,05$ ou $\alpha>5\%$
Non
- La détermination du risque β apparaissait dans la méthode ?
Oui, $\beta=0,20$ ou $\beta=20\%$
Oui, $\beta<0,20$ ou $\beta>20\%$
Oui, $\beta>0,20$ ou $\beta>20\%$
Non

- Un ou des logiciel(s) pour le calcul statistique étaient déclaré(s) par l'auteur dans la méthode ?
Non-défini
Si défini, jusqu'à 4 logiciels étaient relevés
- Le recours à un tiers (individu, équipe) hors du DMG pour l'analyse statistique était déclaré par l'auteur ?
Oui
Non
- L'auteur déclarait des tests de variables continues ? Si un résultat d'allure continue apparaissait dans le résumé (relevé au tri 3), au moins un test de variable continue était attendu.
Non-défini
Si défini, jusqu'à 4 tests étaient relevés
- L'auteur déclarait des tests de variables discrètes/booléennes ? Si un résultat d'allure discret/booléen apparaissait dans le résumé (relevé au tri 3), au moins un test de variable discrète/booléenne était attendu.
Non-défini
Si défini, jusqu'à 4 tests étaient relevés
- L'auteur déclarait la réalisation d'une analyse univariée, bivariée, multivariée ou de régression logistique ?
Si plusieurs analyse étaient déclarées, seule était relevée l'analyse la plus complexe : régression logistique > multivariée > bivariée > univariée.

Synthèse du quatrième tri

Étaient récupérés :

- Lien du texte intégral ou cote Babord du texte papier.
- Calcul du nombre de sujets nécessaires.
- Détermination du risque α .
- Détermination du risque β .
- Logiciel(s) de calcul statistique utilisé(s).
- Déclaration du recours à un tiers hors du DMG pour le calcul statistique.
- Déclaration de test(s) de variables continues.
- Déclaration de test(s) de variables discrètes/booléennes.
- Déclaration d'une analyse statistique complexe, uni/bi/multivariée.

La totalité des étapes de la méthode de saisie des données fut synthétisée sous la forme de diagrammes de flux théoriques, disponibles en annexe (annexe 4).

MODE D'ANALYSE DES DONNÉES

Les effectifs étaient fournis sous forme de nombres entiers.

Les pourcentages étaient fournis sous forme de nombres décimaux avec un chiffre significatif après la virgule.

Les valeurs significatives des tests statiques étaient fournies selon la norme scientifique avec deux chiffres significatifs après la virgule.

Le calcul du nombre de sujet nécessaires n'était pas applicable à cette étude, car la liste de thèses incluse serait exhaustive.

La valeur du risque α était définie à 0,05 soit 5%.

En l'absence de calcul du nombre de sujet nécessaire, la détermination de la puissance et du risque β n'était pas nécessaire.

Concernant l'objectif principal

Les variables de typologie des méthodes furent reportées dans un tableau, trié par effectif décroissant, associé à la fréquence relative. Pour plus de pertinence, cette typologie déduite selon la méthode de Paul FRAPPÉ fut associée à un triple indice de confiance, reflétant le caractère consensuel entre la sous-méthode déduite selon les critères consensuels de Paul FRAPPÉ et la déclaration de typologie du thésard, ou l'absence de déclaration par le thésard.

Concernant les objectifs secondaires

- Les thèses de méthode quantitative furent triées selon l'année universitaire. Pour chaque année fut relevée l'effectif et la fréquence relative. Une test statistique serait réalisé pour comparer les effectifs, le choix de ce test serait un test t de Student ou un test de Mann-Whitney-Wilcoxon si l'effectif était insuffisant.
- Nous avons relevé les variables descriptives de la méthode de calcul statistique utilisée dans chaque thèse dont le texte intégral aurait été lu. Furent compilés sous forme de tableau :
 - Les tests utilisés, sous forme de catalogue construit au fur et à mesure de la saisie, pour variables continues puis pour variables discretisées/booléennes.
 - La détermination du nombre de sujets nécessaires.
 - La détermination du risque α et du risque β .
 - Le recours à une aide méthodologique tierce externe au DMG.
 - La catégorie du directeur de thèse.

Ces variables furent comparées, d'une part avec la déclaration du recours à un tiers pour l'analyse statistique, puis en fonction de la catégorie du directeur de thèse, via un test de Khi-2 de Pearson ou à un test exact de Fisher si l'effectif était insuffisant.

Concernant les autres variables relevées

Les autres variables étaient relevées à titre de composition d'un fichier de données exhaustives, qui pourrait servir de base pour des travaux ultérieurs. L'auteur serait éventuellement amené à analyser ces données, et elles seraient alors jointes en annexe à la thèse. Il était prévu que ce fichier soit diffusé.

Concernant les logiciels utilisés

Les logiciels pour le calcul statistique furent Numbers© qui servait également à la saisie des données, puis biostatgv©⁶ et R©⁷.

Les logiciels pour la rédaction de la thèse furent Pages® version 5.6.2 (traitement de texte sous Mac OS X, équivalent Words).

La retouche des images incluses fut réalisée sous Photoshop® CS5 Extended version 12.0.

La création des diagrammes de flux fut réalisée sous Draw.io©⁸, logiciel libre de création de diagrammes.

ASPECTS RÉGLEMENTAIRES ET ÉTHIQUES

Ce travail n'a pas nécessité de dépôt du projet auprès du Comité de Protection des Personnes (CPP).

Ce travail n'a pas nécessité de dépôt du projet auprès du Comité Consultatif sur le Traitement de l'Information en Matière de Recherche (CCTIRS).

Ce travail n'a pas nécessité de dépôt du projet auprès de la Commission Nationale Informatique et Liberté (CNIL).

⁶ <https://marne.u707.jussieu.fr/biostatgv/>

⁷ <https://cran.r-project.org/>

⁸ <https://www.draw.io/>

RÉSULTATS

POPULATION TOTALE OBTENUE À PARTIR DES DONNÉES D'APOGÉE

Le nombre de thèses soutenues pendant 10 ans du 14 septembre 2005 au 28 juin 2016 récupérées dans APOGÉE était de 1179.

DIAGRAMME DE FLUX DE L'ÉTUDE

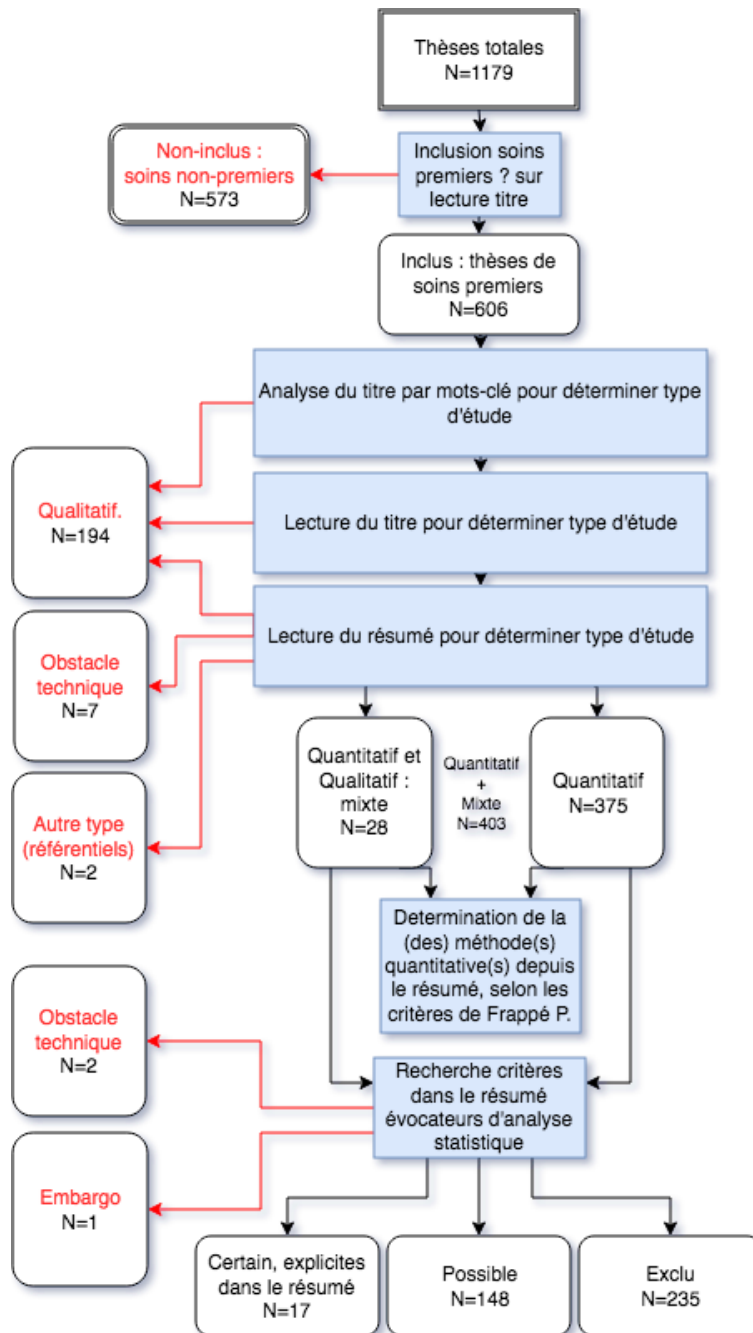


Schéma R1. Flow-Chart de l'étude

POPULATION INCLUSE « SOINS DE SANTÉ PRIMAIRES » À PARTIR DE LA POPULATION TOTALE AVEC RÉPARTITION PAR ANNÉE

Le tableau suivant comporte, pour chaque période, le nombre de thèses présentes dans le document d'APOGÉE relatif à l'année, et le nombre de thèses incluses.

Tableau R1. Répartition des inclusions par promotion.

Promotion (document d'origine)	Thèses incluses « soins de santé primaires »	Total thèses
2006-2007	41 (41,4 %)	99
2007-2008	46 (46,0 %)	100
2008-2009	37 (41,6 %)	89
2009-2010	46 (52,9 %)	87
2010-2011	49 (54,4 %)	90
2011-2012	60 (43,2 %)	139
2012-2013	81 (53,3 %)	152
2013-2014	97 (56,7 %)	171
2014-2015	94 (58,4 %)	161
2015-2016	55 (60,4 %)	91
Total de 2006 à 2016.	606 (51,4 %)	1179

On note que la proportion de thèses de soins de santé primaires suit une pente globalement ascendante sur les 10 dernières années.

RECOURS AU DIRECTEUR DE THÈSE POUR DÉTERMINER L'INCLUSION DES THÈSES « DOUTEUSES »

18 thèses étaient classées « douteuse ». Après discussion avec le directeur de thèses, elles ont été reclassées comme :

- « incluses » : 11 thèses.
- « non-incluses » : 7 thèses.

PREMIÈRE ÉTAPE : TRI DES TITRES DES THÈSES PAR MOTS-CLÉS

Tableau R2. Identification de la méthodologie par mots-clés.

Tri des titres pour méthode employée déduite par mots clés.	Effectif
Total quantitatif :	186 (30,7 %)
Total qualitatif :	12 (2,0 %)
Total douteux :	47 (7,8 %)
Total indéterminable :	361 (59,6 %)
Première déduction déterminante (quantitatif ou qualitatif) :	198 (32,7 %)
Première déduction non-déterminante ou confuse :	408 (67,3 %)
Total inclus :	606 (100 %)

DEUXIÈME ÉTAPE : TRI DES THÈSES PAR LECTURE SYSTÉMATIQUE DU TITRE COMPARAISON AUX RÉSULTATS DU TRI PAR MOTS-CLÉS

Tableau R3. Identification de la méthodologie par lecture systématique du titre, amélioration des effectifs par rapport au tri par mots-clés.

Tri des titres pour méthode employée par relecture systématique	Effectif	Modification de l'effectif par rapport au tri par mots-clés
Total quantitatif :	271 (44,7 %)	+85
Total qualitatif :	61 (10,1 %)	+49
Total douteux :	255 (42,1 %)	+208
Total quantitatif et qualitatif :	1 (0,2 %)	+1
Total indéterminable :	18 (3,0 %)	-343
Total inclus :	606 (100 %)	

TROISIÈME ÉTAPE : TRI DES THÈSES PAR LECTURE DU RÉSUMÉ AVEC CONCLUSION SUR LA MÉTHODE DE LA THÈSE

Tableau R4. Identification de la méthode par lecture systématique du résumé, amélioration des effectifs par rapport au tri par lecture systématique du titre.

Tri des résumés pour méthode employée par relecture systématique	Effectif	Modification de l'effectif par rapport à la lecture seule du titre
Total quantitatif :	375 (61,9 %)	+105
Total qualitatif :	194 (32,0 %)	+133
Total mixte - quantitatif et qualitatif :	28 (4,6 %)	+27
Total autre méthode :	2 (0,3 %)	+2
Total obstacle technique :	7 (1,2 %)	+6
Total douteux :	0 (0 %)	-255
Total indéterminable :	0 (0 %)	-18
Total inclus :	606 (100 %)	

COMPARAISON DE LA PERFORMANCE DE CHAQUE ÉTAPE DE TRI

Tableau R5. Comparaison des tris par mots-clés du titre, lecture du titre, et lecture du résumé.

Conclusions du tri	Tri par mots-clés du titre	Tri par lecture du titre	Tri par lecture du résumé
Total quantitatif :	186 (30,7 %)	271 (44,7 %)	375 (61,9 %)
Total qualitatif :	12 (2,0 %)	61 (10,1 %)	194 (32,0 %)
Total mixte - quantitatif et qualitatif :	0 (0 %)	1 (0,2 %)	28 (4,6 %)
Total autre méthode :	0 (0 %)	0 (0 %)	2 (0,3 %)
Total obstacle technique :	0 (0 %)	0 (0 %)	7 (1,2 %)
Total douteux :	47 (7,8 %)	255 (42,1 %)	0 (0 %)
total indéterminable :	361 (59,6 %)	18 (3,0 %)	0 (0 %)
Total inclus :	606 (100 %)	606 (100 %)	606 (100 %)

Par conséquent, la lecture du résumé est plus performante que la lecture du titre dans son ensemble, toutes deux sont plus performantes que l'identification de mots-clés dans le titre.

DÉTERMINATION DU NOMBRE ANNUEL DES THÈSES CLASSÉES « SOINS DE SANTÉ PRIMAIRES » DE MÉTHODE QUANTITATIVE OU MIXTE

Tableau R6. Comparaison des effectifs annuels des thèses de méthodologie quantitative et mixte, rapport aux thèses incluses « soins premiers ».

Promotion (document d'origine)	Effectif thèses quantitatives	Effectif thèses mixte	Effectif thèses soins de santé primaires
2006-2007	22 (53,7 %)	0 (0 %)	41
2007-2008	28 (60,9 %)	2 (4,4 %)	46
2008-2009	26 (70,3 %)	3 (8,1 %)	37
2009-2010	30 (65,2 %)	5 (10,9 %)	46
2010-2011	27 (55,1 %)	1 (2,0 %)	49
2011-2012	32 (53,3 %)	4 (6,7 %)	60
2012-2013	56 (69,1 %)	1 (1,2 %)	81
2013-2014	61 (62,9 %)	6 (6,2 %)	97
2014-2015	57 (60,6 %)	3 (3,2 %)	94
2015-2016	36 (65,5 %)	3 (5,5 %)	55
Total	375 (61,9 %)	28 (4,6 %)	606

Parmi les thèses de soins de santé primaires la proportion de thèses quantitatives est relativement stable autour de 60% depuis 10 ans.

COMPARAISON DU NOMBRE ANNUELS DE THÈSES CLASSÉES « SOINS DE SANTÉ PRIMAIRES » DE MÉTHODE QUANTITATIVE/MIXTE ET DE MÉTHODE QUALITATIVE

Tableau R7. Comparaison du nombre annuel de thèses de méthodologie quantitative ou mixte, qualitative, et autres méthodes/obstacle technique, rapport aux thèses incluses.

