

FACULTE DE MEDECINE ET DE PHARMACIE DE ROUEN

ANNEE 2017

Thèse pour le
DOCTORAT EN MEDECINE
DIPLOME D'ETAT

Par

Léonard CHANSAVANG

Né le 28/12/1984 à Paris 14^{ème}

Présentée et soutenue publiquement le 28 février 2017

**Création et validation d'un filtre de recherche en
médecine générale sur PubMed**

PRESIDENT DU JURY

Professeur Stéfan DARMONI

DIRECTEUR DE THESE

Docteur Matthieu SCHUERS

MEMBRES DU JURY

Professeur Jacques BENICHOU

Docteur Joël LADNER

D.E.S. de Médecine Générale

ANNEE UNIVERSITAIRE 2016 - 2017
U.F.R. DE MEDECINE ET DE-PHARMACIE DE ROUEN

DOYEN : **Professeur Pierre FREGER**

ASSESEURS : **Professeur Michel GUERBET**

Professeur Benoit VEBER

Professeur Pascal JOLY

Professeur Stéphane MARRET

I - MEDECINE

PROFESSEURS DES UNIVERSITES – PRATICIENS HOSPITALIERS

Mr Frédéric ANSELME	HCN	Cardiologie
Mme Isabelle AUQUIT AUCKBUR	HCN	Chirurgie plastique
Mr Fabrice BAUER	HCN	Cardiologie
Mme Soumeya BEKRI	HCN	Biochimie et biologie moléculaire
Mr Ygal BENHAMOU	HCN	Médecine interne
Mr Jacques BENICHOU	HCN	Bio statistiques et informatique médicale
Mme Bouchra LAMIA	Havre	Pneumologie
Mr Olivier BOYER	UFR	Immunologie
Mr François CARON	HCN	Maladies infectieuses et tropicales
Mr Philippe CHASSAGNE (<i>détachement</i>)	HCN	Médecine interne (gériatrie) – Détachement
Mr Vincent COMPERE	HCN	Anesthésiologie et réanimation chirurgicale
Mr Jean-Nicolas CORNU	HCN	Urologie
Mr Antoine CUVELIER	HB	Pneumologie
Mr Pierre CZERNICHOW (<i>surnombre</i>)	HCH	Epidémiologie, économie de la santé
Mr Jean-Nicolas DACHER	HCN	Radiologie et imagerie médicale

Mr Stéfán DARMONI	HCN	Informatique médicale et techniques de communication
Mr Pierre DECHELOTTE	HCN	Nutrition
Mr Stéphane DERREY	HCN	Neurochirurgie
Mr Frédéric DI FIORE	CB	Cancérologie
Mr Fabien DOGUET	HCN	Chirurgie Cardio Vasculaire
Mr Jean DOUCET	SJ	Thérapeutique - Médecine interne et gériatrie
Mr Bernard DUBRAY	CB	Radiothérapie
Mr Philippe DUCROTTE	HCN	Hépto-gastro-entérologie
Mr Frank DUJARDIN	HCN	Chirurgie orthopédique - Traumatologique
Mr Fabrice DUPARC	HCN	Anatomie - Chirurgie orthopédique et traumatologique
Mr Eric DURAND	HCN	Cardiologie
Mr Bertrand DUREUIL	HCN	Anesthésiologie et réanimation chirurgicale
Mme Hélène ELTCHANINOFF	HCN	Cardiologie
Mr Thierry FREBOURG	UFR	Génétique
Mr Pierre FREGER	HCN	Anatomie - Neurochirurgie
Mr Jean François GEHANNO	HCN	Médecine et santé au travail
Mr Emmanuel GERARDIN	HCN	Imagerie médicale
Mme Priscille GERARDIN	HCN	Pédopsychiatrie
Mr Michel GODIN (<i>surnombre</i>)	HB	Néphrologie
M. Guillaume GOURCEROL	HCN	Physiologie
Mr Dominique GUERROT	HCN	Néphrologie
Mr Olivier GUILLIN	HCN	Psychiatrie Adultes
Mr Didier HANNEQUIN	HCN	Neurologie
Mr Fabrice JARDIN	CB	Hématologie
Mr Luc-Marie JOLY	HCN	Médecine d'urgence
Mr Pascal JOLY	HCN	Dermato - Vénérologie
Mme Annie LAQUERRIERE	HCN	Anatomie et cytologie pathologiques
Mr Vincent LAUDENBACH	HCN	Anesthésie et réanimation chirurgicale
Mr Joël LECHEVALLIER	HCN	Chirurgie infantile
Mr Hervé LEFEBVRE	HB	Endocrinologie et maladies métaboliques
Mr Thierry LEQUERRE	HB	Rhumatologie
Mme Anne-Marie LEROI	HCN	Physiologie

Mr Hervé LEVESQUE	HB	Médecine interne
Mme Agnès LIARD-ZMUDA	HCN	Chirurgie Infantile
Mr Pierre Yves LITZLER	HCN	Chirurgie cardiaque
Mr Bertrand MACE	HCN	Histologie, embryologie, cytogénétique
M. David MALTETE	HCN	Neurologie
Mr Christophe MARGUET	HCN	Pédiatrie
Mme Isabelle MARIE	HB	Médecine interne
Mr Jean-Paul MARIE	HCN	Oto-rhino-laryngologie
Mr Loïc MARPEAU	HCN	Gynécologie - Obstétrique
Mr Stéphane MARRET	HCN	Pédiatrie
Mme Véronique MERLE	HCN	Epidémiologie
Mr Pierre MICHEL	HCN	Hépatogastro-entérologie
M. Benoit MISSET	HCN	Réanimation Médicale
Mr Jean-François MUIR (<i>surnombre</i>)	HB	Pneumologie
Mr Marc MURAINÉ	HCN	Ophtalmologie
Mr Philippe MUSETTE	HCN	Dermatologie - Vénérologie
Mr Christophe PEILLON	HCN	Chirurgie générale
Mr Christian PFISTER	HCN	Urologie
Mr Jean-Christophe PLANTIER	HCN	Bactériologie - Virologie
Mr Didier PLISSONNIER	HCN	Chirurgie vasculaire
Mr Gaëtan PREVOST	HCN	Endocrinologie
Mr Bernard PROUST	HCN	Médecine légale
Mr Jean-Christophe RICHARD (<i>détachement</i>)	HCN	Réanimation médicale - Médecine d'urgence
Mr Vincent RICHARD	UFR	Pharmacologie
Mme Nathalie RIVES	HCN	Biologie du développement et de la reproduction
Mr Horace ROMAN	HCN	Gynécologie - Obstétrique
Mr Jean-Christophe SABOURIN	HCN	Anatomie - Pathologie
Mr Guillaume SAVOYE	HCN	Hépatogastrologie
Mme Céline SAVOYE-COLLET	HCN	Imagerie médicale
Mme Pascale SCHNEIDER	HCN	Pédiatrie
Mr Michel SCOTTE	HCN	Chirurgie digestive
Mme Fabienne TAMION	HCN	Thérapeutique
Mr Luc THIBERVILLE	HCN	Pneumologie

Mr Christian THUILLEZ (<i>surnombre</i>)	HB	Pharmacologie
Mr Hervé TILLY	CB	Hématologie et transfusion
M. Gilles TOURNEL	HCN	Médecine Légale
Mr Olivier TROST	HCN	Chirurgie Maxillo-Faciale
Mr Jean-Jacques TUECH	HCN	Chirurgie digestive
Mr Jean-Pierre VANNIER (<i>surnombre</i>)	HCN	Pédiatrie génétique
Mr Benoît VEBER	HCN	Anesthésiologie - Réanimation chirurgicale
Mr Pierre VERA	CB	Biophysique et traitement de l'image
Mr Eric VERIN	HB	Service Santé Réadaptation
Mr Eric VERSPYCK	HCN	Gynécologie obstétrique
Mr Olivier VITTECOQ	HB	Rhumatologie
Mr Jacques WEBER	HCN	Physiologie

MAITRES DE CONFERENCES DES UNIVERSITES – PRATICIENS HOSPITALIERS

Mme Noëlle BARBIER-FREBOURG	HCN	Bactériologie – Virologie
Mme Carole BRASSE LAGNEL	HCN	Biochimie
Mme Valérie BRIDOUX HUYBRECHTS	HCN	Chirurgie Vasculaire
Mr Gérard BUCHONNET	HCN	Hématologie
Mme Mireille CASTANET	HCN	Pédiatrie
Mme Nathalie CHASTAN	HCN	Neurophysiologie
Mme Sophie CLAEYSSENS	HCN	Biochimie et biologie moléculaire
Mr Moïse COEFFIER	HCN	Nutrition
Mr Manuel ETIENNE	HCN	Maladies infectieuses et tropicales
Mr Serge JACQUOT	UFR	Immunologie
Mr Joël LADNER	HCN	Epidémiologie, économie de la santé
Mr Jean-Baptiste LATOUCHE	UFR	Biologie cellulaire
Mr Thomas MOUREZ	HCN	Virologie
Mme Muriel QUILLARD	HCN	Biochimie et biologie moléculaire
Mme Laëtitia ROLLIN	HCN	Médecine du Travail
Mr Mathieu SALAUN	HCN	Pneumologie
Mme Pascale SAUGIER-VEBER	HCN	Génétique