Promotion (document d'origine)	Effectif thèses quantitatives + mixtes	Effectif thèses qualitatives	Effectif méthode(s), « autre(s) » obstacle technique	Ratio (thèses quantitatives + mixtes) / thèses qualitatives	Effectif thèses soins de santé primaires
2006-2007	22 (53,7 %)	16 (39,0 %)	3 (7,3 %)	1,38	41
2007-2008	30 (65,2 %)	15 (32,6 %)	1 (2,2 %)	2,00	46
2008-2009	29 (78,4 %)	8 (21,6 %)	0 (0 %)	3,63	37
2009-2010	35 (76,1 %)	11 (23,9 %)	0 (0 %)	3,18	46
2010-2011	28 (57,1 %)	21 (42,9 %)	0 (0 %)	1,33	49
2011-2012	36 (60,0 %)	23 (38,3 %)	1 (1,7 %)	1,57	60
2012-2013	57 (70,4 %)	23 (28,4 %)	1 (1,2 %)	2,48	81
2013-2014	67 (69,1 %)	29 (29,9 %)	1 (1,0 %)	2,31	97
2014-2015	60 (63,8 %)	33 (35,1 %)	1 (1,1 %)	1,82	94
2015-2016	39 (70,9 %)	15 (27,3 %)	1 (1,8 %)	2,60	55
total	403 (66,5 %)	194 (32,0 %)	9 (1,5 %)	2,08	606

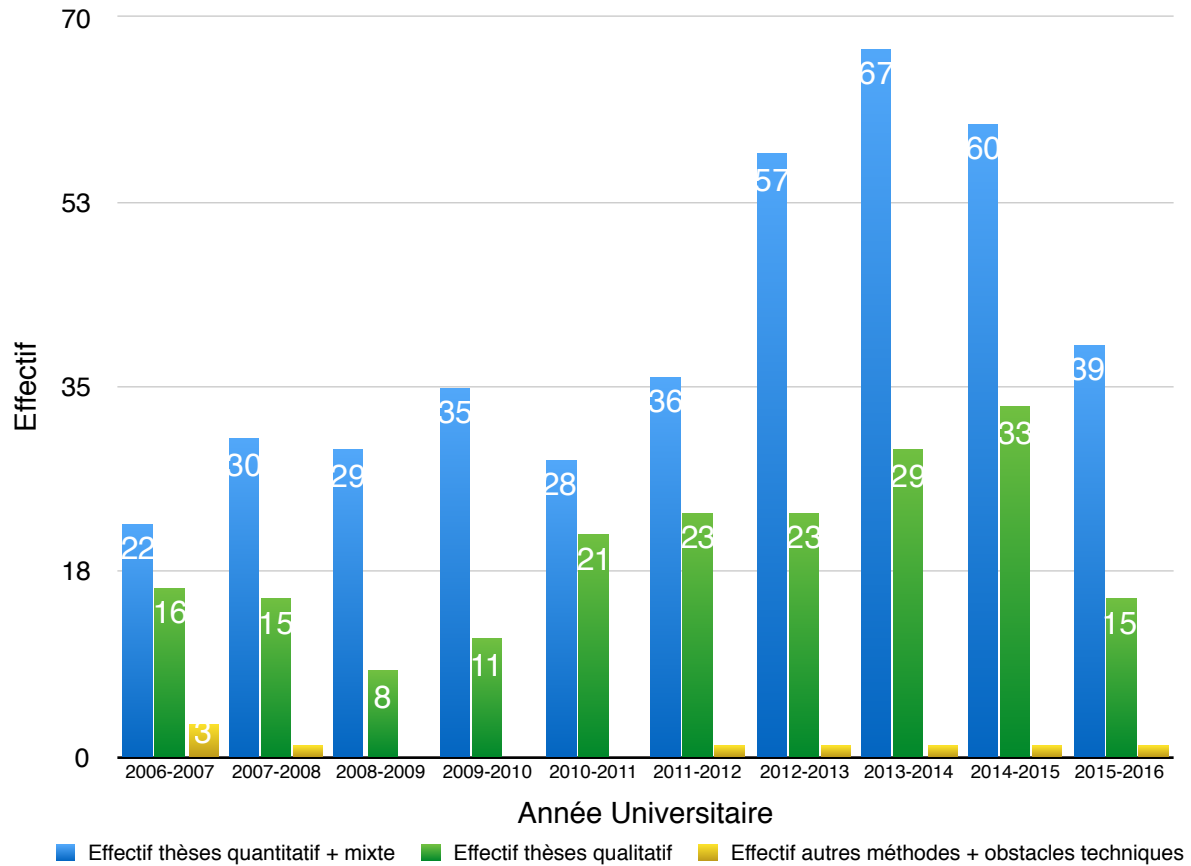
Ratio minimal : 1,33, promotion 2010-2011. (effectif annuel : N=49)

Ratio maximal : 3,63, promotion 2008-2009. (effectif annuel : N=37)

Médiane des ratios : 2,16. (effectif sur 10 ans : N=606)

Il y avait deux fois plus de thèses à méthode quantitative ou mixte que de thèses à méthode qualitative. La comparaison des effectifs année par année sur les 10 années par un test t de Student retrouvait une différence très significative ($p=1,84 \times 10^{-3}$).

Graphique R1. Représentation graphique des effectifs annuels comparés des thèses quantitative/mixtes, qualitatives, et méthodes autres/obstacles techniques.



TYOLOGIE DES MÉTHODES QUANTITATIVES OU MIXTES DÉDUITE DES RÉSUMÉS

MÉTHODE PRINCIPALE RELEVÉE SI PLUS D'UNE MÉTHODE CORRÉLATION AVEC LA DÉCLARATION DU THÉSARD

Au total, les typologies se regroupaient comme suit (quantitatif et mixte confondus) :

Type d'étude	Effectif
Études épidémiologiques/interventionnelles.	220 (54.6 %)
Évaluations de la qualité des soins.	119 (29.5 %)
Entretiens individuels structurés par questionnaire.	27 (6.7 %)
Évaluations/tests/diagnostics.	20 (5.0 %)
Recherches secondaires.	13 (3.2 %)
Études médico-économiques	4 (1.0 %)
Total	403 (100 %)

Tableau R8. Typologie des thèses de méthodologie quantitative ou mixte (N=403).
Corrélation aux déclarations du thésard dans le résumé.

Sous-type d'étude inspirés des critères de Paul FRAPPÉ (7)	Total quantitatif	Total mixte	Total méthodes consensuelles	Total méthodes non-consensuelles	Total méthodes absentes
1.1 : Epidémiologique, étude de cas / Série de cas	10 (2,7 %)	2 (7,1 %)	5 (41.7 %)	5 (41.7 %)	2 (16.7 %)
1.2 : Épidémiologique, étude écologique	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
1.3 : Epidémiologique, étude transversale	105 (28,0 %)	3 (10,7 %)	30 (27.8 %)	69 (63.9 %)	9 (8.3 %)
1.4 : Epidémiologique, Cas/témoins	10 (2,7 %)	0 (0,0 %)	5 (50,0 %)	4 (40,0 %)	1 (10,0 %)
1.5 : Epidémiologique, Cohorte	72 (19,2 %)	2 (7,1 %)	17 (23.0 %)	53 (71.6 %)	4 (5.4 %)
1.6 : Interventionnelle, Essai clinique	14 (3,7 %)	2 (7,1 %)	6 (37.5 %)	10 (62.5 %)	0 (0 %)
2.0 : Evaluation test diagnostic	19 (5,1 %)	1 (3,6 %)	3 (15,0 %)	17 (85,0 %)	0 (0 %)
3.1 : Médico-économique, minimalisation des coûts	1 (0,3 %)	0 (0 %)	0 (0 %)	0 (0 %)	1 (100 %)
3.3 : Médico-économique, coût-utilité	2 (0,5 %)	1 (3,6 %)	0 (0 %)	3 (100,0 %)	0 (0 %)
4.1 : Semi-Quantitatif, processus Delphi	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
4.2 : Semi-Quantitatif, groupe nominal	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	9 (2,4 %)	1 (3,6 %)	1 (10,0 %)	9 (90,0 %)	0 (0 %)
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	28 (7,5 %)	4 (14,3 %)	7 (21.9 %)	23 (71.9%)	2 (6.3 %)
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	10 (2,7 %)	0 (0 %)	1 (10,0 %)	7 (70,0%)	2 (20,0 %)
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	4 (1,1 %)	1 (3,6 %)	1 (20,0 %)	4 (80,0%)	0 (0 %)
5.5 : Evaluation de la Qualité des Soins, audit clinique	16 (4,3 %)	0 (0 %)	4 (25,0 %)	9 (56.3%)	3 (18.8 %)
5.6 : Evaluation de la Qualité des Soins, essai en population	11 (2,9 %)	1 (3,6 %)	2 (16.7 %)	8 (66.7%)	2 (16.7 %)
5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs	7 (1,9 %)	0 (0 %)	0 (0 %)	7 (100 %)	0 (0 %)
5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après	26 (6,9 %)	1 (3,6 %)	6 (22.2 %)	18 (66.7 %)	3 (11.1 %)
6.1 : Qualitative, entretien individuel structuré	20 (5,3 %)	7 (25,0 %)	7 (25.9 %)	18 (66.7 %)	2 (7.4 %)
6.2 : Qualitative, observation directe	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.3 : Qualitative, observation indirecte	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.4 : Qualitative, entretien individuel semi-structuré	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.4 : Qualitative, entretien individuel non structuré	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.6 : Qualitative, focus group	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
7.1 : Recherche secondaire, revue systématique	10 (2,7 %)	2 (7,1 %)	6 (50,0 %)	2 (16.7 %)	4 (33.3 %)
7.2 : Recherche secondaire, méta-analyse quantitative	1 (0,3 %)	0 (0 %)	1 (100 %)	0 (0 %)	0 (0 %)
7.3 : Recherche secondaire, méta-analyse qualitative	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Total :	375 (100 %)	28 (100 %)	102 (25,3 %)	266 (66,0 %)	35 (8,7 %)

Les sous-types dont l'effectif était nul ou correspondant aux méthodologiques qualitatives apparaissent en rouge.

PARAMÈTRES STATISTIQUES DES THÈSES DE MÉTHODE QUANTITATIVE OU MIXTE APPARAISSANT DANS LE RÉSUMÉ PRÉSENTATION EN FONCTION DE LA TYPOLOGIE

Tableau R9. Paramètres statistiques présents dans le résumé des thèses quantitatives et mixtes (N=403).

Sous-type d'étude inspirés des critères de Paul FRAPPÉ (7)	Nombre d'études comparatives	Nombre d'études randomisées	Nombre de p values / [IC]	Nombre d'analyses statistiques probables	Nombre d'analyses statistiques avec méthode précisée
1.1 : Epidémiologique, étude de cas / Série de cas	0 (0 %)	0 (0 %)	0 (0 %)	1 (8.3 %)	0 (0 %)
1.2 : Épidémiologique, étude écologique	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
1.3 : Epidémiologique, étude transversale	8 (7.4 %)	6 (5.6 %)	19 (17.6 %)	39 (36.1 %)	3 (2.8 %)
1.4 : Epidémiologique, Cas/témoins	8 (80,0 %)	0 (0 %)	2 (20,0 %)	4 (40,0 %)	3 (30,0 %)
1.5 : Epidémiologique, Cohorte	12 (16.2 %)	0 (0 %)	24 (32.4 %)	34 (46.0 %)	4 (5.4 %)
1.6 : Interventionnelle, Essai clinique	6 (37.5 %)	2 (12.5 %)	3 (18.8 %)	7 (43.8 %)	0 (0 %)
2.0 : Evaluation test diagnostique	1 (5,0 %)	1 (5,0 %)	5 (25,0 %)	10 (50,0 %)	1 (5,0 %)
3.1 : Médico-économique, minimalisation des coûts	1 (100 %)	0 (0 %)	1 (100 %)	1 (100 %)	0 (0 %)
3.3 : Médico-économique, coût-utilité	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
4.1 : Semi-Quantitatif, processus Delphi	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
4.2 : Semi-Quantitatif, groupe nominal	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	0 (0 %)	0 (0 %)	3 (30,0 %)	4 (40,0 %)	2 (20,0 %)
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	2 (6.3 %)	3 (9.4 %)	4 (12.5 %)	6 (18.8 %)	1 (3.1 %)
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	2 (20,0 %)	3 (30,0 %)	2 (20,0 %)	3 (30,0 %)	1 (10,0 %)
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	4 (80,0 %)	0 (0 %)	2 (40,0 %)	3 (60,0 %)	0 (0 %)
5.5 : Evaluation de la Qualité des Soins, audit clinique	1 (6.3 %)	1 (6.3 %)	3 (18.8 %)	5 (31.3 %)	0 (0 %)
5.6 : Evaluation de la Qualité des Soins, essai en population	2 (16.7 %)	1 (8.3 %)	1 (8.3 %)	3 (25,0 %)	0 (0 %)
5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs	2 (28.6 %)	0 (0 %)	1 (14.3 %)	1 (14.3 %)	0 (0 %)
5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après	3 (11.1 %)	1 (3.7 %)	7 (25.9 %)	14 (51.9 %)	1 (3.7 %)
6.1 : Qualitative, entretien individuel structuré	7 (25.9 %)	2 (7.4 %)	9 (33.3 %)	9 (33.3 %)	1 (3.7 %)
6.2 : Qualitative, observation directe	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.3 : Qualitative, observation indirecte	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.4 : Qualitative, entretien individuel semi-structuré	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.4 : Qualitative, entretien individuel non structuré	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
6.6 : Qualitative, focus group	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
7.1 : Recherche secondaire, revue systématique	1 (8.3 %)	0 (0 %)	0 (0 %)	3 (25,0 %)	0 (0 %)
7.2 : Recherche secondaire, méta-analyse quantitative	0 (0 %)	0 (0 %)	0 (0 %)	1 (100 %)	0 (0 %)
7.3 : Recherche secondaire, méta-analyse qualitative	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Total :	60 (14,9 %)	20 (5,0 %)	86 (21,3 %)	148 (36,7 %)	17 (4,2 %)

PARAMÈTRES STATISTIQUES ISSUS DE LA MÉTHODE APRÈS LECTURE DU TEXTE INTÉGRAL DES THÈSES

162 textes intégraux ont été lus sur les 165 éligibles, thèses dont le résumé comportait des arguments statistiques probables (N=148) ou certains (N=17), 2 thèses étaient introuvables en bibliothèque universitaire, 1 était sous embargo.

Tableau R10. Description des paramètres de l'analyse statistique issus de la méthode du texte intégral (N=162).

Description de l'analyse statistique déclarée par l'auteur :		Effectif
Calcul du NSN	Calcul du NSN présent	8 (4,9 %)
	Calcul du NSN absent	154 (95,1 %)
Détermination du risque α	α non-défini	63 (38,9 %)
	Détermination $\alpha=0,05$	97 (59,9 %)
	Détermination $\alpha<0,05$	1 (0,6 %)
	Détermination $\alpha>0,05$	1 (0,6 %)
Détermination du risque β	β non-défini	159 (98,2 %)
	Détermination $\beta=0,20$	2 (1,2 %)
	Détermination $\beta<0,20$	1 (0,6 %)
	Détermination $\beta>0,20$	0 (0 %)
Déclaration du recours à un tiers pour l'analyse statistique	Recours à un tiers pour l'analyse statistique	37 (22,8 %)
	Pas de recours à un tiers pour l'analyse statistique	125 (77,2 %)
Déclaration d'un logiciel utilisé pour l'analyse statistique	Aucun logiciel d'analyse déclaré	38 (23,5 %)
	Un seul logiciel d'analyse déclaré	90 (55,6 %)
	Deux logiciels d'analyse déclarés	27 (16,7 %)
	Trois logiciels d'analyse déclarés	7 (4,3 %)
Déclaration de l'analyse statistique des données quantitatives (continues)	Aucune analyse quantitative déclarée	79 (48,8 %)
	Une analyse quantitative déclarée	37 (22,8 %)
	Deux analyses quantitatives déclarées	33 (20,4 %)
	Trois analyses quantitatives déclarées	11 (6,8 %)
	Quatre analyses quantitatives déclarées	2 (1,2 %)
Déclaration de l'analyse statistique des données qualitatives (discrétisées, booléennes)	Aucune analyse qualitative déclarée	60 (37,0 %)
	Une analyse qualitative déclarée	54 (33,3 %)
	Deux analyses qualitatives déclarées	50 (30,9 %)
	Trois analyses qualitatives déclarées	1 (0,6 %)
Déclaration d'une analyse statistique complexe	Régression logistique	3 (1,9 %)
	Analyse univariée	5 (3,1 %)
	Analyse bivariée	8 (4,9 %)
	Analyse multivariée	24 (14,8 %)
	Total analyses statistiques complexes	40 (24,7 %)
Nombre total de méthodes lues, texte intégral		162 (100 %)

RECUEIL ET EFFECTIFS DES TESTS STATISTIQUES DE VARIABLES QUANTITATIVES (CONTINUES) DÉCLARÉS PAR L'AUTEUR DANS LA MÉTHODE

Tableau R11. Recueil des tests statistiques de variables quantitatives.

Type de test statistique pour variable quantitative déclaré par l'auteur :	Effectif
Test t de student	45 (25,3 %)
Non-défini	37 (20,8 %)
ANOVA simple test/analyse de variance	30 (16,9 %)
Test de la somme des rangs de Wilcoxon/U test de Mann-Whitney/Test de Wilcoxon Mann-Whitney	24 (13,5 %)
Test de Kruskal-Wallis	12 (6,7 %)
Test t de student pour groupes appariés	8 (4,5 %)
Test de Wilcoxon	6 (3,4 %)
Test des rangs signés de Wilcoxon	2 (1,1 %)
Test Z de différence minimale significative	2 (1,1 %)
Test r de corrélation de Bravais-Pearson	2 (1,1 %)
Test de Tukey	2 (1,1 %)
Test t de student pour données indépendantes	1 (0,6 %)
Test de distribution de Kolmogorov-Smirnov	1 (0,6 %)
Test des rangs de Spearman	1 (0,6 %)
Courbe de survie de Kaplan-Meier	1 (0,6 %)
Test unilatéral de comparaison de moyennes pour séries indépendante	1 (0,6 %)
Test F d'égalité de variance	1 (0,6 %)
Méthode de Kaplan-Meier	1 (0,6 %)
Calcul du coefficient Kappa	1 (0,6 %)
Total	178 (100 %)

L'effectif total (N=178) correspond à l'ensemble des déclarations des thésards, qui pouvaient déclarer entre 0 et 4 tests statistiques.

Les tests les plus fréquemment déclarés étaient le test t de Student, le test ANOVA et le test de la somme des rangs de Wilcoxon.

RECUEIL ET EFFECTIFS DES TESTS STATISTIQUES DE VARIABLES QUALITATIVES (DISCRÉTISÉES, BOOLÉENNES) DÉCLARÉS PAR L'AUTEUR DANS LA MÉTHODE

Tableau R12. Recueil des tests statistiques de variables qualitatives.

Type de test statistique pour variable qualitative déclarée par l'auteur :	Effectif
Test Khi-2 (de Pearson)	87 (44,9 %)
Test exact de Fisher	57 (29,4 %)
Non-défini	37 (19,1 %)
Test Khi-2 qualitatif	3 (1,6 %)
Test Khi-2 de Mc Nemar	2 (1,0 %)
Test de Bland et Altman	2 (1,0 %)
Test Khi-2 de tendance	1 (0,5 %)
Test de Cochran-Mantel-Haenszel	1 (0,5 %)
Test unilatéral de comparaison de valeurs quantitatives pour séries appariées	1 (0,5 %)
Calcul du coefficient intra-classe	1 (0,5 %)
Analyse de correspondance multiple ACM	1 (0,5 %)
Test Khi-2 pour données appariées	1 (0,5 %)
Total	194 (100 %)

L'effectif total (N=194) correspond à l'ensemble des déclarations des thésards, qui pouvaient déclarer entre 0 et 3 tests statistiques différents.

Les tests les plus fréquemment déclarés étaient le test de Khi-2 de Pearson et le test exact de Fisher.

RECUEIL ET EFFECTIFS DES LOGICIELS D'ANALYSE STATISTIQUE UTILISÉS, COMPARAISON AU RECOURS À UN TIERS HORS DU DMG POUR L'ANALYSE STATISTIQUE

Tableau R13. Recueil des logiciels d'analyse statistiques, proportion pour chaque logiciel du recours déclaré à un tiers extérieur intervenant

Nom du logiciel :	Effectif	Nombre de recours à une aide méthodologique externe au DMG pour l'analyse, effectif par logiciel
Non-défini	38 (18,7 %)	5 (13,2 %)
SAS	33 (16,3 %)	15 (45,5 %)
Excel/Numbers/Calc	32 (15,8 %)	6 (18,8 %)
IBM SPSS	26 (12,8 %)	7 (26,9 %)
Epi Info	26 (12,8 %)	8 (30,8 %)
R	13 (6,4 %)	1 (7,7 %)
BiostaTGV	8 (3,9 %)	3 (37,5 %)
STATA	7 (3,5 %)	1 (14,3 %)
EpiData	4 (2,0 %)	1 (25,0 %)
Microsoft Access	3 (1,5 %)	2 (66,7 %)
Sphinx Plus	2 (1,0 %)	1 (50,0 %)
Statview	2 (1,0 %)	0 (0 %)
SaTSCan	1 (0,5 %)	1 (100 %)
Statcal	1 (0,5 %)	0 (0 %)
STATPLUS	1 (0,5 %)	0 (0 %)
Google Forms	1 (0,5 %)	0 (0 %)
Prism	1 (0,5 %)	0 (0 %)
States	1 (0,5 %)	0 (0 %)
Statistica	1 (0,5 %)	1 (100 %)
peddro.org	1 (0,5 %)	0 (0 %)
MedCalc	1 (0,5 %)	0 (0 %)
Total	203 (100 %)	52 (25,6 %)

L'effectif total (N=203) correspond à l'ensemble des déclarations des thésards, qui pouvaient déclarer entre 0 et 3 logiciels différents.

Devant la faiblesse de certains effectifs, aucun test statistique n'a été réalisé.

Les logiciels les plus fréquemment déclarés étaient SAS, un tableur, SPSS et Epi info.

La plupart des thésards (3/4) n'ont pas déclaré de recours à une aide méthodologique externe au DMG pour l'analyse statistique.

RECHERCHE D'UNE CORRÉLATION ENTRE LA CATÉGORIE DU DIRECTEUR DE THÈSE ET LA DÉCLARATION DU RECOURS À UN TIERS POUR L'ANALYSE STATISTIQUE PAR LE THÉSARD

Tableau R14. Recherche d'une corrélation entre le recours à un tiers pour l'analyse statistique et la catégorie du directeur de thèse.

Catégorie directeur de thèse	Effectif thèses de méthode quantitative ou mixte.	Effectif thèses déclarant un recours à une aide méthodologique externe au DMG pour l'analyse statistique.
Enseignant du DMG	75 (18,6 %)	6 (16,2 %)
Médecin généraliste externe au DMG	98 (24,3 %)	11 (29,7 %)
Médecin du CHU	121 (30,0 %)	13 (35,1 %)
Médecin spécialiste externe au CHU	109 (27,1 %)	7 (18,9 %)
Total	403 (100 %)	37 (100 %)

Réalisation d'un test de Khi-2 comparant les effectifs des thèses selon la catégorie du directeur de thèse et les déclaration de recours à un tiers. $p=0,65$. Il n'y avait pas de différence significative.

RECHERCHE D'UNE CORRÉLATION ENTRE LA CATÉGORIE DU DIRECTEUR DE THÈSE ET LA PRÉSENCE DE PARAMÈTRES STATISTIQUES OU ÉVOCATEURS D'ANALYSE STATISTIQUE DANS LE RÉSUMÉ

Tableau 15. Recherche d'une corrélation entre les arguments significatifs de l'analyse statistique dans le résumé et la catégorie du directeur de thèse.

Catégorie directeur de thèse	Effectif thèses de méthode quantitative ou mixte.	Effectif thèses dont l'analyse statistique était possible d'après le résumé	Effectif thèses dont l'analyse statistique était certaine d'après le résumé	Effectif thèses dont l'analyse statistique apparaissait dans le résumé (<i>p values</i> , [IC])
Enseignant du DMG	75 (18,6 %)	23 (15,5 %)	2 (11,8 %)	10 (11,6 %)
Médecin généraliste externe au DMG	98 (24,3 %)	31 (21,0 %)	3 (17,7 %)	16 (18,6 %)
Médecin du CHU	121 (30,0 %)	59 (39,9 %)	10 (58,8 %)	32 (37,2 %)
Médecin spécialiste externe au CHU	109 (27,1 %)	35 (23,6 %)	2 (11,7 %)	28 (32,6 %)
Total	403 (100 %)	148 (100 %)	17 (100 %)	86 (100 %)

Réalisation de plusieurs tests de Khi-2 comparant les effectifs selon la catégorie du directeur de thèse et les paramètres statistiques.

- Concernant la présence des *p values* et des [IC] dans le résumé, $p=0,18$. Il n'y avait donc pas de différence statistiquement significative.
- Concernant la présentation des test statistiques et des résultats significatifs dans le résumé, le test de Khi-2 ne put être réalisé pour cause d'effectif insuffisant ($N < 5$). Un test exact de Fisher le remplaça, retrouvant $p=0,13$. Il n'y avait donc pas de différence statistiquement significative.
- Concernant la présence de mots évocateurs d'analyse statistique dans le résumé (significatif et synonymes), $p=0,19$. Il n'y avait donc pas de différence statistiquement significative.

AUTRE CALCULS

Ont été décrits et analysés d'autres paramètres :

Sexe de l'auteur	(annexe 5.1)
Mention(s)	(annexe 5.2)
Nombre de pages de la thèse	(annexe 5.3)
Critères de formes : structure IMRaD, résumé français ou bilingue, disponibilité en ligne.	(annexe 5.4)
Directeurs de thèses.	(annexe 5.5)

Ces résultats sont joints en annexe (annexe 5).

DISCUSSION

RÉSULTATS PRINCIPAUX

Inclusion soins de santé primaires

1179 thèses furent incluses, et ce nombre est important, bien qu'invérifiable via d'autres sources (SUDOC retrouve des résultats supérieurs, les raisons imputables étant vraisemblablement l'inclusion des thèses postérieures au 28 juin 2016, ainsi que certaines thèses n'étant ni de médecine générale, ni de Bordeaux).

L'inclusion « soins de santé primaires » des titres par mots-clés est plutôt satisfaisante, car le recours au directeur de thèse n'a eu lieu que dans 18 cas, soit 1,5 % des thèses.

Ce taux annuel de thèses de soins de santé primaires a **augmenté progressivement de 19%** sur 10 ans, passant de 41,4 % (41 pour 99) en 2006-2007 - valeur minimale sur la période étudiée - à 60,4 % (55 pour 91) en 2015-2016. La moyenne sur 10 ans est de 51,4 % de thèses de soins de santé primaires, soit 606 thèses, avec une tendance annuelle à l'augmentation.

Détermination de la méthode

Les différents tris ont abouti à la sélection de 403 thèses de soins de santé primaires à méthode quantitative ou mixte, respectivement 375 pour la méthode quantitative et 28 pour une méthode mixte.

Il est important de noter qu'une méthode mixte associe une part de recherche quantitative et une part de recherche qualitative, soit deux méthodes distinctes. Cette distinction n'a pas été faite pour les études incluant deux méthodes différentes de recherche quantitative.

La proportion de thèses mixtes varie entre 0% et 10,9 % (5 pour 46 en 2009-2010), en moyenne 4,6 % sur la période (28 pour 606).

Concernant les thèses de méthode qualitative, la proportion est d'environ un tiers contre deux tiers de méthodes quantitative ou mixte. Ce ratio est **relativement stable** (moyenne 2,23 ; écart-type 0,76 ; médiane 2,16) et semble peu fluctuer dans le temps.

Il serait pertinent de vérifier si ces tendances (à la stabilité) se maintiennent dans les années à venir.

Concernant les obstacles techniques, leur taux est très faible, 1,5 % (9 thèses sur 606 incluses), et ne remet pas en question les résultats obtenus.

Typologie de la méthode

Les typologies les plus représentées sont, pour 403 méthodes déterminées quantitatives ou mixtes, majoritairement quantitatives (220 soit 52,36%), suivies des évaluations de la qualité des soins (119 soit 29,53%) regroupant donc presque 80% des typologies employées par les thésards.

Les thésards privilégieraient donc la recherche quantitative et l'évaluation de la qualité des soins, ce qui est cohérent avec la façon dont nous sommes formés : le raisonnement EBM.

Au sein de ces deux groupes majoritaires, les typologies les plus représentées sont les **études transversales** (108 soit 26,8 %) , suivies des **cohortes** (74 soit 18,4 %), des **enquêtes de pratiques** (32 soit 7,9 %), des **études observationnelles avant-après** (27 soit 6,7 %) et des **entretiens individuels structurés** (27 soit 6,7 %). Le reste des typologies représente les 20% restants.

Un argument vraisemblable étayant ce choix, outre les convictions personnelles du thésard et ses propres motivations, est l'exploitation des bases de données médicales hospitalières qui font partie de l'environnement de travail durant l'internat et le recours à des questionnaires diffusés via le réseau du personnel de la faculté.

Les données étayant ce propos sont compilées dans l'annexe 5 tableau A10.

Certaines typologies proposées par la classification de Paul FRAPPÉ (Annexe 3) n'ont pas été retrouvées, respectivement l'étude écologique, celles de coût-utilité, de coût-bénéfice, semi-quantitatives (processus Delphi et groupe nominal) et méta-analyses qualitatives.

Outre les études semi-quantitatives et qualitatives qui avaient vraisemblablement été classées comme « qualitatives » lors du tri du titre ou du résumé, les autres types d'études n'ont pas été retrouvés car ils ne correspondraient pas aux problématiques habituelles des travaux de thèses.

Paramètres statistiques du résumé

Comparés à l'ensemble des typologies, ces paramètres n'ont pas de grande signification. Ils s'avèrent pertinents selon certaines typologies et améliorent la validité scientifique.

Dans les thèses, les résultats sont globalement en faveur de **designs simples**.

Peu d'études se démarquent par des designs ambitieux. Cependant, le résumé n'est pas forcément le meilleur reflet du travail de thèse.

Cela peut s'expliquer par un impératif de temps dans le travail de thèse, car chaque étudiant a une date limite à ne pas dépasser (6 ans après l'inscription en DES de médecine générale, soit 3 ans après la fin de l'internat de médecine générale dans une majorité de cas).

Méthodes statistiques des textes intégraux

Ont été analysées 162 thèses, toutes typologies confondues.

Outre le calcul du nombre de sujets nécessaires ou du risque β qui dépend du design de l'étude, le risque α était absent de 38,9 % des méthodes, mais il semblait difficile de réaliser des tests statistiques sans définir ce risque, ce qui en ferait probablement un oubli de la part de l'auteur.

De même, 23,5 % des auteurs oubliaient de déclarer le logiciel d'analyse statistique qu'ils avaient utilisé, une majorité n'en déclarait qu'un seul (55,6 %).

Les logiciels les plus déclarés étaient :

- **SAS** (16,3 %), **IBM SPSS** (12,8 %). Logiciels d'analyse statistique **payants**, sous licence.
- **Epi Info** (12,8 %), **R** (6,4 %), **Biostatgv** (3,9 %). Logiciels d'analyse statistique **gratuits**.
- **Excel/Numbers/Calc** (15,8 %). **Tableurs**, présents sur l'ordinateur de tout étudiant, Calc étant un logiciel gratuit.

SAS serait le logiciel le plus utilisé par les thésards, suivi par les tableurs puis SPSS.

Concernant l'absence d'analyse quantitative ou qualitative, cette information est à considérer comme basée sur la supposition de la présence de ces tests d'après les résultats apparaissant dans le résumé, leur valeur est certainement sur-évaluée.

Concernant les analyses statistiques complexes, 40 thèses en ont déclaré au moins une, ce qui est peu comparé à l'effectif global de thèses quantitatives / mixtes (40 pour 403 soit 9,9 %). Sans prendre en compte la nature du travail du thésard, il est difficile de juger cette proportion.

Concernant les tests de variables quantitatives, les plus utilisés étaient le **test t de Student** (25,3 %), le test **ANOVA** (16,9 %) ainsi que le **test de la somme de rangs de Wilcoxon** (13,5 %).

Concernant les tests de variables qualitatives, les plus utilisés étaient le **test Khi-2** (44,9 %) et le **test exact de Fisher** (29,4 %). Ces tests étaient d'ailleurs souvent déclarés dans les méthodes par paire, comme suit : « test de khi-2 ou test exact de Fisher si effectif insuffisant ».

En revanche, les déclarations de la réalisation d'un test t de Student n'étaient que rarement associées à leurs conditions minimales (recherche d'une distribution normale, effectif minimal $N=30$).

À l'exception du test ANOVA, ces quatre tests étaient tous abordés au cours de nos études de médecine dans le tronc commun et revus au cours du DES de médecine générale.

Les thésards déclaraient majoritairement un ou deux tests.

Recours à une aide méthodologique, lien avec la méthode statistique

Les recours à un tiers pour l'aide méthodologique statistique était déclaré dans 52 méthodes, soit 25,6 % des méthodes lues. Ce nombre étant déclaratif, il est possible qu'il soit sous-évalué.

Il semblerait que le recours à un tiers influence le choix du logiciel de calcul statistique, le logiciel SAS étant associé pour 45,5 % au recours à un tiers. De même, certains logiciels comme Statistica ou SaTSCan étant associés à 100% au recours à un tiers, il est par conséquent vraisemblable que ces tiers aient fourni leur logiciel de calcul statistique au thésard ou une licence pour l'utiliser.

Le recours à un tiers ne semble pas influencé par la catégorie du directeur de thèse. Par conséquent ce recours dépendrait au cas-par-cas du travail de thèse sans que l'on ne puisse établir de lien.

Influence du directeur de thèse sur la méthode statistique

La catégorie du directeur de thèse n'influence pas de manière statistiquement significative les arguments de l'analyse statistique, cependant les *p values* obtenues sont toutes inférieures à 0,20. Les thèses dirigées par des médecins du CHU contiendraient plus souvent des arguments statistiques.

Cela s'explique vraisemblablement par une meilleure accessibilité aux données hospitalières, à l'exercice d'une médecine spécialisée, ainsi qu'à un grand nombre de publications médicales issues du CHU.

Il ne faut pas oublier que c'est le thésard qui choisit son directeur de thèse, non le contraire.

COMPARAISON DES RÉSULTATS PRINCIPAUX AVEC LES TRAVAUX ANTÉRIEURS

Les données disponibles sont peu nombreuses et redondantes. Il s'agit surtout de constats et de recommandations pour les enseignants.

Nous avons confronté nos résultats à des thèses de médecine générale, pour la période étudiée, en Aquitaine. Nous avons relevé trois thèses dont les résultats permettaient de critiquer notre travail.

Deux thèses traitaient des thèses de médecine générale

- LES THÈSES EN MEDECINE GÉNÉRALE. ÉTUDE DESCRIPTIVE DES THÈSES DES FUTURS MÉDECINS GÉNÉRALISTES SOUTENUES À BORDEAUX ENTRE 2007 ET 2014, par Maylis LABERNÈDE (2), dirigée par Gérard DUCOS.

Les points essentiels étaient que cette thèse avait une méthode totalement différente concernant les critères de tri comme soins de santé primaires (basée sur la composition du jury) et concernant la lecture du résumé, basée sur les critères WONCA (10). Elle s'est également intéressée aux thèmes.