Mme Anne-Claire TOBENAS-DUJARDIN	HCN	Anatomie
Mr David WALLON	HCN	Neurologie

PROFESSEUR AGREGE OU CERTIFIE

Mme Dominique LANIEZ	UFR	Anglais – retraite 01/10/2016
Mr Thierry WABLE	UFR	Communication

II - PHARMACIE

PROFESSEURS

Mr Thierry BESSON	Chimie Thérapeutique
Mr Jean-Jacques BONNET	Pharmacologie
Mr Roland CAPRON (PU-PH)	Biophysique
Mr Jean COSTENTIN (Professeur émérite)	Pharmacologie
Mme Isabelle DUBUS	Biochimie
Mr Loïc FAVENNEC (PU-PH)	Parasitologie
Mr Jean Pierre GOULLE (Professeur émérite)	Toxicologie
Mr Michel GUERBET	Toxicologie
Mme Isabelle LEROUX - NICOLLET	Physiologie
Mme Christelle MONTEIL	Toxicologie
Mme Martine PESTEL-CARON (PU-PH)	Microbiologie
Mme Elisabeth SEGUIN	Pharmacognosie
Mr Rémi VARIN (PU-PH)	Pharmacie clinique
Mr Jean-Marie VAUGEOIS	Pharmacologie
Mr Philippe VERITE	Chimie analytique

MAITRES DE CONFERENCES

Mme Cécile BARBOT	Chimie Générale et Minérale
Mr Jérémy BELLIEN (MCU-PH)	Pharmacologie
Mr Frédéric BOUNOURE	Pharmacie Galénique
Mr Abdeslam CHAGRAOUI	Physiologie
Mme Camille CHARBONNIER (LE CLEZIO)	Statistiques
Mme Elizabeth CHOSSON	Botanique
Mme Marie Catherine CONCE-CHEMTOB santé	Législation pharmaceutique et économie de la
Mme Cécile CORBIERE	Biochimie

Mr Eric DITTMAR	Biophysique
Mme Nathalie DOURMAP	Pharmacologie
Mme Isabelle DUBUC	Pharmacologie
Mme Dominique DUTERTE- BOUCHER	Pharmacologie
Mr Abdelhakim ELOMRI	Pharmacognosie
Mr François ESTOUR	Chimie Organique
Mr Gilles GARGALA (MCU-PH)	Parasitologie
Mme Nejla EL GHARBI-HAMZA	Chimie analytique
Mme Marie-Laure GROULT	Botanique
Mr Hervé HUE	Biophysique et mathématiques
Mme Laetitia LE GOFF	Parasitologie – Immunologie
Mme Hong LU	Biologie
Mme Marine MALLETER	Toxicologie
Mme Sabine MENAGER	Chimie organique
Mme Tiphaine ROGEZ-FLORENT	Chimie analytique
Mr Mohamed SKIBA	Pharmacie galénique
Mme Malika SKIBA	Pharmacie galénique
Mme Christine THARASSE	Chimie thérapeutique
Mr Frédéric ZIEGLER	Biochimie

PROFESSEURS ASSOCIES

Mme Cécile GUERARD-DETUNCQ	Pharmacie officinale
Mr Jean-François HOUIVET	Pharmacie officinale

PROFESSEUR CERTIFIE

Mme Mathilde GUERIN	Anglais
----------------------------	---------

ASSISTANT HOSPITALO-UNIVERSITAIRE

Mme Sandrine DAHYOT	Bactériologie
----------------------------	---------------

ATTACHES TEMPORAIRES D'ENSEIGNEMENT ET DE RECHERCHE

Mr Souleymane ABDOUL-AZIZE	Biochimie
-----------------------------------	-----------

Mme Hanane **GASMI**

Galénique

Mme Caroline **LAUGEL**

Chimie organique

Mr Romy **RAZAKANDRAINIBE**

Parasitologie

LISTE DES RESPONSABLES DES DISCIPLINES PHARMACEUTIQUES

Mme Cécile BARBOT	Chimie Générale et minérale
Mr Thierry BESSON	Chimie thérapeutique
Mr Roland CAPRON	Biophysique
Mme Marie-Catherine CONCE-CHEMTOB	Législation et économie de la santé
Mme Elisabeth CHOSSON	Botanique
Mr Jean-Jacques BONNET	Pharmacodynamie
Mme Isabelle DUBUS	Biochimie
Mr Loïc FAVENNEC	Parasitologie
Mr Michel GUERBET	Toxicologie
Mr François ESTOUR	Chimie organique
Mme Isabelle LEROUX-NICOLLET	Physiologie
Mme Martine PESTEL-CARON	Microbiologie
Mme Elisabeth SEGUIN	Pharmacognosie
Mr Mohamed SKIBA	Pharmacie galénique
Mr Rémi VARIN	Pharmacie clinique
Mr Philippe VERITE	Chimie analytique

III – MEDECINE GENERALE

PROFESSEUR

Mr Jean-Loup **HERMIL**

UFR Médecine générale

PROFESSEURS ASSOCIES A MI-TEMPS

Mr Emmanuel **LEFEBVRE**

UFR Médecine Générale

Mme Elisabeth **MAUVIARD**

UFR Médecine générale

Mr Philippe **NGUYEN THANH**

UFR Médecine générale

MAITRE DE CONFERENCES ASSOCIE A MI-TEMPS

Mr Pascal **BOULET**

UFR Médecine générale

Mr Emmanuel **HAZARD**

UFR Médecine Générale

Mme Lucile **PELLERIN**

UFR Médecine générale

Mme Yveline **SEVRIN**

UFR Médecine générale

Mme Marie Thérèse **THUEUX**

UFR Médecine générale

ENSEIGNANTS MONO-APPARTENANTS

PROFESSEURS

Mr Serguei FETISSOV (med)	Physiologie (ADEN)
Mr Paul MULDER (phar)	Sciences du Médicament
Mme Su RUAN (med)	Génie Informatique

MAITRES DE CONFERENCES

Mr Sahil ADRIOUCH (med) 905)	Biochimie et biologie moléculaire (Unité Inserm 905)
Mme Gaëlle BOUGEARD-DENOYELLE (med)	Biochimie et biologie moléculaire (UMR 1079)
Mme Carine CLEREN (med)	Neurosciences (Néovasc)
M. Sylvain FRAINEAU (phar)	Physiologie (Inserm U 1096)
Mme Pascaline GAILDRAT (med)	Génétique moléculaire humaine (UMR 1079)
Mr Nicolas GUEROUT (med)	Chirurgie Expérimentale
Mme Rachel LETELLIER (med)	Physiologie
Mme Christine RONDANINO (med)	Physiologie de la reproduction
Mr Antoine OUVRARD-PASCAUD (med)	Physiologie (Unité Inserm 1076)
Mr Frédéric PASQUET	Sciences du langage, orthophonie
Mme Isabelle TOURNIER (med)	Biochimie (UMR 1079)

CHEF DES SERVICES ADMINISTRATIFS : Mme Véronique DELAFONTAINE

HCN - Hôpital Charles Nicolle

HB - Hôpital de BOIS GUILLAUME

CB - Centre Henri Becquerel

CHS - Centre Hospitalier Spécialisé du Rouvray

CRMPR - Centre Régional de Médecine Physique et de Réadaptation

SJ - Saint Julien Rouen

Par délibération en date du 3 mars 1967, la faculté a arrêté que les opinions émises dans les dissertations qui lui seront présentées doivent être considérées comme propres à leurs auteurs et qu'elle n'entend leur donner aucune approbation ni improbation.

Remerciements

Au Professeur Stéfan DARMONI, pour avoir accepté de présider mon jury de thèse.

Au Professeur Jacques BENICHOU et au Docteur Joël LADNER pour avoir accepté d'y participer.

Au Docteur Matthieu SCHUERS pour m'avoir proposé ce travail et l'avoir dirigé.

Aux médecins et professionnels paramédicaux avec qui j'ai eu l'occasion de travailler tout au long de ma formation qui m'ont beaucoup apporté tant sur le plan médical que sur le plan humain.

Au Dr Yann POULINGUE pour ton soutien intarissable, ta bienveillance et pour l'opportunité de travailler dans ton équipe. Sois assuré de ma profonde reconnaissance.

Aux Dr Marie LANDRIN, Modou DIOP et à toute l'équipe du service de Médecine Polyvalente de Louviers, pour l'atmosphère de travail formidable que vous créez.

Aux Dr Laurence BRENET et Bruno VEYRON pour m'avoir montré l'exemple du médecin que j'aspire à devenir.

A tous mes ex-co-externes et ex-co-internes pour avoir fait un bout de chemin à mes côtés.

A mes amis, pour votre soutien inconditionnel.

A tous les VGZ, parce-qu'on le reste à vie.

A ma belle-famille pour m'avoir accepté parmi vous.

A ma famille, mes parents, ma sœur, ma grand-mère pour avoir été toujours là pour moi.

A Guillaume, sans qui mon quotidien serait moins beau.

Cette thèse est présentée sous un format article comprenant deux parties:

- Introduction de l'article de 3585 mots
- Article de 3440 mots

Table des matières

REMERCIEMENTS	14
LEXIQUE/ABBREVIATIONS	17
PREMIERE PARTIE : INTRODUCTION	18
1. LA MEDECINE GENERALE	18
2. LA RECHERCHE EN MEDECINE GENERALE	20
3. BASES DE DONNEES BIBLIOGRAPHIQUES	22
3.1. LE MESH	23
3.2. LES QUALIFICATIFS	23
3.3. LE CHAMP	24
3.4. L'EXPLOSION	24
4. STRATEGIE DE RECHERCHE	25
4.1. OPERATEURS BOOLEENS	25
5. DIFFICULTES RENCONTREES PAR LES MEDECINS GENERALISTES	26
6. LES FILTRES DE RECHERCHE	27
7. OBJECTIF DE L'ETUDE	29
BIBLIOGRAPHIE	30
DEUXIEME PARTIE: ARTICLE	33
1. INTRODUCTION	33
2. MATERIEL ET METHODE	35
2.1. DESIGN PRINCIPAL DE L'ETUDE	35
2.2. CONSTITUTION DU GOLD STANDARD	35
2.2.1. CONSTITUTION DE L'ECHANTILLON	35
2.2.2. ANALYSE MANUELLE DE L'ÉCHANTILLON	35
2.3. DÉVELOPPEMENT DE L'ÉQUATION	36
2.3.1. SELECTION DES TERMES MESH CANDIDATS	36
2.3.2. ANALYSE STATISTIQUE	36
2.4. VALIDATION INTERNE DE L'EQUATION	37
2.5. COMPARAISONS AUX FILTRES DE RECHERCHE EXISTANTS	37
3. RESULTATS	38
3.1. GOLD STANDARD	38
3.2. TERMES MESH	38
3.3. EQUATIONS DE RECHERCHE	39
3.4. COMPARATIF DES FILTRES DE RECHERCHE	40
4. DISCUSSION	41
4.1. PRINCIPAUX RESULTATS	41
4.2. METHODOLOGIE	41
4.3. PERSPECTIVES	43
BIBLIOGRAPHIE	45
FIGURES ET TABLEAUX	48
RESUME	56
ABSTRACT	57

LEXIQUE/ABBREVIATIONS

BMJ	British Medical Journal
CNGE	Collège National des Generalistes Enseignants
EBM	Evidence-Based Medicine
GS	Gold Standard
ISSG	InterTASC Information Specialists' Sub-Group
JAMA	Journal of the American Medical Association
LSTRC	Literature Selection Technical Review Committee
MeSH	Medical Subject Headings
NCBI	National Center for Biotechnology Information
NEJM	New England Journal of Medicine
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NLM	U.S. National Library of Medicine
PMC	PubMed Central
SFMG	Société Française de Médecine Générale
WONCA	World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians

Première partie : Introduction

1. La médecine générale

La définition de la médecine générale n'a pas cessé d'évoluer. En 1974, le groupe Leeuwenhorst proposait une définition centrée sur le rôle et les modalités d'exercice du médecin généraliste dans sa communauté. Il définissait le généraliste comme un professionnel médical formé pour prodiguer des soins de première ligne et assurer le suivi d'individus appartenant à une communauté indépendamment de leur âge ou de leur sexe (1). Dans les années 90, on pouvait trouver dans un référentiel une définition uniquement centrée sur les caractéristiques du généraliste. Ainsi on insistait sur le suivi médical des individus dans une communauté, les qualités d'écoute du généraliste, la dimension familiale (2).

Le concept d'Evidence Based Medicine (EBM), médecine fondée sur les preuves, développé dans les années 80, s'est intégré à la pratique clinique médicale (3). L'EBM consiste à baser son raisonnement clinique en tenant compte de plusieurs composantes définies par le paradigme de l'EBM qui combine : les preuves (données de la recherche), la connaissance du médecin (expérience clinique), le choix du patient (ses préférences). Ainsi, la décision médicale doit se prendre en tenant compte de ces 3 paramètres (3–5). (Figure 1)

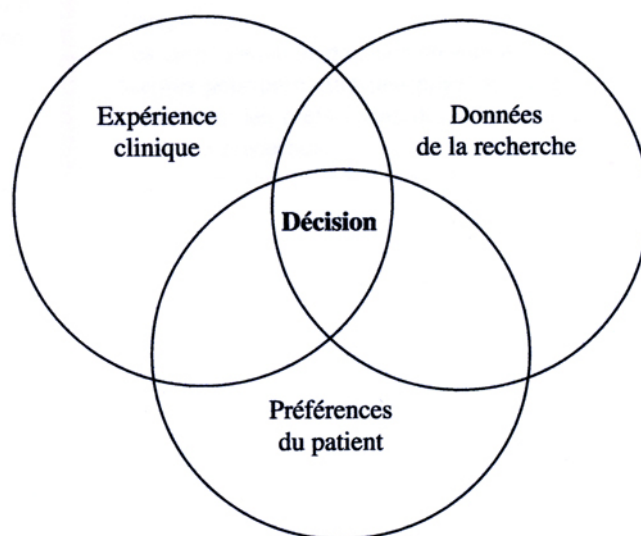


Figure 1 Paradigme de l'EBM

L'intégration de la dimension scientifique de l'EBM en médecine générale a donné naissance à d'autres définitions à la fin des années 90 soulignant le besoin développer la recherche en médecine générale ou en soins primaires (6,7). La notion de soins primaires, définie par la World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA), comme « le cadre du système de santé qui est, généralement, la communauté à laquelle le patient appartient et où s'établit le premier contact avec un professionnel de la santé » (8). Elle intègre, en plus de la médecine générale, les soins infirmiers, la kinésithérapie, l'odontologie, la santé publique ou encore la psychologie.

L'intégration de l'EBM à la pratique implique de définir le contenu de base de la discipline avec une nécessité de délimiter les frontières de la médecine générale. Pour produire des données validées par la science en médecine générale il convient d'abord de définir le cadre de recherche. Cependant en médecine, les frontières entre spécialités sont parfois floues. Par sa position d'acteur de santé de première ligne, de « gatekeeper » avec une position de pivot entre le patient et les autres spécialités, le généraliste fait face à des situations cliniques transversales. Définir la médecine générale selon son contenu de base, c'est accepter les chevauchements avec d'autres spécialités (9).

La médecine générale occupe dans de nombreux systèmes de santé comme le nôtre, une place particulière de première ligne. Le médecin généraliste est confronté à un panel de problèmes de santé transversaux dont la recherche est couverte par les autres spécialités. Mais sa position d'acteur de soins de santé primaires justifie l'existence d'un questionnement scientifique spécifique, non pris en charge par les autres spécialités, à l'exception partielle de la santé publique (10–12). Pour répondre aux questions cliniques que le médecin généraliste se pose, transversales mais aussi spécifiques des soins primaires, il doit s'appuyer sur des données scientifiques validées produites par la recherche en médecine générale.

2. La recherche en médecine générale

Le médecin généraliste a besoin de données scientifiquement validées pour prendre une décision médicale selon l'EBM. En 1961, un article intitulé "The Ecology of Medical Care" estimait, pour les Etats-Unis et la Grande Bretagne, que dans une population de 1 000 habitants, 750 signalaient une maladie, 250 consultaient un médecin, 9 étaient hospitalisés, 5 dirigés vers un autre médecin et 1 hospitalisé dans un centre universitaire (13). Green et al. ont réactualisé ces données en 2001 et ont trouvé des résultats similaires (14). Ces résultats ont conduit à préconiser un rééquilibrage de la formation médicale et de la recherche clinique de l'hôpital et les spécialités vers la médecine de première ligne. L'étude des problèmes de santé les plus fréquents dans la population devenait une priorité.

Les données validées en médecine générale étaient peu présentes dans les bases de données numériques au début des années 2000 (15). Les travaux de Starfield soulignant l'impact positif des soins primaires dans l'amélioration de l'état de santé des populations dans de nombreux pays ont renforcé la nécessité d'accentuer le développement de la recherche en médecine générale (16,17). L'impulsion de développer la recherche en médecine générale a été suivie par les institutions.

En 2003, la WONCA organisait une conférence : « Améliorer la santé mondiale : la nécessité de la recherche en médecine générale ». Elle faisait le constat que le développement d'une médecine générale de haute qualité fondée sur les preuves nécessitait de développer la recherche dans le domaine de la médecine générale et des soins primaires à tous les niveaux de formations. En 2006, le rapport de De Pourville décrivait la nécessité de produire des données scientifiques pertinentes pour les professionnels de santé situés en première ligne afin de répondre aux besoins du patient. Il rappelait que l'activité des médecins généralistes recouvrait une part importante des soins dispensés dans notre système de santé. Il précisait également qu'à l'instar de la recherche hospitalière et des spécialités d'organes, la recherche en médecine générale devait intégrer les dimensions économiques, sociales et psychologiques. Elle devait également développer l'analyse des comportements des médecins et des patients. Plus globalement, la recherche en médecine générale devait s'intégrer à la recherche en soins primaires impliquant le développement d'équipes pluridisciplinaires (10).