- THÈMES, OBJETS ET MÉTHODES DE RECHERCHE DES THÈSES DES INTERNES DE MÉDECINE GÉNÉRALE DE LA SUBDIVISION OCÉAN INDIEN SOUTENUES ENTRE 2007 ET 2014, par Alix DE CUSSY-BOUVET (12), dirigée par Jean-Marc FRANCO.

Les points essentiels étaient que cette thèse n'avait inclus que les internes de la subdivision océan indien, pour un total de 144 thèses, en revanche elle a également inclus les DESC médecine d'urgence, représentant 31 thèses soit 22 % de la population étudiée. Elle s'est également intéressée à la composition du jury et aux thèmes.

Une comparaison exhaustive avec ces deux thèses a été réalisée en annexe (annexe 6). Il en ressort que les résultats ne sont pas comparables.

Cette différence peut s'expliquer par un paradoxe de Simpson (11), des méthodes différentes, d'un côté une évaluation qualitative dans notre travail basée sur des mots-clés et concepts, d'un autre côté une évaluation quantitative plus objective basée sur la composition du jury et la proportion de médecins généralistes. D'autre part, la classification des sous-types de méthodes de Maylis LABERNÈDE ne correspondait pas exactement à celle que nous avons utilisée. De même, les classifications proposées en cours de DES de médecine générale étaient encore différentes, et aucune ne semble universelle.

En revanche, concernant la proportion de thèses de soins de santé primaires nos observations sont cohérentes avec l'augmentation de la proportion de médecins généralistes, de médecins généralistes enseignants et de membres du Département de Médecine Générale dans les jurys de thèses, comme observé par Maylis LABERNÈDE dans son travail (1).

Une thèse traitait des difficultés rencontrées par les internes dans la rédaction de leur thèse

- DIFFICULTÉS RENCONTRÉES PAR LES MÉDECINS GÉNÉRALISTES POUR DIRIGER UNE THÈSE DE MÉDECINE GÉNÉRALE EN AQUITAINE, par Jean-Baptiste HARRIAGUE (13), dirigée par Jean-Philippe JOSEPH.

Dans la première partie, une analyse des thèses de médecine générale fut réalisée. L'amplitude allait de 2003-2004 à 2007-2008, soit deux années en commun, 2006-2007 et 2007-2008. La proportion de médecins généralistes directeurs de thèse avait également été établie.

Concernant l'année 2006-2007, le nombre total de thèses était le même : 99.

Concernant l'année 2007-2008, le nombre total de thèses était également le même : 100.

Des résultats différents furent notés, Jean-Baptiste HARRIAGUE retrouvait respectivement 23 et 26 thèses de soins primaires (dénomination ancienne des soins de santé primaires), contre 41 et 46 selon la classification du présent travail. Cette différence s'explique vraisemblablement par une plus grande tolérance des critères utilisés dans notre travail, en particulier « recherche initiée par le DMG » et « L'étude est réalisée en coopération avec une autre spécialité, ou une autre profession, mais intéresse les soins de santé primaires ». Ces résultats concordent avec l'observation précédente.

LIMITES DU TRAVAIL DE RECHERCHE

La majorité du travail fut relevée à la main, l'automatisation n'a eu lieu qu'au terme de la saisie. Cependant, n'ayant été réalisé que par une seule personne, en dépit de multiples relectures, certaines erreurs ont pu passer inaperçues dans celui-ci.

Concernant la population

La population est considérée comme exhaustive, mais nous n'avons pas mis en évidence de moyen de vérifier le nombre total supposé.

Concernant le tri des titres

L'analyse des titres selon le sens des mots-clés était la seule option possible pour déterminer l'inclusion. Les critères tirés de la définition WONCA des soins de santé primaires (10), proposés dans la méthode, ont été retravaillés pour être les plus discriminants possibles tout en étant peu nombreux pour limiter les erreurs, gardant pour cible l'intérêt pour tout praticien dans l'exercice des soins de santé primaires.

Concernant le tri des résumés

L'analyse des résumés posa plus de problèmes que celle des titres. Le sens des phrases dans leur ensemble était indispensable à leur tri. Cependant, les résumés suivaient une structure globale continue entre « indicatif » et « informatif » (9), sans pour autant présenter de réelle dichotomie entre les deux. Nous avons choisi les variables d'une part pour leurs disponibilités sur SUDOC, et d'autre part à titre d'appréciation et de facilité de lecture du résumé, établies comme critères de forme indépendants du contenu scientifique mais offrant un confort de lecture.

Notre relevé de données était plus ciblé que la lecture critique proposée par la grille STROBE (17), car nous avons choisi les variables à relever en fonction des besoins pour l'enseignement des étudiants, et leur choix avait été discuté avec le directeur de thèse.

Concernant le critère de jugement principal

Il serait plus juste de le qualifier de « proposition de tri selon des critères supposés consensuels ». La classification de Paul FRAPPÉ est une classification qui nous semblait détaillée et exhaustive, couvrant le plus grand nombre de typologies possibles.

Cependant, la cohérence - entre les déclarations du thésard concernant la typologie de sa méthode et celle retrouvée lors de notre lecture systématique - est globalement mauvaise. Les déclarations ne sont pas consensuelles dans 66,0 % des cas.

Néanmoins, si nous avons reporté strictement ce que le thésard déclarait dans son résumé, une partie des typologies aurait nettement manqué de précision. De plus, les déclarations non

consensuelles regroupaient d'une part des contradictions avec nos déductions, d'autre part une faible rigueur de dénomination, laquelle paraissait majoritaire sans que nous ne puissions en déterminer le pourcentage.

Dans 8,68% des cas le thésard ne déclarait aucune typologie de méthode dans son résumé. Nous déduisons la typologie du design vraisemblable de son étude, et nous la confrontons par une lecture du texte intégral quand il était disponible, ce qui prenait énormément de temps.

Concernant la lecture des textes intégraux

L'analyse des textes intégraux fut inspirée par les recommandations du directeur de thèse et regroupait les éléments clés directement liés au niveau de preuve des travaux de recherche (16, 17). Cependant, l'inclusion des thèses pour lecture du texte intégral reposait sur des critères subjectifs et il est vraisemblable que certains travaux comportant une analyse statistique complexe aient été exclus à tort.

La seule méthode pragmatique permettant d'avoir une vision objective de l'ensemble des thèses serait de toutes les lire dans leur intégralité, ce qui constituerait un pari risqué, difficile à mener à terme, surtout pour une seule personne.

Concernant les paramètres statistiques

Les paramètres ont été relevés de manière exhaustive sur l'ensemble des textes intégraux disponibles. Leur choix avait été discuté avec le directeur de thèse, conforté par la grille de lecture STROBE (17).

Concernant la déclaration de recours à un tiers

Cette déclaration ne préjuge pas que le travail ait réellement été réalisé par une seule personne, sans recours à un tiers (petit ami, connaissance), par conséquent la valeur relevée doit être sous-estimée.

Concernant le directeur de thèse

Notre travail n'a pas étudié la composition des jurys, ni les directeurs des thèses hors soins primaires ou qualitatives. Nos conclusions en ce sens sont limitées.

Concernant la validité interne

Ce travail fut long, saisi à la main et en dépit de nombreuses relectures nous supposons que le coefficient kappa serait faible.

Les travaux de ce genre sont actuellement peu nombreux car ils nécessitent beaucoup d'investissement personnel de la part de l'auteur.

BIAIS

Biais de selection

Il n'y en a aucun a priori.

L'effectif est considéré comme exhaustif mais nous ne pouvons vérifier ce postulat.

Biais de classement

Il y en a trois successifs, à chaque tri.

Les trois étapes clés de l'avancée du travail reposaient sur la perception du sens d'une phrase ou d'un paragraphe :

- Une évaluation subjective du titre pour l'inclusion « soins de santé primaires », guidée par les critères adaptés WONCA. Nous avons cependant démontré que la lecture systématique du titre était plus performante que la recherche de mots-clés (tableaux R2 et R3). L'inclusion a vraisemblablement été réalisée par excès.
- Une évaluation subjective du résumé pour déterminer si la méthode était quantitative ou mixte.
- Une détermination subjective de la typologie - critère de jugement principal - dirigée par les critères de Paul FRAPPÉ (annexe 3) et l'identification de la question de recherche (illustration M1).

De plus, l'identification de mots-clés déterminant des critères évocateurs d'analyse statistique dans le résumé, pour lecture du texte intégral, reposait sur la recherche de ces derniers. Leur éventail était le plus large possible mais représentait un risque d'inclure par excès des thèses ne comportant aucune analyse statistique. Inversement, nous avons vraisemblablement exclu des thèses qui en comportaient une.

Nous ne pouvons fournir les valeurs de sensibilité ou de spécificité (versus Gold-Standard : lecture systématique de tous les textes intégraux) de notre tri, dans son ensemble ou en fonction de l'étape considérée. Nous n'avons pas trouvé d'autre solution et avons conscience que ces biais étaient impossibles à limiter davantage.

Biais de confusion

Il existe certainement un biais de confusion entre la catégorie du directeur de thèse et la présence d'arguments statistiques dans le résumé car les *p values* sont inférieures à 0,20 .

Nous avons réalisé peu de tests statistiques et les effectifs étaient de faible taille.

PERSPECTIVES

Notre travail a permis de réaliser un catalogue se voulant le plus exhaustif possible du contenu des thèses. Il permettra de fournir des arguments pour l'enseignement, afin de déterminer ce dont les futurs thésards pourraient avoir besoin pour leurs travaux, en se basant sur les choix de leurs prédécesseurs. Une liste des logiciels les plus utilisés et des tests statistiques les plus souvent réalisés est disponible pour les enseignants.

Nous avons fait le choix d'examiner les thèses selon les paramètres les plus pragmatiques possibles, pour les comparer, rechercher des tendances et des liens, sans parvenir à fournir de résultat statistiquement significatif. Ce travail s'inscrit dans une liste, certes modeste, mais grandissante de « thèses sur les thèses », dont la finalité sera d'avoir de multiples sources à inclure dans une analyse secondaire ultérieure, pour vérifier les résultats obtenus et améliorer la validité scientifique.

Notre classification, basée sur les travaux de Paul FRAPPÉ, ne saurait être universelle. Elle confirme cependant que 80% des travaux de recherche à méthode quantitative ont une typologie quantitative « pure » ou d'évaluation de la qualité des soins.

Il apparaît que les étudiants et les enseignants définissent les méthodes de recherche selon une terminologie qui ne répond à aucun consensus, alors qu'il serait pertinent d'en établir un, par exemple à l'occasion d'une conférence ou lors de groupes de travail au sein des collègues d'enseignement.

D'autre part, nous avons mis en évidence des oublis, tant dans les résumés que dans les méthodes des textes intégraux. Une attention particulière devrait y être portée, car ils sont préjudiciables pour la qualité des mémoires.

Enfin, la recherche quantitative représente environ 2/3 des travaux de recherche, contre 1/3 de travaux de recherche qualitative. Par conséquent, cela devrait avoir un impact sur la répartition du temps nécessaire à la formation des étudiants.

Au total, notre travail intéressera vraisemblablement les enseignants et permettra peut-être d'influencer leurs décisions dans la façon d'organiser la formation.

Si des travaux ultérieurs devaient être réalisés, le fichier du tableur ayant permis la réalisation du présent mémoire est disponible sur demande.

CONCLUSION

Cette étude descriptive des 1179 thèses couvrant la période du 14 septembre 2005 au 28 juin 2016 issues du répertoire de la faculté de médecine de Bordeaux a permis de mettre en évidence les points suivants :

- Les soins de santé primaires sont une thématique de plus en plus présente dans les thèses de médecine générale, avec une augmentation sur 10 ans de 19%.
- Environ deux tiers de thèses de ces travaux avaient une méthode quantitative contre un tiers de méthodes qualitatives, avec un ratio relativement constant.
- Au sein des travaux quantitatifs, 40,2 % bénéficiaient d'une analyse statistique, dont environ un quart d'analyse complexe (régression logistique, analyse uni/multivariée).
- Les typologies de méthodes quantitatives les plus représentées étaient les études transversales (26,8 %), les cohortes (18,4 %), les enquêtes de pratiques (7,9 %), les études observationnelles avant-après (6,7 %) et les entretiens individuels structurés (6,7 %).
- Le calcul du nombre de sujets nécessaires et du risque β était peu réalisé.
- Dans la limite des résumés et textes intégraux lus, des oublis ont été relevés, comme l'absence de déclaration du risque α (38,9 % des cas), du logiciel de travail (23,5 % des cas) ou des tests statistiques réalisés. La méthode n'était pas déclarée par le thésard dans le résumé dans 8,7 % des cas.
- Les logiciels les plus utilisés étaient SAS (16,3 %), Excel/Number/Calc (15,8 %), SPSS (12,8 %), Epi Info (12,8 %) et R (6,4 %).
- Les tests statistiques quantitatifs les plus utilisés étaient le test t de Student (25,3 %), le test ANOVA (18,9 %) et le test de la somme des rangs de Wilcoxon (13,5 %).
- Les tests statistiques qualitatifs les plus utilisés étaient le test Khi-2 (44,9 %) et le test exact de Fisher (29,4 %).
- Le recours à un tiers externe au DMG pour l'analyse statistique n'aurait pas de lien avec le directeur de thèse. Cependant, notre travail ne peut déterminer les facteurs influençant ce recours, d'autres travaux pourraient s'y intéresser.
- Un directeur de thèse membre du CHU pourrait avoir une influence sur la présence d'arguments statistiques dans le résumé, sans lien statistiquement significatif. Une étude sur un plus grand échantillon de thèses, et une analyse des jurys en fonction de l'année, pourrait y apporter une réponse.
- N'oublions pas que l'étudiant choisit son jury, non le contraire.

Au total, nous avons observé 10 années de thèses, et répondu à la majorité de nos objectifs. Seuls le recours à un tiers pour l'analyse méthodologique et le choix du directeur de thèse n'ont pas apporté de réponse significative, mais pourraient donner lieu à d'autres travaux.

RÉFÉRENCES

- (1) Décret n° 2004-67 du 16 janvier 2004 relatif à l'organisation du troisième cycle des études médicales, Legifrance. 2004 [Internet] [consulté le 20 mars 2017] Disponible sur : <https://www.dumg-toulouse.fr/uploads/4883b73b7e072133b41447e363bc16e2a74bc6c7.pdf>
- (2) Labernede M. Les thèses dans la filière de médecine générale : étude descriptive des thèses des futurs Médecin généralistes soutenues à Bordeaux entre 2007 et 2014. Thèse de médecine Bordeaux: Université de Bordeaux; 2016.
- (3) Haller DM et al. Recherche scientifique en médecine de famille : expériences, barrières et besoins des praticiens. *Revue Médicale Suisse*; 2011;1089–1094.
- (4) ARS Ile-de-France, Affectation des étudiants en médecine reçus aux ECN 2015. Service de la Démographie, Direction de l'Offre de Soins et Médico-Sociale; ARS; 2016 [Internet] [consulté le 20 mars 2017] Disponible sur : http://www.iledefrance.paps.sante.fr/fileadmin/ILE-DE-FRANCE/PAPS/Je_me_forme/Devenir_medecin/Affectations-etudiants-medecine-ECN-2015.pdf
- (5) Rivière E. Analyse du choix des 7658 étudiants en médecine après les ECN 2012. Mémoire de DU de pédagogie dans l'enseignement supérieur, Université de Bordeaux; 2013 [Internet] [consulté le 20 mars 2017] Disponible sur : <https://dumas.ccsd.cnrs.fr/dumas-00989665/document>
- (6) Instruction interministerielle N° DGOS/RH1/DGRH/A2-3/2016/198 du 16 juin 2016 relative à l'ouverture, en 2016, de postes de chefs de clinique universitaires de médecine générale dans le cadre du pacte territoire santé II. Legifrance 2016 [Internet] [consulté le 20 mars 2017] Disponible sur : http://circulaire.legifrance.gouv.fr/pdf/2016/06/cir_41055.pdf
- (7) Aubin-Auger I, Stalnikiewicz B, Mercier A, Lebeau JP, Baumann L. Diriger une thèse qualitative : difficultés et solutions possibles. *Exercer* 2010;93:111-4.
- (8) Teisset Y, Lehr-Drylewicz AM. Perception du travail de thèse par les internes de médecine générale de la faculté de Tours. *Exercer* 2012;102:127-8

- (9) de Pouvoirville G. Développer la recherche en médecine générale et en soins primaires en France: Propositions. Rapport remis à Monsieur le Ministre de la Santé et à Monsieur le Ministre Délégué à la Recherche par Gérard de Pouvoirville, Directeur de Recherche au CNRS, coordonnateur du Comité d'Interface INSERM/Médecine Générale. 2006 [Internet] [consulté le 20 mars 2017] Disponible sur : http://extranet.inserm.fr/index.php/layout/set/print/content/download/25726/163582/file/Rapport_Pouvoirville.pdf
- (10) Gay B. La définition européenne de la médecine générale-médecine de famille. Centre Universitaire de Médecine Générale, UCL, Bruxelles, Belgique 2002.
- (11) Frappe P. Initiation à la recherche. Paris : co ed GMS/CNGE, 2011.
- (12) de Cussy-Bouvet A. Thèmes, objets et méthodes de recherche des thèses des internes de médecine générale de la subdivision Océan Indien soutenues entre 2007 et 2014. Thèse de médecine Bordeaux: Université de Bordeaux; 2016.
- (13) Harriague JB. Difficultés rencontrées par les médecins généralistes pour diriger une thèse de médecine générale en Aquitaine. Thèses de médecine Bordeaux: Université de Bordeaux; 2010.
- (14) Quinton A, Emeriau JP. Comment faire une thèse ? Document pédagogique du CRAME; DMG Bordeaux. 2013 [Internet] [consulté le 20 mars 2017] Disponible sur : <https://fr.scribd.com/document/82819191/Comment-Faire-Unethese>
- (15) Maisonneuve H. Guide pratique du thésard, 7ème édition. Ressource Internet, Editions Scientifique LC, SANOFI. 2010 [Internet] [consulté le 20 mars 2017] Disponible sur : <http://www.reims.cnge.fr/IMG/pdf/guide.pdf>
- (16) HAS. Etat des lieux Niveau de preuve et gradation des recommandations de bonnes pratiques. Haute Autorité de Santé. 2013 [Internet] [consulté le 20 mars 2017] Disponible sur : http://www.has-sante.fr/portail/upload/docs/application/pdf/2013-06/etat_des_lieux_niveau_preuve_gradation.pdf
- (17) Gedda M, Riche B. Traduction française des lignes directrices SAMPL pour l'écriture et la lecture des méthodes et analyses statistiques. Kinésithérapie la Revue. 2015;15(157): 69-74.