Ainsi, progressivement, la production des données de recherche dans le champ de la médecine générale a augmenté en quantité et en qualité. Une étude récente montrait que le nombre de publications en médecine générale passait de 793 en 1980 à plus de 2 000 en 2010, avec un record de plus de 2 500 articles publiés en 2014 (18). Parmi les plus gros

producteurs, on pouvait citer les Etats-Unis, le Royaume-Uni ou encore l'Australie. On pouvait trouver les données dans des revues généralistes comme le « British Medical Journal » (BMJ) ou spécialisées en médecine générale comme « British Journal of General Practice » ou encore « Family Medicine » (18).

3. Bases de données bibliographiques

Les données validées par la recherche se trouvent dans des bases de données bibliographiques. Les praticiens et les chercheurs peuvent accéder à ces bases de données sur internet. Il en existe plusieurs, selon le domaine de recherche ou la langue de production des références. MEDLINE (Medical Literature Analysis and Retrieval System Online) est la plus importante base de données couvrant les domaines de la médecine, les soins infirmiers, la chirurgie dentaire, la science vétérinaire, des sciences de la santé en général. Elle est la plus utilisée. Elle est produite par la U.S. National Library of Medicine (NLM) de la National Institutes of Health (NIH) (19). L'accès à la base de données se fait via une interface numérique (moteur de recherche). Il en existe plusieurs, PubMed, NLM Gateway ou Medscape proposant un accès gratuit. Ovid, Knowledge Finder requièrent un accès payant.

PubMed est le moteur de recherche le plus populaire. Il en accès libre et est développé par le National Center for Biotechnology Information (NCBI). Il utilise la base de données MEDLINE. PubMed donne accès à des références bibliographiques incluses ou en cours d'indexation dans MEDLINE mais également à des liens vers le texte intégral, disponible gratuitement ou payant sur le site de l'éditeur. PubMed contient plus de 26 millions de références de la littérature biomédicale internationale de MEDLINE. PubMed couvre les domaines de la médecine et des sciences de la santé, des sciences de la vie, des sciences comportementales, des sciences chimiques et de la bioingénierie. Les références sont issues de plus de 5 000 revues sélectionnées et indexées par le Literature Selection Technical Review Committee (LSTRC). PubMed est mis à jour quotidiennement et environ un million de notices sont indexées annuellement, c'est à dire identifiées et classées. Son utilisation se fait intégralement en anglais (20,21)

Lorsqu'on lance une requête dans PubMed, on obtient une liste de références. Une référence, également appelée notice bibliographique, est un ensemble d'éléments permettant d'identifier le document original. Elle est composée d'un titre, d'un résumé et d'un bordereau d'indexation. Dans ces derniers, on trouve des descripteurs MeSH associés à des qualificatifs et des descriptions du type de l'étude (étude comparative; revue de la littérature) ou le genre de la population étudiée. Dans la référence se trouve également un lien vers le document original. Ce dernier peut être gratuit, lorsqu'il est présent dans les archives PubMed Central (PMC). Cependant son accès peut être payant sur le site de l'éditeur, parfois les abonnements des institutions (université, hôpital, laboratoire) couvrent l'accès à ces revues (21,22).

3.1. Le MeSH

Le point de départ d'une recherche bibliographique est la construction de la requête utilisée pour interroger la base de données. Elle doit être rédigée en respectant la syntaxe de l'interface utilisée afin d'être optimale, c'est à dire retrouver le plus grand nombre de notices bibliographiques pertinentes pour le sujet étudié. Le contenu de MEDLINE est indexé par les termes du thésaurus utilisant un vocabulaire contrôlé, le MeSH (Medical Subject Headings). L'utilisation des termes MeSH plutôt que du texte libre (« text word ») dans une requête PubMed augmente la probabilité de retrouver des références précises pour un sujet donné (23).

Un thésaurus est une liste de termes normalisés. Ces termes sélectionnés sont appelés descripteurs ou mots-clés, servant à l'indexation de documents. Le thésaurus MeSH est organisé selon des relations hiérarchiques. Les descripteurs MeSH sont regroupés en 16 catégories (anatomie, organismes, maladies...). Chaque catégorie est divisée en concepts. La définition de ces derniers est également précisée par des termes plus spécifiques. Il existe jusque 11 niveaux hiérarchiques et le terme appartenant au niveau le plus élevé sera le plus précis. Un descripteur peut appartenir à plusieurs arborescences.

Les termes MeSH sont également organisés selon des relations sémantiques. Un seul terme MeSH est indexé pour décrire un concept. Les termes rejetés, non-descripteurs, ou synonymes, appelés également « Entry terms » sont liés aux descripteurs par une relation d'équivalence.

3.2. Les qualificatifs

Il existe aussi des qualificatifs « Subheadings » qui sont des concepts généraux. Il en existe à ce jour 83 et ont aussi des relations hiérarchiques entre eux. Ils précisent le sens d'un descripteur et accroissent sa spécificité. Par exemple, pour le descripteur MeSH « Alzheimer's disease », on trouvera 32 synonymes comme « senile dementia » ou encore « Alzheimer Syndrome » et parmi les qualificatifs : « diagnosis », « genetics » ou encore « drug therapy ». Dans une notice bibliographique, les qualificatifs suivent un descripteur et sont précédés d'un slash « / » (21,24).

3.3. Le champ

La notion de champ s'applique à un mot et peut préciser soit sa nature s'il s'agit d'un descripteur, d'un auteur ou d'un type de publication, soit sa localisation : mot du titre ou du résumé par exemple (22). Le champ suit un mot et s'écrit entre crochets « [] ». (Figure 2)

MesH Terms	[mh]	Descripteur du thésaurus
All fields	[all]	Inclut tous les champs de PubMed. (par défaut)
Author	[au]	Nom de l'auteur +/- initiales
Journal title	[ta]	Nom du journal
Language	[la]	Langue de publication
MeSH Subheadings	[sh]	Qualificatif
Publication date	[dp]	Date de publication, format a/m/j
Publication type	[pt]	Type de publication
Text Word	[tw]	Mots des titres, résumés et bordereaux d'indexation
Title	[ti]	Mot des titres
Title/Abstract	[tiab]	Mot des titres et résumés

Figure 2. Principaux champs utilisés dans PubMed

3.4. L'explosion

La notion d'explosion d'un descripteur correspond à l'extension de la requête effectuée sur le descripteur à ses descendants hiérarchiques. L'explosion concerne également les qualificatifs. Dans PubMed, elle se fait par défaut. Dans ce cas, la requête ne se limite pas seulement au descripteur, mais est étendue aux descendants hiérarchiques (22).

Le thésaurus MeSH est spécialisé en sciences biomédicales et en médecine. En 2016 il compte 27 883 descripteurs avec plus de 87 000 synonymes. Il est renouvelé annuellement par la NLM avec en moyenne 400 nouveaux descripteurs (21,25).

4. Stratégie de recherche

L'utilisation de PubMed est différente de celle d'un moteur de recherche comme Google, Yahoo ou Bing. Lancer une recherche avec un langage courant (libre), en opposition au langage contrôlé du thesaurus, aboutit à un nombre trop élevé de résultats. La liste de références obtenue est difficilement exploitable et contient des résultats ne correspondant pas au sujet recherché (bruit). Une formulation de requête correcte prend en compte les notions de descripteur, de qualificatif, de champ et d'explosion. L'utilisation du thesaurus MeSH est donc recommandée pour une recherche pertinente (23).

L'équation de recherche est formulée de façon linéaire en utilisant le langage MeSH. Les résultats d'une première requête sont rarement satisfaisants et il est nécessaire de prendre en compte les notions de sensibilité et de précision lorsqu'on affine sa requête. Pour cela, on utilise des opérateurs booléens AND, OR ou NOT, écrits en majuscules.

La formulation linéaire est organisée de manière séquentielle où se succèdent mots-clés, associés à leur champ, et opérateurs booléens.

Le module terminologique de l'équipe CISMéF (Catalogue et Index des Sites Médicaux Francophones) du CHU de Rouen, HeTOP (Health Terminology / Ontology Portal) en anglais ou Portail Terminologique de Santé (PTS), permet de traduire un mot clé en descripteur MeSH. Il peut traduire en anglais à partir de 42 langues différentes. Il peut également lancer une requête via son outil Doc'CISMéF.

4.1. Opérateurs booléens

Les opérateurs booléens permettent d'associer des termes MeSH :

L'opérateur AND (ET) permet de réaliser l'intersection de deux ou plusieurs éléments qui doivent figurer dans les références recherchées, il restreint la recherche.

L'opérateur OR (OU) permet de réaliser la réunion de deux ou plusieurs éléments. Il est utile pour élargir les recherches, il réalise une addition.

L'opérateur NOT (SAUF) permet d'exclure les résultats liés au terme suivant l'opérateur.

Il est également utile de préciser que l'utilisation d'opérateurs booléens différents dans une même requête implique impérativement l'utilisation de parenthèses « () », comme on écrit le produit d'une somme (22).

En pratique, on pourra utiliser le « PubMed search builder » pour construire une requête composée via le MeSH Database (base de données MeSH de PubMed) ou entrer directement la requête dans l'« Advanced search » à partir de la page d'accueil.

5. Difficultés rencontrées par les médecins généralistes

Le médecin généraliste fait face à un éventail de situations cliniques dans sa pratique. Leur prise en charge amène le praticien à se poser beaucoup de questions cliniques (26). On peut trouver des réponses à ces questions sur internet, les données validées par la recherche sont présentes dans les bases de données. Cependant une étude a montré qu'au début des années 2000, les médecins généralistes ne consultaient les bases de données électroniques que très rarement (27,28). En effet bien que l'information soit disponible, le médecin généraliste continue à rencontrer des difficultés lorsqu'il recherche une réponse à une question clinique (15,16). Malgré la production et l'utilisation grandissante du contenu en ligne, les praticiens peinent à trouver réponse à leurs questions (30).

Parmi les difficultés rencontrées par les médecins généralistes lors de la recherche d'information en ligne, le manque de temps est majoritairement cité dans la littérature (26, 27). Lors d'une consultation, le praticien dispose de peu de temps pour rechercher des réponses à ses questions (33). Alors que les bases de données en ligne offrent des informations validées en soins primaires, l'accès aux réponses à des questions précises n'est pas assez rapide pour un généraliste dans sa pratique quotidienne (26). Lorsqu'ils recherchent une information sur internet, les médecins généralistes passent 2,4 à 6,5 minutes en ligne pour trouver une réponse adéquate (26).

La difficulté à utiliser les bases de données, le volume important de références à trier et l'hétérogénéité de la qualité de l'information est également fréquemment rapportée (31–33). L'utilisation de PubMed pour accéder à MEDLINE peut s'avérer complexe sans initiation préalable. Le manque de compétence dans la construction de requêtes de recherche est aussi souligné, influençant directement la qualité de l'information retrouvée (12). En effet, l'utilisation correcte des termes MeSH et des opérateurs booléens n'est pas évidente et peut peser sur la pertinence des résultats. Les médecins généralistes rencontrent également des difficultés à identifier l'information pertinente parmi le volume important de références retrouvées lors d'une requête (12,15).

6. Les filtres de recherche

Certaines difficultés ont pu être surmontées. Une base de données bibliographiques de la Littérature Scientifique en Santé de langue française (LiSSa) est proposée par l'équipe de CISMeF, permettant ainsi de supprimer la barrière de la langue. (34) Un autre outil en langue française, le Constructeur de Requêtes Bibliographiques Médicales (CRBM) permet de faciliter l'utilisation des mots-clés MeSH et des opérateurs booléens pour construire une requête (35).

Une solution a également été proposée pour faciliter l'accès aux informations pertinentes pour une question donnée : l'utilisation d'un filtre de recherche. Un filtre de recherche est une requête complexe prédéfinie qui a pour but de retrouver un type de références précisé par la composition de l'équation. Il peut être combiné avec une autre équation. Les filtres de recherche permettent d'affiner, d'orienter la recherche. Ils sont développés pour augmenter le rappel ou optimiser la précision. Ils représentent un outil précieux lors d'une recherche bibliographique. En effet, ils font gagner du temps au chercheur en réduisant le bruit (36).

L'équipe britannique de la InterTASC Information Specialists' Sub-Group (ISSG), travaillant avec la National Institute for Health and Care Excellence (NICE), et son collectif spécialisé dans les filtres de recherche, représente une référence en termes d'identification, de développement et d'évaluation de filtres de recherche (36–40). Ils ont notamment publié une check-list pour évaluer la qualité d'un filtre de recherche mais également contribué au développement de filtres de recherche (36).

Ainsi en utilisant la méthodologie décrite par le InterTASC ISSG, des filtres de recherche ont été créés dans plusieurs disciplines. Une équipe australienne a développé un filtre de recherche en soins palliatifs, en oncologie pulmonaire et en soins primaires (39–41). Des canadiens ont également élaboré un filtre de recherche en soins primaires (38). Concernant la médecine générale, nous n'avons retrouvé qu'un seul article d'origine néerlandaise. Dans cette étude, les auteurs avaient développé deux filtres de recherche. Un premier « sensible », avec un rappel élevé destiné aux chercheurs. Le deuxième « spécifique », utilisable par les cliniciens pour répondre à une question précise lors d'une consultation (38–40,42). Tous ces filtres ont été construits via une analyse textuelle d'un Gold Standard de références pertinentes pour la médecine générale. Une analyse de fréquence des termes a permis de sélectionner et de construire des équations de recherche à partir des termes les plus fréquents. La validité de l'équation a ensuite été testée en calculant des indicateurs statistiques. Dans ces études, les termes composant l'équation de recherche étaient

similaires. Il s'agissait de termes décrivant le rôle ou le cadre d'exercice du médecin généraliste. Par exemple : soins de santé primaire, médecin généraliste, médecine de famille (38,39). Ces équations nous semblaient trop restrictives pour trouver toute l'information pertinente. Dans notre étude l'idée était de trouver d'autres termes pour l'équation afin de trouver les articles pertinents pour la médecine générale non retrouvés par les filtres existants.

Dans ces études, les critères d'inclusion des références dans le Gold Standard nous paraissaient trop restrictifs. En effet, les lettres, les éditoriaux étaient exclus alors qu'ils peuvent contenir des informations pertinentes pour la médecine générale. Il nous paraissait intéressant d'élargir les critères d'inclusion pour notre étude.

Les termes candidats pour la constitution des équations ont été sélectionnés manuellement, ils devaient correspondre au concept de soins primaires selon les auteurs. Dans ce cas on peut être amené à penser que le filtre reflète indirectement la définition des soins primaires des auteurs. Il nous paraissait intéressant de constituer un filtre de recherche à partir de notre vision de la médecine générale et d'employer une approche purement sémantique. L'idée était d'utiliser uniquement les termes MeSH, les concepts englobés par ces descripteurs pour refléter au mieux notre vision de la médecine générale. Mais aussi de ne pas se restreindre aux termes décrivant une dimension trop partielle de la médecine générale, rôle ou cadre d'exercice, et d'inclure d'autres concepts dans l'équation de recherche.

7. Objectif de l'étude

Les médecins généralistes se posent beaucoup de questions cliniques en raison de leur activité variée. Pour y répondre, ils doivent se baser sur les données validées de la science. Ces données sont produites par la recherche en médecine générale ou en soins primaires. La recherche en médecine générale s'est progressivement développée en quantité et en qualité. Ces données sont présentes dans des bases de données bibliographiques utilisables à l'aide d'outils numériques. La plus utilisée est MEDLINE accessible via PubMed mais son utilisation est complexe et le volume bibliographique à trier est important. Cela rend difficile pour le médecin l'identification des informations pertinentes pour sa pratique. Le filtre de recherche constitue un outil pouvant faciliter l'accès à ces données. Des équipes ont développé des filtres de recherche en médecine générale mais ne répondent pas aux besoins des médecins généralistes. L'objectif de cette étude est de construire un nouveau filtre de recherche afin de permettre au praticien de retrouver les informations pertinentes en médecine générale pour un sujet précis.

Bibliographie

1. Heyrman J, Spreeuwenbergh C, editors. Vocational training in general practice. Leuven: Katholieke Universiteit Leuven; 1987.
2. McWhinney IR. A textbook of family medicine. 2nd ed. New York: Oxford University Press; 1997.
3. Guyatt G, Cairns J, Churchill D, Cook D, Haynes B, Hirsh J, et al. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. *JAMA*. 1992 4;268(17):2420–5.
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996 13;312(7023):71–2.
5. Sackett DL. The Cochrane Collaboration. *ACP J Club*. 1994;120 Suppl 3:A11.
6. Institute of Medicine (US) Committee on the Future of Primary Care. Primary Care: America's Health in a New Era [Internet]. Donaldson MS, Yordy KD, Lohr KN, Vanselow NA, editors. Washington (DC): National Academies Press (US); 1996 [cited 2017 Feb 16]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK232643/>
7. Weel C. General practice: a suitable place for clinical research. *Eur J Gen Pract*. 1995;1:6–7.
8. WONCA, Definition de la medecine generale.
9. Olesen F, Dickinson J, Hjortdahl P. General practice—time for a new definition. *BMJ*. 2000 5;320(7231):354–7.
10. De Pourville, Rapport_Pourville.pdf.
11. Problems in the “Evidence” of “Evidence-Based Medicine”. ResearchGate [Internet]. [cited 2017 Feb 6]; Available from: https://www.researchgate.net/publication/13801263_Problems_in_the_Evidence_of_Evidence-Based_Medicine
12. Schuurs M, Griffon N, Kerdelhue G, Foubert Q, Mercier A, Darmoni SJ. Behavior and attitudes of residents and general practitioners in searching for health information: From intention to practice. *Int J Med Inf*. 2016;89:9–14.
13. White KL, Williams TF, Greenberg BG. The ecology of medical care. *N Engl J Med*. 1961 2;265:885–92.
14. Green LA, Fryer GE, Yawn BP, Lanier D, Dovey SM. The ecology of medical care revisited. *N Engl J Med*. 2001 28;344(26):2021–5.
15. Collard-Siedlecki Clémence, Thèmes et tendances de recherche en médecine générale.
16. Smith J. Primary care: balancing health needs, services and technology. *Int J Integr Care* [Internet]. 2001 [cited 2017 Feb 13];1. Available from:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1484414/>