TABLES ET INDEX

TABLE DES TABLEAUX

	Page
Table des Tableaux relatifs aux résultats.	
Tableau R1. Répartition des inclusions par promotion.	23
Tableau R2. Identification de la méthode par mots-clés.	24
Tableau R3. Identification de la méthode par lecture systématique du titre, amélioration des effectifs par rapport au tri par mots-clés.	24
Tableau R4. Identification de la méthode par lecture systématique du résumé, amélioration des effectifs par rapport au tri par lecture systématique du titre.	25
Tableau R5. Comparaison des tris par mots-clés du titre, lecture du titre, et lecture du résumé.	25
Tableau R6. Comparaison des effectifs annuels des thèses de méthode quantitative et mixte par rapport aux thèses incluses « soins de santé primaires ».	26
Tableau R7. Comparaison du nombre annuel de thèses de méthode quantitative ou mixte, qualitative, et autres méthodes/obstacle technique, rapport aux thèses incluses.	27
Tableau R8. Typologie des thèses de méthode quantitative ou mixte (N=403). Corrélation aux déclarations du thésard dans le résumé.	29
Tableau R9. Paramètres statistiques présents dans le résumé des thèses quantitatives et mixtes (N=403).	30
Tableau R10. Description des paramètres de l'analyse statistique issus de la méthode du texte intégral.	31
Tableau R11. Recueil des tests statistiques de variables quantitatives.	32
Tableau R12. Recueil des tests statistiques de variables qualitatives.	33
Tableau R13. Recueil des logiciels d'analyse statistiques, proportion pour chaque logiciel du recours déclaré à un tiers extérieur intervenant	34
Tableau R14. Recherche d'une corrélation entre le recours à un tiers pour l'analyse statistique et la catégorie du directeur de thèse.	35
Tableau R15. Recherche d'une corrélation entre les arguments significatifs de l'analyse statistique dans le résumé et la catégorie du directeur de thèse.	36

Table des Tableaux relatifs aux annexes.	Page
Tableau A1. Comparaison des effectifs d'hommes et de femmes, dans les thèses incluses et dans les thèses totales, calcul des sex-ratios.	66
Tableau A2. Comparaison des effectifs hommes/femmes en fonction de la méthode du résumé	66
Tableau A3. Comparaison des effectifs hommes/femmes en fonction du sous-type de méthode déduite par le résumé	67
Tableau A4. Comparaison des mentions obtenue des thèses incluses « soins de santé primaires » en fonction du sexe.	68
Tableau A5. Comparaison des thèses selon la mention en fonction de la sous-méthode déduite par le résumé	69
Tableau A6. Description statistique du nombre de pages des thèses, relevées lors du tri des résumés, en fonction de la mention.	70
Tableau A7. Répartition des moyennes de nombre de pages en fonction de l'année et de la mention.	70
Tableau A8. Paramètres de formes des thèses issues de la lecture du résumé, comparés par année universitaire.	71
Tableau A9. Paramètres de formes des thèses issues de la lecture du résumé, comparés en fonction de la sous-méthode.	72
Tableau A10. Source des données de l'étude en fonction de la typologie	73
Tableau A11. Répartition des directeurs de thèses en fonction du nombre de thèses dirigées/co-dirigées. Effectif et proportion en fonction de la catégorie du directeur de thèse.	74
Tableau A12. Comparaison des méthodes et résultats de cette thèse avec les travaux antérieurs comparables de M.LABERNÈDE et A. de CUSSY-BOUVET.	75
Tableau A13. Comparaison des résultats de sous-typage des méthodes de cette thèse avec les résultats de M. LABERNÈDE.	76
Tableau A14. Synthèse des variables disponibles dans la feuille de travail	77

TABLE DES ILLUSTRATIONS ET SCHÉMAS

Table des illustrations	Page
Illustration M1 : Critères de choix de type d'étude	15
Illustration A1 : Exemple d'une fiche SUDOC de thèse à résumé et texte intégral papiers	59
Illustration A2 : Exemple d'une fiche SUDOC de thèse à résumé numérique et texte intégral papier	60
Illustration A3 : Exemple d'une fiche SUDOC de thèse à résumé et texte intégral numériques	61

Table des schémas	Page
Schéma R1. Flow-Chart de l'étude	22
Schéma A1. Flow-Chart du tri des titres de thèses, déterminant le rapport aux soins de santé primaires.	63
Schéma A2. Flow-Chart du tri des méthodes selon le titre des thèses incluses « soins de santé primaires ».	63
Schéma A3. Flow-Chart du tri des méthodes selon les résumées des thèses incluses « soins de santé primaires ».	64
Schéma A4. Flow-Chart du tri des types de méthodes quantitatives selon les résumées des thèses de méthode quantitative ou mixte-quantitatives et qualitatives.	64
Schéma A5. Flow-Chart du tri des thèses de méthode quantitative ou mixte supposées comporter une analyses statistique, et détail des critères relevés.	65

TABLE DES REPRÉSENTATIONS GRAPHIQUES

Table des représentations graphiques	Page
Graphique R1. Représentation graphique des effectifs annuels comparés des thèses quantitative/mixtes, qualitatives, et méthodes autres/obstacles techniques.	28

LISTE DES ABRÉVIATIONS

Liste des abréviations

APOGÉE	Application pour l'organisation et la gestion des enseignements et des étudiants
CHU	Centre Hospitalier Universitaire
DES	Diplôme d'études supérieures
DMG	Département de Médecine Générale (de Bordeaux)
DUMAS	Dépot Universitaire des Mémoires Après Soutenance
IMRaD	Anglais : Introduction, Methods, Results and Discussion (Introduction, Matériel et Méthodes, Résultats et Discussion)
[IC]	Intervalle de Confiance
MeSH	Anglais : Medical Subject Headings (en-têtes de sujets médicaux)
NSN	Nombre de Sujets Nécessaires
PDF	Anglais : Portable Document File. Format de fichier informatique, équivalent numérique de documents papier, sous licence libre d'Adobe®.
SUDOC	Système Universitaire de DOCumentation
WONCA	World Organization of National Colleges, Academies and Academic Associations of general practitioners/family physicians Organisation Mondiale des Médecins Généralistes

ANNEXES

Annexe 1. Algorithmes utilisés	56
Annexe 2. Aspects de SUDOC	59
Annexe 3. Classification des méthodes des thèses selon les critères de Paul FRAPPÉ	62
Annexe 4. Diagrammes de flux théoriques de saisie des données.	63
Annexe 5. Résultats complémentaires	66
Annexe 6. Comparaison avec les travaux de Maylis LABERNÈDE et Alix De CUSSY-BOUVET	75

ANNEXE 1. ALGORITHMES UTILISÉS

1.1 Algorithme de déduction comparant l'évaluation des titres des thèses incluses « soins de santé primaires » par mots-clés et l'évaluation des titres par relecture systématique

```
SI($Q125="non-inclus";"non-inclus";  
SI(ET($U125=1;$X125=1);"inchangé, quantitatif 1 -> 1";  
SI(ET($U125=2;$X125=2);"inchangé, douteux 2 -> 2";  
SI(ET($U125=3;$X125=3);"inchangé, qualitatif 3 -> 3";  
SI(ET($U125=4;$X125=4);"inchangé, indéterminable 4 -> 4";  
SI(ET($U125=1;$X125=2);"initialement quantitatif, changé pour douteux 1 -> 2";  
SI(ET($U125=1;$X125=3);"initialement quantitatif, changé pour qualitatif 1 -> 3";  
SI(ET($U125=1;$X125=4);"initialement quantitatif, changé pour quantitatif 1 -> 4";  
SI(ET($U125=2;$X125=1);"initialement douteux, changé pour quantitatif 2 -> 1";  
SI(ET($U125=2;$X125=3);"initialement douteux, changé pour qualitatif 2 -> 3";  
SI(ET($U125=2;$X125=4);"initialement douteux, changé pour indéterminable 2 -> 4";  
SI(ET($U125=3;$X125=1);"initialement qualitatif, changé pour quantitatif 3 -> 1";  
SI(ET($U125=3;$X125=2);"initialement qualitatif, changé pour douteux 3 -> 2";  
SI(ET($U125=3;$X125=4);"initialement qualitatif, changé pour indéterminable 3 -> 4";  
SI(ET($U125=4;$X125=1);"initialement indéterminable, changé pour quantitatif 4 -> 1";  
SI(ET($U125=4;$X125=2);"initialement indéterminable, changé pour douteux 4 -> 2";  
SI(ET($U125=4;$X125=3);"initialement indéterminable, changé pour qualitatif 4 -> 3";  
SI(ET($U125=1;$X125=5);"initialement qualitatif, changé pour Quantitatif ET Qualitatif 1 -> 5";  
SI(ET($U125=2;$X125=5);"initialement douteux, changé pour Quantitatif ET Qualitatif 2 -> 5";  
SI(ET($U125=3;$X125=5);"initialement qualitatif, changé pour Quantitatif ET Qualitatif 3 -> 5";  
SI(ET($U125=4;$X125=5);"initialement indéterminable, changé pour Quantitatif ET Qualitatif 4 ->  
5";"impossible""))))))))))))))))
```

Où :

125 correspondait à la ligne de tableur dont était issu cet exemple.

\$Q contenait le paramètre booléen d'inclusion : « inclus » ; « non-inclus ».

\$U contenait le paramètre discrétisé de l'évaluation par mots-clés : 1 ; 2 ; 3 ; 4 ; où 1= « Quantitatif » ; 2 =« Douteux » ; 3= « Qualitatif » ; 4= « Indéterminable ».

\$X contenait le paramètre discrétisé de l'évaluation par relecture systématique : 1 ; 2 ; 3 ; 4 ; 5 ; où 1= « Quantitatif » ; 2 =« Douteux » ; 3= « Qualitatif » ; 4= « Indéterminable » ; 5= « Quantitative et Qualitative ».

Les conditions étaient incrémentées les une dans les autres, ce qui explique les nombreuses parenthèses en fin d'algorithme.

Le dernier argument SINON « impossible » avait un rôle de recherche de données manquantes, et son absence assurait que toutes les données étaient bien rentrées dans le tableur.

1.2 Algorithme de vérification du lien pérenne issu de SUDOC

```
SI(Q91="non-inclus";"non-inclus";  
SI(AE91=1;  
SI(CHERCHE("TRM";AF91)=45;"Cohérent";"risque erreur »)))
```

Où :

91 correspondait à la ligne de tableur dont était issu cet exemple.

\$Q contenait le paramètre booléen d'inclusion : « inclus » ; « non-inclus ».

\$AE contenait le paramètre booléen de disponibilité en ligne du résumé : 1 ; 2 ; où 1= « Oui »; 2= « Non ».

\$AF contenait le lien SUDOC supposé pérenne.

Une structure normale attendue pour un lien pérenne était :

<http://www.sudoc.abes.fr/DB=2.1/SRCH?IKT=12&TRM=XXXXXXXXXX>, où XXXXXXXXXXXX était un nombre à 9 chiffres, lié de manière unique à chaque fiche SUDOC.

Par conséquent, la recherche de la chaîne TRM en position 45 permettait de vérifier que le lien récupéré était bien pérenne.

<http://www.sudoc.abes.fr/DB=2.1/SRCH?IKT=12&TRM=XXXXXXXXXX>

Les caractères en rouge sont au nombre de 44, par conséquent TRM est bien en position 45, argument donné à la fonction CHERCHE.

1.3. Algorithme de génération du lien Babord+ d'une thèse à partir du nom et du prénom du thésard

```
SI(ET(Q107="inclus";AE107=2;OU(AB107="Nécessaire car quantitatif";AB107="Nécessaire pour déterminer type d'étude"));  
CONCATENER("https://babordplus.u-bordeaux.fr/resultat.php?  
type_rech=rs&index[]=fulltext&bool[]=&value[]=";SUBSTITUE(H107;"  
";"+");"&spec_expand=0&spec_expand=1");"NR")
```

Où :

107 correspondait à la ligne de tableur dont était issu cet exemple.

\$Q contenait le paramètre booléen d'inclusion : « inclus » ; « non-inclus ».

\$AE contenait le paramètre booléen de disponibilité en ligne du résumé : 1 ; 2 ; où 1= « Oui » ; 2= « Non ».

\$H contenait les noms et prénoms du thésard, séparés d'un espace.

\$AB contenait une déduction de l'appel de lecture du résumé. L'algorithme de la colonne \$AB était :
SI(Q107="inclus";SI(OU(X107=1;X107=3);SI(X107=1;"Nécessaire car quantitatif";"non-nécessaire car quantitatif");"Nécessaire pour déterminer type d'étude");"Non-inclus")

Où \$X contenait le paramètre discrétisé de l'évaluation par relecture systématique : 1 ; 2 ; 3 ; 4 ; 5 ; où 1= « Quantitatif » ; 2= « Douteux » ; 3= « Qualitatif » ; 4= « Indéterminable » ; 5= « Quantitative et Qualitative ».

La structure attendue d'un lien babord+ était :

```
https://babordplus.u-bordeaux.fr/resultat.php?  
type_rech=rs&index[]=fulltext&bool[]=&value[]=<nom>+<prénom>&spec_expand=0&spec_expand=1
```

La fonction CONCATENER assemblait 3 segments :

- https://babordplus.u-bordeaux.fr/resultat.php?type_rech=rs&index[]=fulltext&bool[]=&value[]=
- SUBSTITUE(H107;" « ; »+ »), qui renvoyait une chaîne composée des noms et prénoms reliés par des « + ».
- &spec_expand=0&spec_expand=1 »

Ainsi, les liens avaient une forte probabilité d'être fonctionnels, dans la limite des certains noms qui n'étaient pas cohérents entre les données d'APOGEE et celles de Babord+, mais c'était marginal, il suffisait de refaire une recherche dans la fenêtre de bâbord plus indiquant que la page n'existait pas (code d'erreur 404).

ANNEXE 2. ASPECTS DE SUDOC

2.1 présentation type sur SUDOC d'une thèse avec résumé papier et texte intégral papier

rechercher (et) Tous les mots trier par pertinence Recherche

Liste des résultats | **Notice détaillée** | **Où trouver ce document ?** Copyright © 2017 ABES / OCLC

rechercher (et) (PPN) | 1 résultat(s)

Prêt Entre Bibliothèques
Prêt
Photocopie

Services
Mon panier / Export

Identifiant pérenne de la notice :

Titre :

Alphabet du titre :

Auteur(s) :

Date(s) :

Langue(s) :

Pays :

Editeur(s) :

Description : 1 vol. (142 f.) ; 30 cm

Num. national de thèse :

Thèse :

Sujets :

Origine de la notice :

Lien(s) externe(s)
Worldcat :

Google

1

Illustration A1. Exemple d'une fiche SUDOC de thèse à résumé et texte intégral papiers

ANNEXE 3. CLASSIFICATION DES MÉTHODES DES THÈSES SELON LES CRITÈRES DE PAUL FRAPPÉ

- 1 Quantitative
 - Epidémiologique
 - 1.1 étude de cas / Série de cas
 - 1.2 étude écologique
 - 1.3 étude transversale
 - 1.4 Cas-témoins
 - 1.5 Cohorte
 - Interventionnelle
 - 1.6 Essais cliniques
- 2 Évaluation de test diagnostique
 - 2.0 Étude de test diagnostique
- 3 Médico-économique
 - 3.1 Minimisation des coûts
 - 3.2 Coût-efficacité
 - 3.3 Coût-utilité
 - 3.4 Coût-bénéfice
- 4 Semi-quantitatif
 - 4.1 Processus-Delphi
 - 4.2 Groupe nominal
- 5 Évaluation de la qualité des soins
 - 5.1 Enregistrement systématique
 - 5.2 Enquête de pratiques
 - 5.3 Enquête analytique de pratiques
 - 5.4 Étude ici-ailleurs quasi-expérimentale
 - 5.5 Audit clinique
 - 5.6 Essai en population
 - 5.7 Étude observationnelle ici-ailleurs
 - 5.8 Étude observationnelle avant-après
- 6 Qualitative
 - 6.1 Entretien individuel structuré
 - 6.2 Observation directe
 - 6.3 Observation indirecte
 - 6.4 Entretien individuel semi-structuré
 - 6.5 Entretien individuel non structuré
 - 6.6 Focus group
- 7 Recherche secondaire
 - 7.1 Revue systématique
 - 7.2 Méta-analyse quantitative
 - 7.3 Méta-analyse qualitative

ANNEXE 4. DIAGRAMMES DE FLUX THÉORIQUES DE SAISIE DES DONNÉES.