17. Starfield B. Primary care: an increasingly important contributor to effectiveness, equity, and efficiency of health services. *SESPAS report 2012. Gac Sanit.* 2012;26 Suppl 1:20–6.
18. Thèmes et tendances de recherche en médecine générale : analyse de la littérature extraite de Pubmed.
19. Fact SheetMEDLINE® [Internet]. [cited 2017 Feb 16]. Available from: <https://www.nlm.nih.gov/pubs/factsheets/medline.html>
20. PubMed Help [Internet]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.How_do_I_search_PubMed
21. PubMed [Internet]. Available from: <http://www.biusante.parisdescartes.fr/ressources/pdf/pubmed1-tutoriel-biusante.pdf>
22. Eveillard P, Hannedouche T. Recherche bibliographique médicale avec Medline–Pubmed. Une approche pratique basée sur l'exemple. *Néphrologie Thérapeutique.* 2007 ;3(7):475–85.
23. Clarke M, Greaves L, James S. MeSH terms must be used in Medline searches. *BMJ.* 1997 19;314(7088):1203.
24. Glossaire – CISMef [Internet]. [cited 2017 Feb 5]. Available from: <http://www.chu-rouen.fr/cismef/aide/glossaire/>
25. Fact SheetMedical Subject Headings (MeSH®) [Internet]. [cited 2017 Feb 4]. Available from: <https://www.nlm.nih.gov/pubs/factsheets/mesh.html>
26. Coumou HCH, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. *J Med Libr Assoc.* 2006 ;94(1):55–60.
27. Verhoeven AA, Boerma EJ, Meyboom-de Jong B. Which literature retrieval method is most effective for GPs? *Fam Pract.* 2000 ;17(1):30–5.
28. Wilson P, Droogan J, Glanville J, Watt I, Hardman G. Access to the evidence base from general practice: a survey of general practice staff in Northern and Yorkshire Region. *Qual Health Care QHC.* 2001 ;10(2):83–9.
29. Bernard E, Arnould M, Saint-Lary O, Duhot D, Hebbrecht G. Internet use for information seeking in clinical practice: a cross-sectional survey among French general practitioners. *Int J Med Inf.* 2012 ;81(7):493–9.
30. K. Ramos, R. Linscheid, S. Schafer Real-time information-seeking behavior of residency physicians *Fam. Med.*, 35 (2003), pp. 257–260.
31. E.W. Westberg, A.M. Randolph The basis for using the internet to support the information needs of primary care *J. Am. Med. Inform. Assoc.*, 6 (1) (1999), pp. 6–25..
32. Bennett NL, Casebeer LL, Kristofco R, Collins BC. Family physicians' information seeking behaviors: A survey comparison with other specialties. *BMC Med Inform Decis Mak.*

2005;5:9.

33. E.W. Westberg, A.M. Randolph The basis for using the internet to support the information needs of primary care *J. Am. Med. Inform. Assoc.*, 6 (1) (1999), pp. 6–25.
34. Lancement de LiSSa, base de données de la Littérature Scientifique en Santé – CISMef [Internet]. [cited 2017 Feb 17]. Available from: <http://www.chu-rouen.fr/cismef/2016/01/lancement-de-lissa-base-de-donnees-de-la-litterature-scientifique-en-sante/>
35. CISMef. Constructeur de Requêtes Bibliographiques Médicales [Internet]. [cited 2017 Feb 17]. Available from: <http://crbm.chu-rouen.fr/querybuilder/>
36. Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. *J Med Libr Assoc JMLA*. 2008 ;96(4):356–61.
37. Pols DHJ, Bramer WM, Bindels PJE, Laar FA van de, Bohnen AM. Development and Validation of Search Filters to Identify Articles on Family Medicine in Online Medical Databases. *Ann Fam Med*. 2015 1;13(4):364–6.
38. Brown L, Carne A, Bywood P, McIntyre E, Damarell R, Lawrence M, et al. Facilitating access to evidence: Primary Health Care Search Filter. *Health Inf Libr J*. 2014 ;31(4):293–302.
39. Gill PJ, Roberts NW, Wang KY, Heneghan C. Development of a search filter for identifying studies completed in primary care. *Fam Pract*. 2014 1;31(6):739–45.
40. Sladek R, Tieman J, Fazekas BS, Abernethy AP, Currow DC. Development of a subject search filter to find information relevant to palliative care in the general medical literature. *J Med Libr Assoc*. 2006 94(4):394–401.
41. Sladek RM, Tieman J, Currow DC. Improving search filter development: a study of palliative care literature. *BMC Med Inform Decis Mak*. 2007;7:18.
42. Lung Cancer Search Filters [Internet]. [cited 2017 Feb 5]. Available from: <https://www.caresearch.com.au/caresearch/tabid/1940/Default.aspx>

DEUXIEME PARTIE: ARTICLE

1. INTRODUCTION

La médecine générale a longtemps été définie uniquement autour de sa pratique médicale, centrée sur le médecin. Une vision axée autour du généraliste au centre d'une communauté d'individus dont il répond aux besoins et assure le suivi. La recherche était laissée au profit des hospitaliers et des universitaires (1,2).

L'Evidence-Based Medicine (EBM) a progressivement modifié notre pratique clinique en intégrant la nécessité de consulter les données validées par la recherche pour prendre une décision médicale. Ainsi l'EBM a intégré une dimension scientifique à la pratique clinique. Le paradigme de l'EBM veut qu'une décision médicale soit prise en prenant en compte les données de la recherche, l'expérience du praticien et les choix du patient (3–5). La recherche devient donc indispensable pour développer une médecine générale de haute qualité (3,4,6).

Au début du siècle, des travaux ont montré l'impact positif des soins primaires dans l'amélioration de l'état de santé des populations dans de nombreux pays, renforçant la nécessité d'accentuer le développement de la recherche en médecine générale (7,8). L'impulsion a été suivie au niveau institutionnel. En 2003 la World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA) a fait la promotion de la recherche en médecine générale. Des recommandations ont été émises telles que le développement des conférences et revues de recherche en médecine familiale avec l'amélioration de la représentation de ces revues dans les bases de données bibliographiques ; le développement d'instituts nationaux de recherche et de départements universitaires de médecine familiale ou encore la facilitation des financements de la recherche collaborative internationale en médecine familiale (9–11). On a pu observer une augmentation de la représentation des revues de médecine générale dans les bases de données bibliographiques et le développement d'une filière de recherche universitaire dans de nombreux pays. (12)

Les données validées par la recherche en médecine générale sont disponibles sur internet dans des bases de données bibliographiques. La plus connue, MEDLINE, est accessible via une interface numérique comme PubMed. PubMed contient plus de 26 millions de références de la littérature biomédicale internationale de MEDLINE et environ un million de références sont indexées annuellement (13,14). Cependant les publications en médecine générale sont peu nombreuses et dispersées dans de nombreuses revues rendant leur

accès difficile pour les praticiens (12).

La recherche en ligne d'articles relevant de la médecine générale peut poser des difficultés notamment en terme d'utilisation des outils de recherche (15,16). Le médecin généraliste dispose de peu de temps pour faire le tri parmi ces résultats et trouver ce qui est pertinent (16,17). Il est également peu formé lorsqu'il s'agit de formuler une requête de recherche adéquate en utilisant les mots-clés du thesaurus MeSH (Medical Subject Headings) (18–20).

Il existe des outils pour surmonter certaines de ces difficultés. Par exemple, l'outil terminologique HeTOP (Health Terminology/Ontology Portal ou Portail Terminologique de Santé en français) développé par CISMef (Catalogue et Index des Sites Médicaux de langue Française) permet de trouver les mots-clés MeSH anglais à partir du français ou d'une autre langue (21). Un filtre de recherche permet de faciliter l'accès aux données en sélectionnant les références pertinentes pour un sujet donné parmi les millions de références indexées dans PubMed (22).

Des filtres de recherche ont été élaborés en soins primaires par des équipes australiennes et canadiennes (23,24). Une équipe néerlandaise a développé deux filtres en médecine générale. Un filtre « sensible » pour les chercheurs et un filtre « spécifique » pour les cliniciens (25). Cependant, ces filtres étaient composés de mots-clés décrivant uniquement le médecin généraliste en tant qu'acteur de soins ou selon son mode d'exercice. Les critères d'inclusion pour la constitution du Gold Standard de références nous paraissaient trop restrictifs influençant indirectement la composition de l'équation de recherche. Les auteurs de ces études ont tous effectué une analyse textuelle des références pour sélectionner les termes candidats à la constitution de l'équation. Ils ont également choisi les termes de l'équation de manière subjective. Ces filtres nous paraissaient insuffisants pour trouver une information de qualité. (23–25)

L'objectif principal de cette étude était de développer un filtre de recherche en médecine générale permettant de retrouver le plus grand nombre des références pertinentes.

2. MATERIEL ET METHODE

2.1. Design principal de l'étude

Nous avons créé le filtre en suivant le modèle de Sladek et Haynes et en nous appuyant sur la checklist de l'InterTASK ISSG (26–28). Notre étude se décompose en quatre étapes: création d'un Gold Standard (GS) ; développement des équations de recherches candidates à partir d'une analyse de fréquence; validation du filtre de recherche interne et externe au GS; comparaisons aux filtres existants.

2.2. Constitution du Gold Standard

2.2.1. Constitution de l'échantillon

Nous avons lancé une requête afin d'obtenir toutes les publications parues en 2014 dans une sélection de 8 revues indexées dans PubMed. La requête initiale était la suivante: ((((((("BMJ"[Journal] OR "JAMA"[Journal]) OR "The New England journal of medicine"[Journal]) OR "Annals of internal medicine"[Journal]) OR "Family practice"[Journal]) OR "Family medicine"[Journal]) OR "Canadian family physician Medecin de famille canadien"[Journal]) OR "Australian family physician"[Journal] AND ("2014/01/01"[PDAT] : "2015/01/01"[PDAT])). Les revues ont été choisies par un panel de médecins généralistes et comprenaient 4 revues généralistes (British Medical Journal BMJ, Journal of the American Medical Association JAMA, New England Journal of Medicine NEJM, Annal of internal Medicine) et 4 revues spécialisées en médecine générale (Family Practice, Family Medicine, Canadian Family Physician, Australian family physician). Ces journaux spécialisées avaient la plus grande proportion d'articles relevant de la médecine générale (12). Tous les types de publications ont été inclus.

2.2.2. Analyse manuelle de l'échantillon

L'échantillon a été analysé manuellement par 2 relecteurs (A.G, L.C). Ils ont analysé le titre et les abstracts de chaque article et identifié ceux qui relevaient de la médecine générale. L'intégralité de l'article a été analysée si besoin. Suite à cette analyse, l'échantillon a été randomisé en 2 groupes d'articles. Le groupe 1 « équation », a permis de constituer notre équation de recherche. Le groupe 2 « validation », a permis de tester l'équation. Les articles relevant de la médecine générale constituaient le sous-groupe "MG". Les publications ne relevant pas de la médecine générale constituaient le sous-groupe "Non MG". Dans le

groupe 1 “équation”, le sous-groupe d’articles relevant de la médecine générale constituait notre Gold Standard (GS) de références.

2.3. Développement de l’équation

2.3.1. Sélection des termes MeSH candidats

Les termes MeSH indexés pour chaque référence relevant de la médecine générale dans le groupe 1 ont été extraits. Les termes ont été isolés et classés par ordre de fréquence, les récurrences ont été regroupées. Les termes en rapport avec le type de publication, l’âge, le sexe et le pays ont été exclus. Les 30 termes MeSH les plus fréquents ont été retenus.

2.3.2. Analyse statistique

Nous avons analysé la performance des requêtes lancées dans le groupe 1 avec chacun des 30 termes MeSH. Nous avons calculé des indicateurs statistiques selon la méthode adaptée de Haynes (27)(Tableau 1). Nous avons calculé la sensibilité, la spécificité, la précision et le F-score des requêtes pour chaque terme MeSH candidat afin de déterminer les plus pertinents pour la composition du filtre.

2.3.2.1. *Sensibilité/Rappel*

La sensibilité, également appelée rappel, est le rapport entre le nombre d’articles pertinents retrouvés à l’aide du terme MeSH étudié sur le total des articles pertinents identifiés manuellement (GS).

2.3.2.2. *Spécificité*

La spécificité correspond au rapport entre le nombre d’articles non retrouvés non pertinents par la requête sur le total des articles non pertinents, il s’agit de la probabilité qu’un article soit correctement exclu.

2.3.2.3. Précision

La précision, également appelée valeur prédictive positive, correspond au rapport du nombre d'articles pertinents retrouvés sur le total d'articles retrouvés par la requête.

2.3.2.4. F-score

Le F-score ou F-mesure, est une moyenne harmonique du rappel et de la précision. Le rappel, la précision et le F-score sont des indicateurs souvent utilisés dans la littérature pour évaluer un filtre de recherche.

2.3.2.5. Risque Relatif Medline

Le Risque Relatif Medline, RR, correspond au risque de trouver un article pertinent pour la médecine générale dans le groupe 1 en utilisant une requête avec un terme MeSH donné par rapport au risque dans l'ensemble des publications MEDLINE indexées en 2014.

Les termes MeSH avec un F-Score supérieur à 5% étaient candidats pour la constitution de l'équation. Parmi ces termes, ceux avec un Risque Relatif Medline supérieur à 10 ont été retenus pour la constitution de la première équation. Les termes MeSH étaient associés entre eux par l'opérateur booléen « OU » et la requête était lancée dans le groupe 1. La deuxième équation était composée de MeSH avec un Risque Relatif Medline supérieur à 7. Le rappel, la précision, la spécificité et le F-score ont été calculés pour les deux équations dans le groupe 1.

2.4. Validation interne de l'équation

Les deux équations de recherche ont été testées sur le groupe 2 « validation ». Le rappel, la précision, la spécificité et le F-score ont été calculés.

2.5. Comparaisons aux filtres de recherche existants

Les équations de recherche de notre étude ont été comparées aux filtres de recherche d'autres études. Les équipes de Pols & al. et Gill & al. ont chacune constitué un filtre de recherche en médecine générale sensible et un autre spécifique (23,24). L'équipe de Brown & al. a élaboré un filtre de recherche équilibré (25). Le rappel, la précision, le F-score et la spécificité de ces filtres dans le groupe 2 ont été calculés.

3. RESULTATS

3.1. Gold Standard

La requête initiale nous a permis d'obtenir un échantillon de 7615 références. 7602 articles ont été analysés manuellement par les deux reviewers. 994 références ont été identifiées pertinentes pour la médecine générale. 6608 références ne relevaient pas de la médecine générale. Pour 13 articles, l'analyse du titre et de l'abstract n'a pas permis de statuer sur la pertinence pour la médecine générale et il n'a pas été possible d'accéder à l'article intégral. Ils ont donc été exclus. Les trois journaux généralistes réunissaient 84% des articles identifiés pertinents pour la médecine générale. Les articles pertinents représentaient 13,08% de l'échantillon de 7602 références (Tableau 2).

Après randomisation, le groupe 1 comprenait 3802 articles dont 512 relevaient de la médecine générale. Le groupe 2 comprenait 3800 articles dont 482 étaient pertinents pour la médecine générale. Les 512 articles traitant de médecine générale du groupe 1 (sous-groupe "MG" du groupe 1) constituaient notre Gold Standard (Figure 1).

3.2. Termes MeSH

11698 termes MeSH ont été extraits des bordereaux d'indexation des 994 articles relevant de la médecine générale. On trouvait dans le groupe 1, 6099 termes soit 1469 termes différents, après regroupement des termes identiques (« exact match »). La liste des 30 termes MeSH les plus fréquents figure dans le tableau 3. Ils représentaient 8,7% du total des termes MeSH du groupe 1. Les 13 termes MeSH qui avaient un F-score > 5% et ont été retenus comme candidats pour la constitution du filtre de recherche. Les termes candidats représentaient 4,4% de l'effectif total des 6099 termes.

Parmi ces termes, 8 décrivaient la médecine générale (personnel soignant, organisation de soins, élément professionnel): médecine de famille[MH] ; médecins[MH] ; médecine générale[MH] ; soins de santé primaires[MH] ; relations médecin-patient[MH] ; politique de santé[MH] ; assurance maladie[MH]. Les 5 autres descripteurs faisaient référence à des pathologies : maladies cardiovasculaires[MH] ; tabagisme[MH] ; grossesse[MH] ; infarctus du myocarde[MH] ; hypertension artérielle[MH].

3.3. Equations de recherche

L'équation 1 était composée de 5 termes MeSH avec un risque relatif MEDLINE (RR) supérieur à 10 parmi les 13 termes candidats. Ils étaient les suivants : médecine de famille[MH] (RR=43,59) ; médecine générale[MH] (RR=27,52) ; relations médecin-patient[MH] (RR=12,45) ; médecins[MH] (RR=11,76) ; assurance maladie[MH] (RR=10,51) (Tableau 4). Dans le groupe 1, le filtre de recherche obtenait un rappel à 18,16%, une précision à 15,15%, un F-score à 16,52% et une spécificité à 84,16%. Lors de la validation dans le groupe 2, il obtenait un rappel à 13,28%, une précision à 11,49%, un F-score à 12,32% et une spécificité à 85% (Tableau 5).