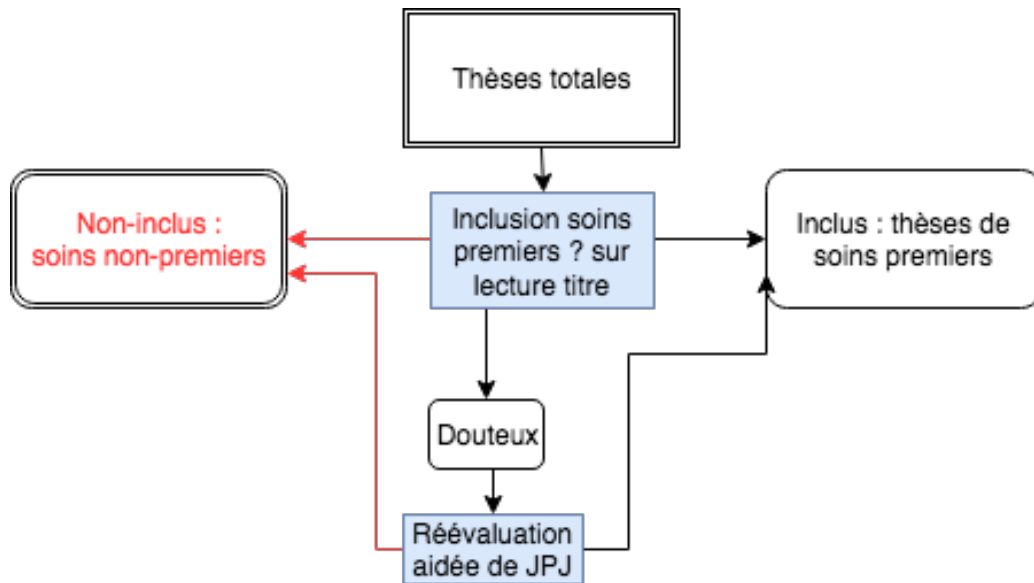


Schéma A1. Flow-Chart du tri des titres de thèses, déterminant le rapport aux soins de santé

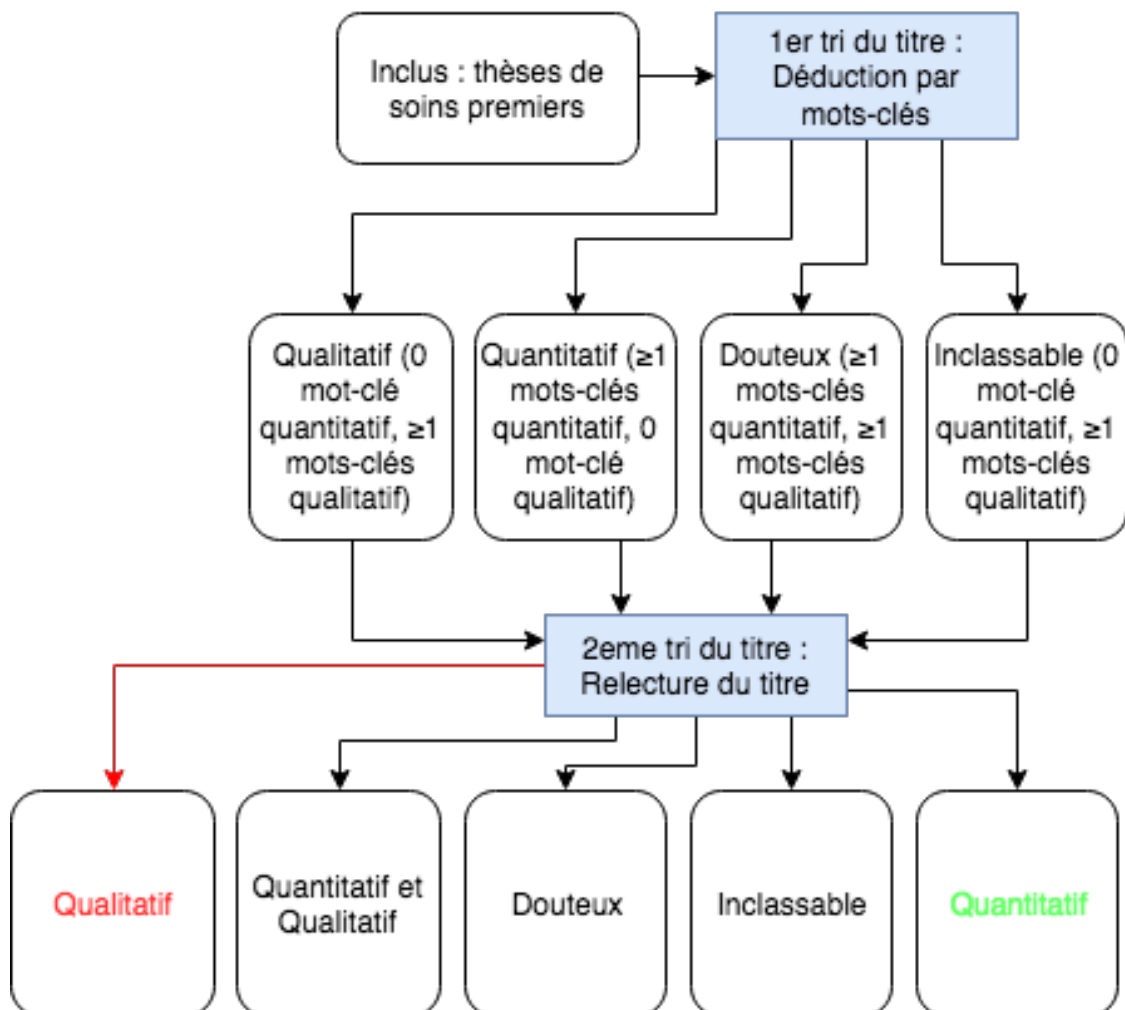


Schéma A2. Flow-Chart du tri des méthodes selon le titre des thèses incluses « soins de santé »

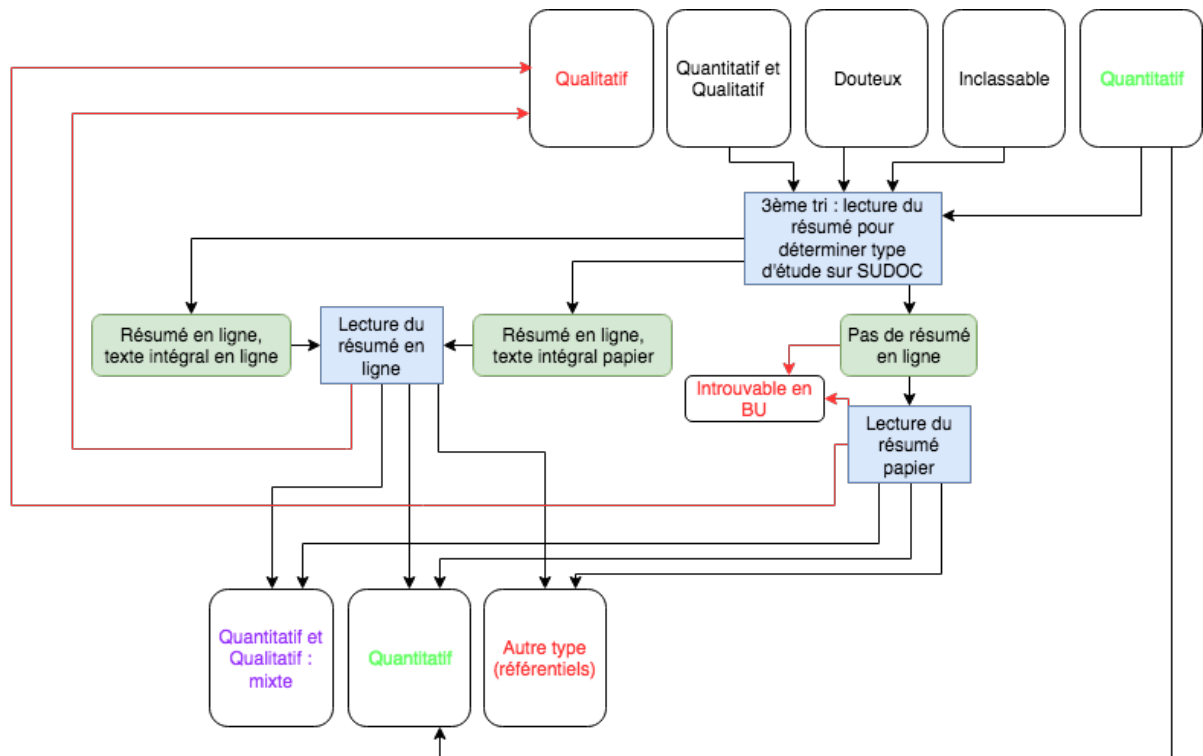


Schéma A3. Flow-Chart du tri des méthodes selon les résumés des thèses incluses « soins de santé

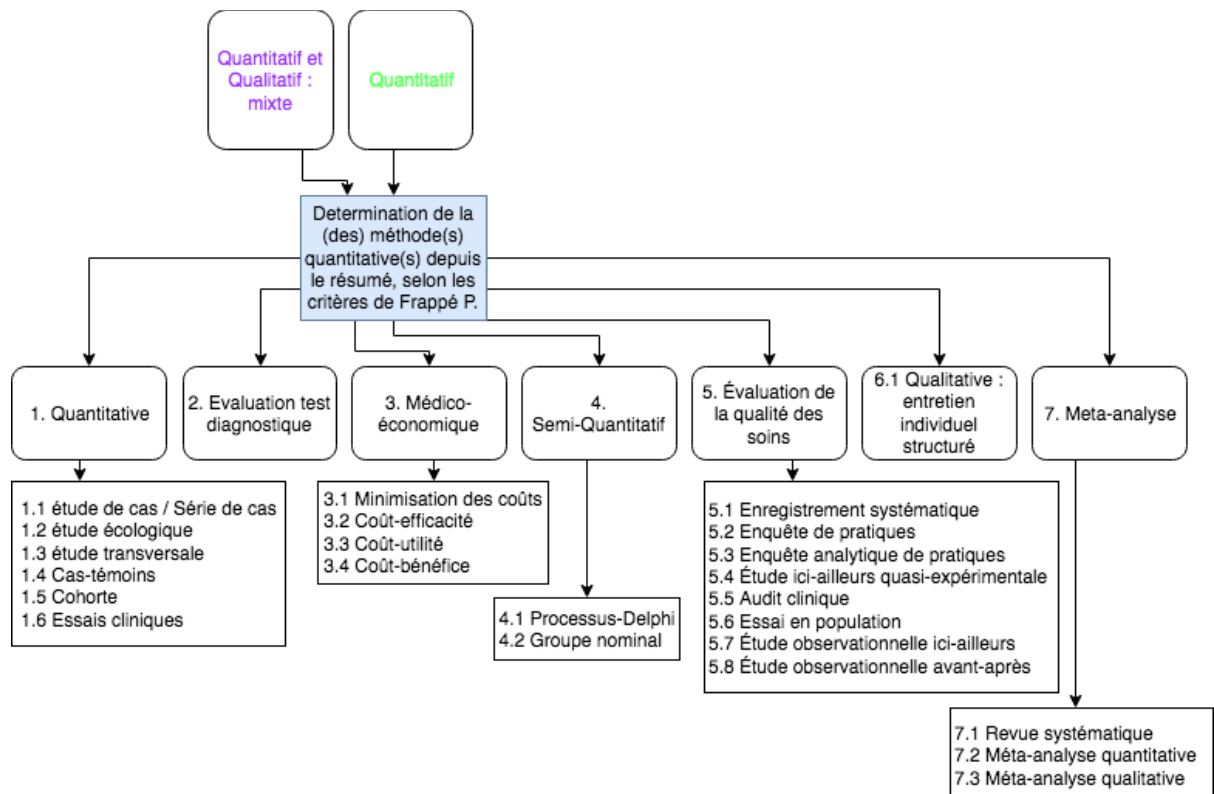


Schéma A4. Flow-Chart du tri des types de méthodes quantitatives selon le résumé des thèses de méthode quantitatives ou mixtes-quantitatives et qualitatives

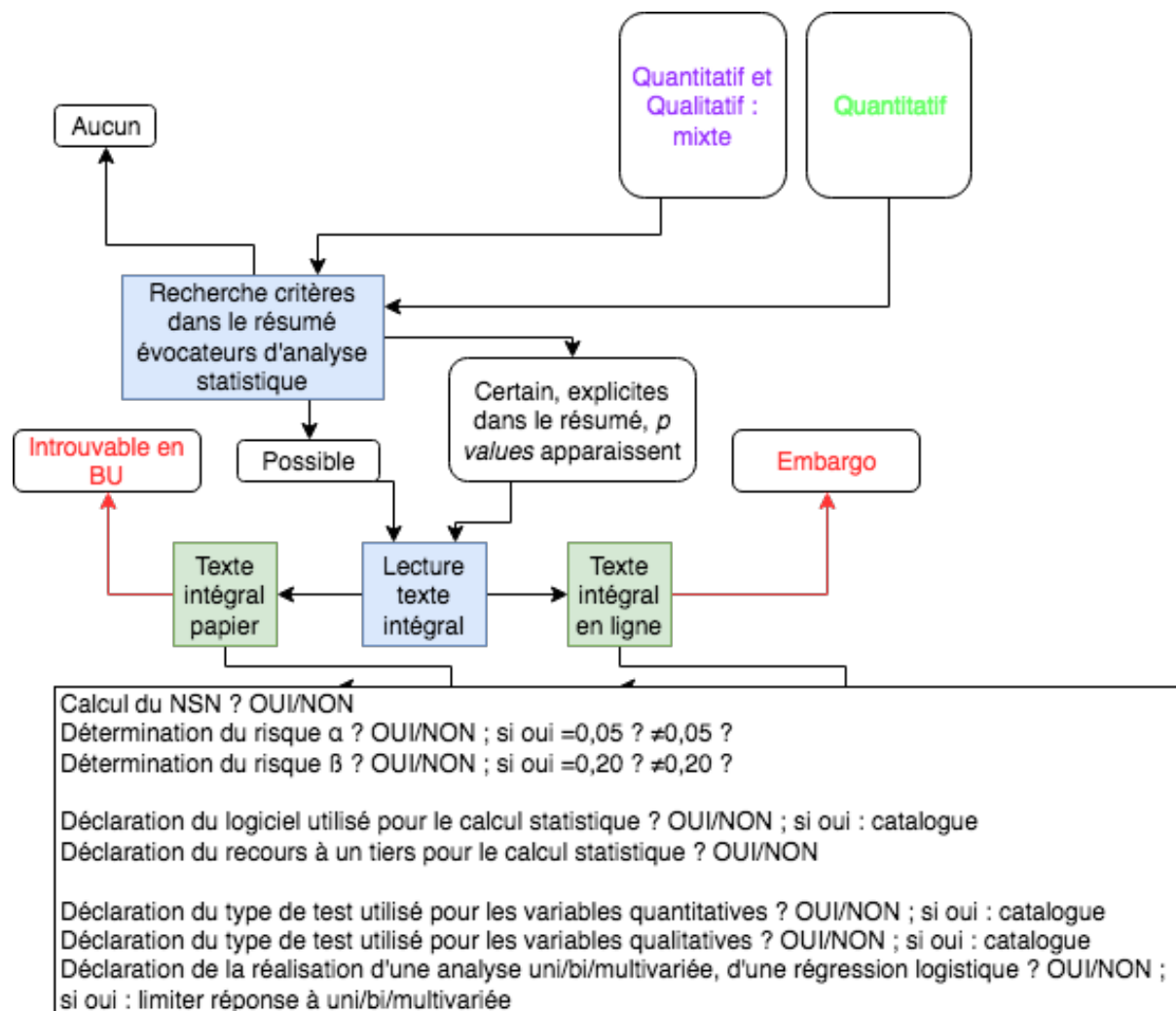


Schéma A6. Flow-Chart du tri des thèses de méthode quantitative ou mixte supposées comporter une analyses statistique, et détail des critères relevés

ANNEXE 5. RÉSULTATS COMPLÉMENTAIRES

5.1 Concernant le sexe de l'auteur(e)

Tableau A1. Comparaison des effectifs d'hommes et de femmes, dans les thèses incluses et dans les thèses totales, calcul des sex-ratios.

Promotion (document d'origine)	Hommes		Femmes		Sex-ratio (hommes/femmes)	
	Hommes inclus	Hommes totaux	Femmes incluses	Femmes totaux	Sex-ratio inclus	Sex-ratio total
2006-2007	11 (28,2 %)	39	30 (50,0 %)	60	0,37	0,65
2007-2008	21 (41,2 %)	51	25 (51,0 %)	49	0,84	1,04
2008-2009	13 (36,1 %)	36	24 (45,3 %)	53	0,54	0,68
2009-2010	17 (50,0 %)	34	29 (54,7 %)	53	0,59	0,64
2010-2011	21 (50,0 %)	42	28 (58,3 %)	48	0,75	0,88
2011-2012	23 (37,7 %)	61	37 (47,4 %)	78	0,62	0,78
2012-2013	31 (47,7 %)	65	50 (57,5 %)	87	0,62	0,75
2013-2014	35 (50,7 %)	69	62 (60,8 %)	102	0,56	0,68
2014-2015	32 (52,5 %)	61	62 (62,0 %)	100	0,52	0,61
2015-2016	22 (66,7 %)	33	33 (56,9 %)	58	0,67	0,57
total	226 (46,0 %)	491	380 (55,2 %)	688	0,59	0,71

Le test t de Student comparant les effectifs totaux d'hommes et de femme, par année, retrouvait : $p=2,41 \times 10^{-2}$. Il existait donc une différence statistiquement significative.

Le test t de Student comparant les effectifs inclus d'hommes et de femmes, par année, retrouvait : $p=9,20 \times 10^{-3}$. Il existait donc une différence statistiquement significative.

Tableau A2. Comparaison des effectifs hommes/femmes en fonction de la méthode du résumé

Résultat tri des résumés	Effectifs hommes	Effectifs femmes	Effectifs tous genres
total quantitatif	136 (22,4 %)	239 (39,4 %)	375 (61,9 %)
total qualitatif	71 (11,7 %)	123 (20,3 %)	194 (32,0 %)
total mixte	12 (2,0 %)	16 (2,6 %)	28 (4,6 %)
total autre méthode	1 (0,2 %)	1 (0,2 %)	2 (0,3 %)
total inclus	226 (37,3 %)	380 (62,7 %)	606 (100 %)

Tableau A3. Comparaison des effectifs hommes/femmes en fonction du sous-type de méthode déduite par le résumé

Sous-type de méthodes	Effectifs hommes selon la sous-méthode	Effectif femmes selon la sous-méthode	Effectif tous genres selon la sous-méthode
1.1 : Epidémiologique, étude de cas / Série de cas	5 (3,4 %)	7 (2,8 %)	12 (3,0 %)
1.3 : Epidémiologique, étude transversale	37 (25,0 %)	71 (27,8 %)	108 (26,8 %)
1.4 : Epidémiologique, Cas/témoins	4 (2,7 %)	6 (2,4 %)	10 (2,5 %)
1.5 : Epidémiologique, Cohorte	25 (16,9 %)	49 (19,2 %)	74 (18,4 %)
1.6 : Interventionnel, Essai clinique	7 (4,7 %)	9 (3,5 %)	16 (4,0 %)
2.0 : Evaluation test diagnostic	6 (4,1 %)	14 (5,5 %)	20 (5,0 %)
3.1 : Médico-économique, minimalisation des coûts	0 (0 %)	1 (0,4 %)	1 (0,3 %)
3.3 : Médico-économique, coût-utilité	2 (1,4 %)	1 (0,4 %)	3 (0,7 %)
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	5 (3,4 %)	5 (2,0 %)	10 (2,5 %)
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	8 (5,4 %)	24 (9,4 %)	32 (7,9 %)
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	4 (2,7 %)	6 (2,4 %)	10 (2,5 %)
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	1 (0,7 %)	4 (1,6 %)	5 (1,2 %)
5.5 : Evaluation de la Qualité des Soins, audit clinique	8 (5,4 %)	8 (3,1 %)	16 (4,0 %)
5.6 : Evaluation de la Qualité des Soins, essai en population	7 (4,7 %)	5 (2,0 %)	12 (3,0 %)
5.7 : Evaluation de la Qualité des Soins, etude observationnelle ici-ailleurs	4 (2,7 %)	3 (1,2 %)	7 (1,7 %)
5.8 : Evaluation de la Qualité des Soins, etude observationnelle avant-après	10 (6,8 %)	17 (6,7 %)	27 (6,7 %)
6.1 : Qualitative, entretien individuel structuré	11 (7,4 %)	16 (6,3 %)	27 (6,7 %)
7.1 : Recherche secondaire, revue systématique	4 (2,7 %)	8 (3,1 %)	12 (3,0 %)
7.2 : Recherche secondaire, méta-analyse quantitative	0 (0 %)	1 (0,4 %)	1 (0,3 %)
Total :	148 (100 %)	255 (100 %)	403 (100 %)

5.2 Concernant les mentions obtenues

Tableau A4. Comparaison des mentions obtenue des thèses incluses
« soins premiers » en fonction du sexe.

Mention	Effectifs hommes	Effectifs femmes	Effectifs tous genres
admis	3 (1,3 %)	1 (0,3 %)	4 (0,7 %)
admis honorable	61 (27,0 %)	63 (16,3 %)	124 (20,5 %)
admis très honorable	121 (53,5 %)	248 (65,3 %)	369 (60,9 %)
admis très honorable avec félicitations	41 (18,1 %)	68 (17,9 %)	109 (18,0 %)
total	226 (100 %)	380 (100 %)	606 (100 %)

Le test t de Student comparant les effectifs inclus d'hommes et de femme, selon la mention, retrouvait : $p=0,54$. Il n'existait donc pas de différence statistiquement significative.

Tableau A5. Comparaison des thèses selon la mention en fonction de la sous-méthode déduite par le résumé

Sous-type de méthode	total mentions admis selon méthode	total mentions admis honorables selon méthode	total mentions admis très honorables selon méthode	total mentions admis très honorables avec félicitations selon méthode
1.1 : Epidémiologique, étude de cas / Série de cas	0 (0 %)	1 (1,3 %)	9 (3,7 %)	2 (2,6 %)
1.3 : Epidémiologique, étude transversale	1 (25,0 %)	20 (25,0 %)	67 (27,6 %)	20 (26,0 %)
1.4 : Epidémiologique, Cas/témoins	0 (0 %)	1 (1,3 %)	7 (2,9 %)	2 (2,6 %)
1.5 : Epidémiologique, Cohorte	1 (25,0 %)	19 (23,8 %)	40 (16,5 %)	14 (18,2 %)
1.6 : Interventionnel, Essai clinique	1 (25,0 %)	4 (5,0 %)	7 (2,9 %)	4 (5,2 %)
2.0 : Evaluation test diagnostic	0 (0 %)	4 (5,0 %)	12 (4,9 %)	4 (5,2 %)
3.1 : Médico-économique, minimalisation des coûts	0 (0 %)	0 (0 %)	1 (0,4 %)	0 (0 %)
3.3 : Médico-économique, coût-utilité	0 (0 %)	2 (2,5 %)	1 (0,4 %)	0 (0 %)
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	0 (0 %)	0 (0 %)	7 (2,9 %)	3 (3,9 %)
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	0 (0 %)	7 (8,8 %)	22 (9,1 %)	3 (3,9 %)
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	0 (0 %)	2 (2,5 %)	7 (2,9 %)	1 (1,3 %)
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	0 (0 %)	0 (0 %)	3 (1,2 %)	2 (2,6 %)
5.5 : Evaluation de la Qualité des Soins, audit clinique	0 (0 %)	4 (5,0 %)	10 (4,1 %)	2 (2,6 %)
5.6 : Evaluation de la Qualité des Soins, essai en population	0 (0 %)	0 (0 %)	10 (4,1 %)	3 (3,9 %)
5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs	0 (0 %)	0 (0 %)	4 (1,7 %)	3 (3,9 %)
5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après	1 (25,0 %)	4 (5,0 %)	11 (4,5 %)	11 (14,3 %)
6.1 : Qualitative, entretien individuel structuré	0 (0 %)	8 (10,0 %)	19 (7,8 %)	0 (0 %)
7.1 : Recherche secondaire, revue systématique	0 (0 %)	3 (3,8 %)	6 (2,5 %)	3 (3,9 %)
7.2 : Recherche secondaire, méta-analyse quantitative	0 (0 %)	1 (1,3 %)	0 (0 %)	0 (0 %)
Total :	4 (100 %)	80 (100 %)	243 (100 %)	77 (100 %)

Les sous-types de méthode dont l'effectif était nul n'ont pas été reportés dans le tableau.

5.3 Concernant le nombre de pages des thèses

Tableau A6. Description statistique du nombre de pages des thèses, relevées lors du tri des résumés, en fonction de la mention.

Paramètres Statistiques	Mentions			
	Admis	Admis Honorable	Admis Très Honorable	Admis Très Honorable Avec Félicitations
Moyenne	60,8	97,7	109,3	126,9
Écart-Type	14,9	33,2	44,0	44,9
Valeur Minimale	43,0	44,0	29,0	46,0
Quartile 25%	52,0	74,0	81,0	95,0
Mediane	61,5	91,0	101,0	122,0
Quartile 75%	70,3	112,5	125,5	149,0
Valeur Maximale	77,0	203,0	377,0	347,0
Mode	absent	74,0	98,0	85,0
Nombre de thèses concernées	4	112	327	97

Tableau A7. Répartition des moyennes de nombre de pages en fonction de l'année et de la mention.

Promotion (document d'origine)	Moyenne du nombre de pages des thèse	Moyenne du nombre de pages des thèse selon la mention			
		Admis	Admis Honorable	Admis très Honorable	Admis très Honorable avec Félicitations
2006-2007	133,1	<i>aucune thèse</i>	111,0	139,9	116,3
2007-2008	113,1	<i>aucune thèse</i>	87,7	111,3	147,6
2008-2009	110,8	<i>aucune thèse</i>	81,8	114,3	121,0
2009-2010	114,7	55,0	104,8	121,2	114,0
2010-2011	116,1	<i>aucune thèse</i>	105,0	103,2	170,1
2011-2012	112,4	77,0	109,7	106,8	129,2
2012-2013	111,1	<i>aucune thèse</i>	102,6	109,8	128,0
2013-2014	106,4	<i>aucune thèse</i>	96,1	102,4	122,6
2014-2015	95,5	68,0	89,4	94,7	108,3
2015-2016	101,4	43,0	78,8	106,4	114,7
Moyenne sur 10 ans	109,7	60,8	96,7	111,0	127,2
Effectif total	606	4	112	327	97

5.4 Concernant les critères de formes du résumé : structure IMRaD ; résumé français ou bilingue ; disponibilité en ligne

Tableau A8. Paramètres de formes des thèses issues de la lecture du résumé,
comparés par année universitaire.

Promotion (document d'origine)	Résumés Bilingues	Structure IMRaD	Structure IMRaD très explicite	Résumé en ligne	Texte intégral en ligne	Triples positifs*	total année
2006-2007	5 (12,2 %)	8 (19,5 %)	7 (17,1 %)	9 (22,0 %)	0 (0 %)	0 (0 %)	41
2007-2008	2 (4,4 %)	9 (19,6 %)	6 (13,0 %)	32 (69,6 %)	0 (0 %)	0 (0 %)	46
2008-2009	3 (8,1 %)	12 (32,4 %)	12 (32,4 %)	36 (97,3 %)	0 (0 %)	0 (0 %)	37
2009-2010	7 (15,2 %)	18 (39,1 %)	18 (39,1 %)	46 (100 %)	0 (0 %)	0 (0 %)	46
2010-2011	13 (26,5 %)	9 (18,4 %)	9 (18,4 %)	49 (100 %)	1 (2,0 %)	0 (0 %)	49
2011-2012	10 (16,7 %)	25 (41,7 %)	17 (28,3 %)	59 (98,3 %)	0 (0 %)	0 (0 %)	60
2012-2013	30 (37,0 %)	35 (43,2 %)	2 (2,5 %)	81 (100 %)	20 (24,7 %)	9 (11,1 %)	81
2013-2014	32 (33,0 %)	47 (48,5 %)	11 (11,3 %)	97 (100 %)	68 (70,1 %)	41 (42,3 %)	97
2014-2015	60 (63,8 %)	55 (58,5 %)	3 (3,2 %)	88 (93,6 %)	67 (71,3 %)	48 (51,1 %)	94
2015-2016	37 (67,3 %)	38 (69,1 %)	0 (0 %)	53 (96,4 %)	42 (76,4 %)	37 (67,3 %)	55
Total	199 (32,8 %)	256 (42,2 %)	85 (14,0 %)	550 (90,8 %)	198 (32,7 %)	135 (22,3 %)	606

Triples positifs : association d'une structure IMRaD ou IMRaD très explicite, d'un texte intégral disponible en ligne, et d'un résumé bilingue.

Cette variable, reflet global de critères de forme idéaux du résumé et d'accessibilité en ligne de la thèse, est nulle avant 2012, puis présente une pente forte, pour culminer à 67% en 2016.

Tableau A9. Paramètres de formes des thèses issues de la lecture du résumé, comparés en fonction de la sous-méthode.

Sous-type de méthode	Nombre de thèses bilingues	Nombre de résumés IMRaD	Nombre de résumés IMRaD très explicites	Nombre de textes intégraux en ligne	Total selon sous-type
1.1 : Epidémiologique, étude de cas / Série de cas	4 (33,3 %)	4 (33,4 %)	2 (16,7 %)	4 (33,3 %)	12
1.3 : Epidémiologique, étude transversale	53 (49,1 %)	63 (58,4 %)	17 (15,7 %)	52 (48,2 %)	108
1.4 : Epidémiologique, Cas/témoins	3 (30,0 %)	5 (50,0 %)	3 (30,0 %)	3 (30,0 %)	10
1.5 : Epidémiologique, Cohorte	25 (33,8 %)	46 (62,2 %)	18 (24,3 %)	28 (37,8 %)	74
1.6 : Interventionnel, Essai clinique	5 (31,3 %)	7 (43,8 %)	2 (12,5 %)	7 (43,8 %)	16
2.0 : Evaluation test diagnostic	8 (40,0 %)	9 (45,0 %)	5 (25,0 %)	6 (30,0 %)	20
3.1 : Médico-économique, minimalisation des coûts	0 (0 %)	1 (100 %)	1 (100 %)	0 (0 %)	1
3.3 : Médico-économique, coût-utilité	1 (33,3 %)	1 (33,3 %)	1 (33,3 %)	1 (33,3 %)	3
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	2 (20,0 %)	7 (70,0 %)	3 (30,0 %)	2 (20,0 %)	10
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	13 (40,6 %)	11 (34,4 %)	4 (12,5 %)	14 (43,8 %)	32
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	2 (20,0 %)	2 (20,0 %)	1 (10,0 %)	3 (30,0 %)	10
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	1 (20,0 %)	2 (40,0 %)	2 (40,0 %)	2 (40,0 %)	5
5.5 : Evaluation de la Qualité des Soins, audit clinique	5 (31,3 %)	6 (37,5 %)	2 (12,5 %)	5 (31,3 %)	16
5.6 : Evaluation de la Qualité des Soins, essai en population	7 (58,3 %)	4 (33,3 %)	2 (16,7 %)	3 (25,0 %)	12
5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs	4 (57,1 %)	4 (57,1 %)	1 (14,3 %)	4 (57,1 %)	7
5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après	10 (37,0 %)	15 (55,6 %)	2 (7,4 %)	11 (40,7 %)	27
6.1 : Qualitative, entretien individuel structuré	13 (48,2 %)	18 (66,7 %)	5 (18,5 %)	8 (29,6 %)	27
7.1 : Recherche secondaire, revue systématique	2 (16,7 %)	2 (16,7 %)	1 (8,3 %)	4 (33,3 %)	12
7.2 : Recherche secondaire, méta-analyse quantitative	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	1
total	158 (39,2%)	207 (51,4%)	72 (17,87 %)	157 (39,0 %)	403

Tableau A10. Source des données de l'étude en fonction de la typologie

Sous-type d'étude selon critères de Paul FRAPPÉ	Différentes sources de données des études									Total	
	0 non-précisé	1 bases de données médicales hospitalières	2 bases de données médicales non-hospitalières	3 bases de données de santé publique (isped, inserm, CFAM, autres caisses primaires)	4 dossier médical du médecin traitant/du spécialiste	5 questionnaire	6 multiples, 1+/-2+/-3+/-4, sans questionnaire	7 multiples, 1+/-2+/-3+/-4, avec questionnaire	8 données issue d'une autre étude		9 Autres bases de données, Faculté ou en rapport avec les thèses
1.1 : Epidémiologique, étude de cas / Série de cas	0 (0,00 %)	3 (25,00 %)	2 (16,67 %)	0 (0,00 %)	1 (8,33 %)	1 (8,33 %)	2 (16,67 %)	1 (8,33 %)	0 (0,00 %)	2 (16,67 %)	12
1.3 : Epidémiologique, étude transversale	0 (0,00 %)	40 (37,04 %)	3 (2,78 %)	1 (0,93 %)	8 (7,41 %)	45 (41,67 %)	5 (4,63 %)	2 (1,85 %)	2 (1,85 %)	2 (1,85 %)	108
1.4 : Epidémiologique, Cas/témoins	1 (10,00 %)	4 (40,00 %)	0 (0,00 %)	0 (0,00 %)	2 (20,00 %)	1 (10,00 %)	1 (10,00 %)	1 (10,00 %)	0 (0,00 %)	0 (0,00 %)	10
1.5 : Epidémiologique, Cohorte	1 (1,25 %)	39 (52,70 %)	5 (6,76 %)	0 (0,00 %)	3 (4,05 %)	17 (22,97 %)	2 (2,70 %)	4 (5,41 %)	3 (4,05 %)	0 (0,00 %)	74
1.6 : Interventionnel, Essai clinique	0 (0,00 %)	4 (25,00 %)	4 (25,00 %)	0 (0,00 %)	3 (18,75 %)	3 (18,75 %)	0 (0,00 %)	1 (6,25 %)	1 (6,25 %)	0 (0,00 %)	16
2.0 : Evaluation test diagnostic	0 (0,00 %)	10 (50,00 %)	0 (0,00 %)	0 (0,00 %)	1 (5,00 %)	6 (30,00 %)	0 (0,00 %)	1 (5,00 %)	2 (10,00 %)	0 (0,00 %)	20
3.1 : Médico-économique, minimisation des coûts	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	1 (100,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	1
3.3 : Médico-économique, coût-utilité	0 (0,00 %)	0 (0,00 %)	1 (33,33 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	2 (66,67 %)	0 (0,00 %)	0 (0,00 %)	3
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	0 (0,00 %)	4 (40,00 %)	1 (10,00 %)	2 (20,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	3 (30,00 %)	0 (0,00 %)	0 (0,00 %)	10
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	0 (0,00 %)	1 (3,13 %)	0 (0,00 %)	0 (0,00 %)	2 (6,25 %)	28 (87,50 %)	0 (0,00 %)	1 (3,13 %)	0 (0,00 %)	0 (0,00 %)	32
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	0 (0,00 %)	2 (20,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	7 (70,00 %)	0 (0,00 %)	1 (10,00 %)	0 (0,00 %)	0 (0,00 %)	10
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	0 (0,00 %)	1 (20,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	2 (40,00 %)	0 (0,00 %)	2 (40,00 %)	0 (0,00 %)	0 (0,00 %)	5
5.5 : Evaluation de la Qualité des Soins, audit clinique	0 (0,00 %)	5 (31,25 %)	4 (25,00 %)	0 (0,00 %)	3 (18,75 %)	3 (18,75 %)	0 (0,00 %)	0 (0,00 %)	1 (6,25 %)	0 (0,00 %)	16
5.6 : Evaluation de la Qualité des Soins, essai en population	0 (0,00 %)	0 (0,00 %)	2 (16,67 %)	0 (0,00 %)	0 (0,00 %)	8 (66,67 %)	1 (8,33 %)	0 (0,00 %)	1 (8,33 %)	0 (0,00 %)	12
5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs	0 (0,00 %)	0 (0,00 %)	1 (14,29 %)	0 (0,00 %)	1 (14,29 %)	4 (57,14 %)	0 (0,00 %)	1 (14,29 %)	0 (0,00 %)	0 (0,00 %)	7
5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après	0 (0,00 %)	5 (18,52 %)	4 (14,81 %)	1 (3,70 %)	0 (0,00 %)	11 (40,74 %)	2 (7,41 %)	2 (7,41 %)	2 (7,41 %)	0 (0,00 %)	27
6.1 : Qualitative, entretien individuel structuré	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	22 (81,48 %)	0 (0,00 %)	5 (18,52 %)	0 (0,00 %)	0 (0,00 %)	27
7.1 : Recherche secondaire, revue systématique	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	1 (8,33 %)	0 (0,00 %)	2 (16,67 %)	1 (8,33 %)	1 (8,33 %)	6 (50,00 %)	1 (8,33 %)	12
7.2 : Recherche secondaire, méta-analyse quantitative	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	0 (0,00 %)	1 (100,00 %)	0 (0,00 %)	1
Total :	2 (0,50 %)	118 (29,28 %)	27 (6,70 %)	5 (1,24 %)	25 (6,20 %)	160 (39,70 %)	14 (3,47 %)	28 (6,95 %)	19 (4,71 %)	5 (1,24 %)	403

5.5 Concernant les directeurs de thèses

Le nombre total de directeurs de thèses à méthode quantitative ou mixte différents de 2006 à 2016 ans était de 263.

Tableau A11. Répartition des directeurs de thèses en fonction du nombre de thèses dirigées/co-dirigées. Effectif et proportion en fonction de la catégorie du directeur de

Nombre de thèses dirigées ou co-dirigées par directeur de thèse	Effectif des directeurs de thèses en fonction du nombre de thèses dirigée				
	Enseignant du DMG	Médecin généraliste externe au DMG	Médecin du CHU	Médecin spécialiste	Total directeurs de thèse
1	1	60	56	77	194
2	6	9	12	6	33
3	1	3	10	5	19
4	3	3	3	0	9
5	1	0	0	1	2
6	1	0	0	0	1
7	1	0	0	0	1
8	1	0	1	0	2
11	1	0	0	0	1
14	1	0	0	0	1
Total par catégorie de directeur de thèse	17	75	82	89	263

Une large majorité des directeurs de thèses ont donc dirigé une seule thèse (N=194).

Cependant, aucun test statistique n'a pu être réalisé pour manque d'effectif.

Les directeurs de thèses ayant dirigé le plus grand nombre de thèse sont des enseignants du Département de Médecine Générale.

ANNEXE 6. COMPARAISON AVEC LES TRAVAUX DE MAYLIS LABERNÈDE ET ALIX DE CUSSY-BOUVET

Tableau A12. Comparaison des méthodes et résultats de cette thèse avec les travaux antérieurs comparables de M.LABERNÈDE et A. de CUSSY-BOUVET.

Item comparé	Cette thèse	M. LABERNÈDE	A. de CUSSY-BOUVET
Année soutenance	2017	2016	2016
Période	2006-2016	2007-2014	2004-2014
Type de thèse	Ensembles des thèses du DMG de Bordeaux	Ensemble des thèses du DMG de Bordeaux	Thèses subdivision Océan Indien incluent DESC MU
Origine des données	APOGÉE	APOGÉE	APOGÉE recoupée avec le DUMG
Objectif principal	Typologie des thèses de recherche quantitative	Proportion de thèses considérées comme médecine générale, selon composition du jury.	Décrire thèmes, objets et méthodes des thèses de médecine générale subdivision Océan Indien
Objectifs secondaires	<ul style="list-style-type: none"> - Déterminer la proportion de thèses ayant une méthode de recherche quantitative en soins de santé primaires et évolution sur les 10 dernières années. - Comparaison de cette proportion aux thèses à méthode qualitative. - Déterminer la catégorie du directeur de thèse, et recherche l'existence d'un lien entre la catégorie du directeur de thèse et la méthode statistique. - Détailler les méthodes statistiques : détermination du nombre de sujet nécessaire, calcul du risque α, calcul du risque β. - Déterminer le recours à une aide méthodologique hors du DMG et recherche l'existence d'un lien entre le recours à une aide méthodologique hors du DMG et la méthode statistique. - Déterminer les tests statistiques utilisés. 	<ul style="list-style-type: none"> - Proportion de directeurs de thèses et de présidents de jury médecins généralistes et la proportion de jurés généralistes par jury des thèses soutenues à Bordeaux de 2007 à 2014 - Informations sur les caractéristiques des thèses : type d'étude, thème de recherche, population, problème de santé étudié, accessibilité du résumé et de la thèse électronique, structure du résumé et traduction en anglais. - Évaluer la grille de critères ayant permis de discriminer les thèses de médecine générale dans ce travail après analyse des résultats. 	<ul style="list-style-type: none"> - Nombre de médecins généralistes enseignants, médecins généralistes maîtres de stage des universités et médecins généralistes dans les jurys de thèses, - Nombre de directeurs de thèse étant des médecins généralistes, - Définir le nombre de médecins généralistes parmi les rapporteurs, - Analyser les thèmes et objets des thèses soutenus par les internes inscrits en Diplôme d'Etudes Spécialisées Complémentaires (DESC) de Médecine d'Urgence, - Définir le délai entre le début du troisième cycle et la soutenance de thèse, - Donner les mentions obtenues, - Montrer une éventuelle évolution des critères mentionnés ci-dessus depuis la création du DUMG.
Tri soins de santé primaires	Critères simplifiés en 6 items selon WONCA	Critères WONCA/CISP-2	Critères WONCA/CISP-1
Tri méthodes de recherche	Liste P. Frappé.	Partiellement comparable à liste de P. Frappé, sous-type pris en compte	Qualitatif. Quantitatif, 4 classes. - Descriptif - Analytique - Interventionnel - Revue de la littérature sous-type non-pris en compte.
Effectif thèses totales	1179	925	144
Effectif thèses soins de santé primaires	606 (51,4%)	360 (38,9%) (médecine générale + ambiguës)	48 (33,3%)
Effectifs quantitatif+mixte	403 (66,5%) (375+28)	301 (83,6%) (médecine générale + ambiguës)	141 sur thèses totales (138+3)
Effectif qualitatif	194 (32,0%)	55 (15,3%) (médecine générale + ambiguës)	3 sur thèses totales
Effectif thèse dirigée par un médecin généraliste	173 (28,6%) (pour quantitatif et mixte)	199 (55,3%) (pour thèses soins de santé primaires, médecine générale + ambiguës)	non-recherché

Tableau A13. Comparaison des résultats de sous-typage des méthodes de cette thèse avec les résultats de M. LABERNÈDE.

Sous-type d'étude selon critères de Paul FRAPPÉ	total cette thèse quantitatif + mixte	total M. LABERNÈDE	comparaison, <i>p value</i> test exact de Fisher
1.1 : Epidémiologique, étude de cas / Série de cas	12 (2,3 %)	5 (2,4 %)	0.80
1.3 : Epidémiologique, étude transversale	108 (26,8 %)	65 (31,1 %)	0.42
1.4 : Epidémiologique, Cas/témoins	10 (2,5 %)	15 (7,2 %)	1.03x10 ⁻²
1.5 : Epidémiologique, Cohorte	74 (18,4 %)	ND	
1.6 : Interventionnel, Essai clinique	16 (4,0 %)	ND	
2.0 : Evaluation test diagnostic	20 (5,0 %)	ND	
3.1 : Médico-économique, minimalisation des coûts	1 (0,3 %)	ND	
3.3 : Médico-économique, coût-utilité	3 (0,7 %)	ND	
5.1 : Evaluation de la Qualité des Soins, enregistrement systématique	10 (2,5 %)	ND	
5.2 : Evaluation de la Qualité des Soins, enquête de pratique	32 (7,9 %)	99 (47,4 %)	3.01x10 ⁻¹⁸
5.3 : Evaluation de la Qualité des Soins, enquête analytique de pratiques	10 (2,5 %)	ND	
5.4 : Evaluation de la Qualité des Soins, étude ici-ailleurs quasi expérimentale	5 (1,2 %)	ND	
5.5 : Evaluation de la Qualité des Soins, audit clinique	16 (4,0 %)	ND	
5.6 : Evaluation de la Qualité des Soins, essai en population	12 (3,0 %)	ND	
5.7 : Evaluation de la Qualité des Soins, étude observationnelle ici-ailleurs	7 (1,7 %)	0 (0 %)	0.10
5.8 : Evaluation de la Qualité des Soins, étude observationnelle avant-après	27 (6,7 %)	10 (4,8 %)	0.47
6.1 : Qualitative, entretien individuel structuré	27 (6,7 %)	4 (1,9 %)	1.73x10 ⁻²
7.1 : Recherche secondaire, revue systématique	12 (3,0 %)	11 (5,3 %)	0.19
7.2 : Recherche secondaire, méta-analyse quantitative	1 (0,3 %)	0 (0 %)	1.00
Total :	403 (100 %)	209 (100 %)	

ND : donnée non-disponible car non-concordante avec les classes établies.

Le test exact de Fisher fut privilégié par rapport aux effectifs très faibles de certains résultats.

A noter que la classe « 5.2 : Evaluation de la Qualité des Soins, enquête de pratique » n'est pas strictement concordante entre la thèse de M. LABERNÈDE et le présent travail, nous avons détaillé les Evaluations de la qualité des soins sur davantage de catégories.

Tableau A13. Synthèse des variables disponibles dans la feuille de travail

Variable	Type de thèse concernée	Effectifs disponibles dans le tableau de données	
Nom(s) et Prénom(s) thésard	toutes	1179	
Sexe	toutes	1179	
Mention obtenue	toutes	1179	
Titre de la thèse	toutes	1179	
Recherche du résumé sur sudoc	études triées quantitativement à partir du titre	271	545
	études dont méthode à déterminer par le résumé	274	
Lien sudoc	liens pérennes trouvés lors de la recherche du résumé sur SUDOC	482	
Nombre de pages	retrouvé sur la fiche SUDOC	540	
Langue résumé	Bilingue	199	539
	Français	340	
Thèse en ligne si résumé en ligne	Non	285	
	Oui	198	
Directeur de thèse	quantitatif + mixte	407	
Datégorie directeur de thèse	quantitatif + mixte	407	
Typologie de sous-méthode	quantitatif + mixte	404	
Taille de l'échantillon disponible	quantitatif + mixte, présente dans le résumé ou le texte intégral en ligne	364	403
	quantitatif + mixte, absente du résumé et texte intégral indisponible en ligne	39	
Source des données de l'étude	quantitatif + mixte	404	
Nombre de groupes	quantitatif + mixte	404	
Type de variable	quantitatif + mixte	404	
Valeurs significatives	quantitatif + mixte	404	
Lecture des méthodes des textes intégraux	quantitatif ou mixte, analyse statistique possible ou certaine	165	
Détermination paramètres des méthodes des textes intégraux	quantitatif ou mixte, analyse statistique possible ou certaine	162	

Furent lus en Bibliothèque Universitaire : 55 résumés, 93 textes intégraux.

SERMENT MÉDICAL

Au moment d'être admis à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences. Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer leurs consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera.

Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis dans l'intimité des personnes, je tairai les secrets qui me sont confiés. Reçu à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs.

Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité.

Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses : que je sois déshonoré et méprisé si j'y manque.

Résumé / Summary

Titre : Méthode quantitative utilisée dans les thèses portant sur les soins de santé primaires. Étude rétrospective en Aquitaine sur 10 ans.

Introduction

Au sein des méthodes de recherche quantitative disponibles pour leur travaux en soins de santé primaires, quels étaient les choix des étudiants de médecine générales ces 10 dernières années en Aquitaine ? L'analyse statistique est-elle fréquente dans ces travaux ? L'étudiant a-t-il recours à un tiers pour l'assister dans cette analyse ? Le directeur de thèse influence-t-il ce choix ?

Méthodes

Etude descriptive transversale de l'ensemble des données issues d'APOGÉE de 2006 à 2016. Inclusions selon des tris successifs des titres puis des résumés. Lecture des textes intégraux pour appréciation de l'analyse statistique. Analyse comparative des paramètres obtenus, recherche de liens entre les thèses, le recours à un tiers et la catégorie du directeur de thèse.

Résultats

1179 travaux furent inclus, 606 (51,40%) traitaient des soins de santé primaires dont 375 (61,88%) avaient une méthode quantitative, 162 (40,20%) une méthode qualitative et 28 (4,62%) une méthode mixte. La proportion de thèses de soins de santé primaires a augmenté de 19% sur 10 ans. Le ratio entre méthodes quantitatives et qualitatives semble stable. Les typologies les plus représentées étaient les études transversales (26,80%), les cohortes (18,36%), les enquêtes de pratiques (7,94%), les études observationnelles avant-après (6,70%) et les entretiens individuels structurés (6,70%), 162 (40,20%) déclaraient une analyse statistique, 40 (24,69%) déclaraient une analyse statistique complexe. Le recours à un tiers n'aurait pas de lien avec le directeur de thèse. Un directeur de thèse membre du CHU pourrait avoir une influence sur la présence d'arguments statistiques dans le résumé.

Discussion

Les étudiants ont besoin d'être formés spécifiquement pour les soins de santé primaires, cette tendance devrait augmenter régulièrement. Le Département de Médecine Générale de Bordeaux dispose de pistes pour l'enseignement des futurs thésards et la valorisation de la recherche. La typologie de méthode devrait faire l'objet d'un consensus.

Title : Quantitative methodology used in primary care related thesis. Retrospective study on 10 years in Aquitaine.

Introduction

Among the different quantitative methods available for their research work regarding primary care, which choices general medicine student made on the last 10 years in Aquitaine ? Is statistical analysis frequent ? Does the student gets help from a third-party for analysis ? Does the thesis director influent that choice ?

Methods

Transversal study of the whole data collected from APOGÉE from 2006 to 2016. Eligibility regarding successive sorting of titles then summaries. Integral texts red for statistical analysis investigation. Comparative analysis of the results, search for links between results, third-party resort and thesis director category.

Results

1179 thesis were included, 606 (51,40%) regarding primary care, of which 375 (61,88%) had a quantitative method, 162 (40,20%) a qualitative method and 28 (4,62%) a mixed method. Primary care thesis proportion has increased by 19% on 10 years. Quantitative versus qualitative methods ratio seems to be stable. Most frequent typologies were transversal studies (26,80%), cohorts (18,36%), practice surveys (7,94%), before-after observational studies (6,70%), structured individual interviews (6,70%), 162 (40,20%) disclosed a statistical analysis, 40 (24,69%) disclosed a complex statistical analysis. Resort to a third-party for statistical analysis has no link with the thesis director. A teaching hospital thesis director would influence statistical analysis elements presence in the summary.

Discussion

Students require specific educational background for primary care, this trend is likely to regularly increase. Bordeaux General medicine department dispose of leads for students teaching and promote research. Method typology should undergo a consensus.

Mots-clés : Thèse d'exercice, Médecine générale, DMG Bordeaux, Département de Médecine Générale de Bordeaux, Recherche en médecine, Université de Bordeaux, Thèse sur les thèses.