L'équation 2 était composée des 8 termes MeSH avec un risque relatif MEDLINE supérieur à 7. Elle était composée des 5 termes de la première équation et des termes suivants : politique de santé[MH] (RR=9,29) ; tabagisme[MH] (RR=7,59) ; soins de santé primaires[MH] (RR=7,10)(Tableau 4). Dans le groupe 1, le filtre de recherche obtenait un rappel à 28,91%, une précision à 17,19%, un F-score à 21,56% et une spécificité à 78,33%. Lors de la validation dans le groupe 2, il obtenait un rappel à 22,61%, une précision à 13,16%, un F-score à 16,64% et une spécificité à 78% (Tableau 5).

L'équation 3 était composée des 13 termes candidats et présentait un rappel à 52,12%, une spécificité à 56,72% et un F-score à 24,24% dans le groupe 1. Lorsque la requête était exécutée dans le groupe 2, elle obtenait un rappel à 49,38%, une précision à 13,59%, un F-score à 21,32% et une spécificité à 54,40% (Tableau 5).

3.4. Comparatif des filtres de recherche

Utilisé dans le groupe 2, le filtre de Brown & al. avait un rappel à 11%, un F-score à 11,22% et une spécificité à 88%. Les filtres « sensibles » de Pols & al. et Gill & al. ont respectivement obtenu un rappel à 29,88% et 22,61%, un F-score à 16,67% et 15,77% et une spécificité à 67% et 76%. Les filtres « spécifiques » ont obtenu un rappel à 13,49% et 2,28%, un F-score à 12,18% et 3,69% et une spécificité à 84% et 97% (Tableau 5).

4. DISCUSSION

4.1. Principaux résultats

Nous avons développé et validé un filtre de recherche en médecine générale pour PubMed à réalisant une analyse sémantique à partir d'un Gold Standard de 512 références en médecine générale sélectionnées manuellement parmi tous les articles indexés en 2014 issus d'une sélection de 8 revues. L'équation 1 composée de 5 termes obtenait un rappel à 13,28%, une précision à 11,49%, un F-score à 12,32% et une spécificité à 85%. L'équation 2 composée de 8 termes obtenait un rappel à 22,61%, une précision à 13,16%, un F-score à 16,64% et une spécificité à 78%. L'objectif était de réaliser un filtre de recherche équilibré pour le rappel et la spécificité. Au final nous avons obtenu 2 filtres de recherche. L'équation 1 a obtenu des valeurs similaires au filtre spécifique en médecine générale de Pols & al. et au filtre équilibré pour les soins primaires de Brown & al. pour le F-score et la spécificité. L'équation 2 a obtenu des valeurs similaires au filtre sensible de Pols & al. et au filtre sensible pour les soins primaires de Gill & al (23-25).

La sensibilité de nos filtres plafonnait à environ 50% avec l'équation comprenant tous les termes MeSH candidats (équation 3). Ce résultat pouvait être expliqué par le fait que la liste des termes candidats ne représentait que 4,4% de l'ensemble des termes MeSH extraits du groupe 1. Ainsi on peut penser qu'ajouter des termes à la liste et dans la requête permettrait d'augmenter la sensibilité.

Par ailleurs, nous avons considéré que les 13 références (sur un total de 76150) exclues lors de la constitution du GS, par défaut d'accessibilité au texte intégral, n'auraient pas eu d'impact sur les termes MeSH sélectionnés et par conséquent les équations finales.

4.2. Méthodologie

Notre avons constitué notre échantillon d'articles de départ en incluant toutes les publications de l'année 2014 dans une sélection de revues généralistes indexées dans PubMed. Notre avons émis l'hypothèse qu'il existait des publications relevant de la médecine générale dans les publications autres que les essais, méta-analyses et les revues de littérature (17,29). Ainsi, les lettres, les commentaires, les « clinical reviews », les éditoriaux ou d'autres types de publications issus des revues sélectionnées ont été incluses

dans notre étude. Les autres études avaient exclu ces types de publications pour la constitution de l'échantillon de départ. Nous avons obtenu des résultats similaires aux autres études pour nos filtres de recherche. Nous manquons de données concernant les caractéristiques des références selon le type de publications.. Il serait intéressant d'analyser ces données afin d'évaluer leur impact dans la constitution d'un filtre de médecine générale.

Nous avons fait le choix de réaliser une analyse sémantique pour identifier les termes candidats pour la constitution de l'équation de recherche par opposition à l'analyse textuelle des autres études. Les auteurs des autres filtres de recherche ont analysé le « text word » en plus du MeSH pour identifier les termes candidats. La recherche via « text word » permet de rechercher l'information dans des données qui ne sont pas encore indexées ou en cours d'indexation, sont utilisation est plus libre (utilisation de la troncature et des astérisques) (30). Elle augmente la sensibilité mais au détriment du bruit en retrouvant plus de faux positifs (31). Se focaliser uniquement sur les termes MeSH a probablement permis de réduire le bruit.

L'analyse manuelle de l'échantillon de départ à été réalisée par deux relecteurs. Ils avaient une vision semblable de la définition de la médecine générale et de son champ d'action. Cela pouvait constituer un biais de sélection dans la constitution de notre échantillon. Constituer un panel représentatif des médecins généralistes, panachant cliniciens et chercheurs pourrait contribuer à limiter ce biais, de même que la définition de critères d'inclusion plus précis pour les références concernant la médecine générale.

Une des forces de notre étude réside dans la démarche rigoureuse de notre méthodologie. Cette démarche avait pour but de minimiser les décisions subjectives afin d'être reproductible. Nous avons suivi les étapes méthodologiques pour la constitution d'un filtre de recherche éditée par la InterTASK ISSG (26). Nous avons utilisé des indicateurs statistiques pour la sélection de nos termes candidats afin d'être le plus rigoureux possible.

Nous avons fait le choix d'utiliser des indicateurs statistiques validés pour évaluer la performance de notre outil de recherche d'information (32). Nous avons utilisé la F-score

pour sélectionner les termes candidats. C'est une mesure combinant le rappel et la précision qui a été validée par les programmes internationaux d'évaluation (32). Nous avons utilisé une mesure qui était absente dans les autres études, le risque relatif d'habitude utilisé dans les études épidémiologiques (33). Il s'agissait plus précisément du risque de trouver un article pertinent en médecine générale en utilisant un mot-clé donné, dans le GS, par rapport au risque de trouver une référence pertinente dans l'ensemble des publications MEDLINE de 2014. Un terme avec un RR élevé était a priori plus spécifique pour la médecine générale. Dans notre étude, les termes aux RR les plus élevés ont été sélectionnés pour la constitution des filtres et ont obtenu des performances semblables aux autres études. Il semblerait que le risque relatif soit un bon indicateur pour évaluer la spécificité.

Dans la littérature, les filtres spécifiques sont développés pour aider le clinicien à trouver une réponse à une question précise. Les filtres sensibles vont théoriquement plus intéresser les chercheurs (22,34). Peu d'information est disponible dans la littérature concernant l'utilisation par les médecins généralistes de ces filtres de recherche. Les différentes études autour des filtres de recherche en soins primaires et en médecine générale n'ont pas permis d'élaborer un filtre unique, optimal, utilisable par tous (24,25). Il est possible que notre filtre de recherche reste peu utilisé par le généraliste ou le chercheur en médecine générale. A contrario, il pourrait être utile pour la veille bibliographique des références en médecine générale ou à un documentaliste pour aider à la constitution des bases de données bibliographiques en médecine générale.

4.3. Perspectives

Définir la médecine générale par des termes MeSH constitue un challenge. Les termes MeSH existants nous paraissaient trop restrictifs pour définir des aspects spécifiques de la médecine générale telles que la dimension sociale ou la transversalité des situations cliniques rencontrées. Ainsi des travaux au sur le plan sémantique pourraient être réalisés afin d'améliorer la définition de la discipline par les descripteurs MeSH.

Des travaux complémentaires semblent nécessaires pour créer un outil utilisable par les médecins généralistes pour retrouver la littérature pertinente. Des efforts sont notamment à réaliser dans la formation des médecins à la recherche d'informations bibliographiques. Concernant les filtres de recherches existants, ils peuvent être utilisés par les documentalistes pour constituer des bases de données. Le filtre que nous avons développé

pourrait notamment être utilisé dans la base bibliographiques en français en cours de création, LiSSA (35)

BIBLIOGRAPHIE

1. Heyrman J, Spreeuwenbergh C, editors. Vocational training in general practice. Leuven: Katholieke Universiteit Leuven; 1987.
2. McWhinney IR. A textbook of family medicine. 2nd ed. New York: Oxford University Press; 1997.
3. Guyatt G, Cairns J, Churchill D, Cook D, Haynes B, Hirsh J, et al. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. JAMA. 1992 4;268(17):2420–5.
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996 13;312(7023):71–2.
5. Sackett DL. The Cochrane Collaboration. ACP J Club. 1994;120 Suppl 3:A11.
6. Weel C. General practice: a suitable place for clinical research. Eur J Gen Pract. 1995;1:6–7.
7. Starfield B. Primary care: an increasingly important contributor to effectiveness, equity, and efficiency of health services. SESPAS report 2012. Gac Sanit. 2012;26 Suppl 1:20–6.
8. Smith J. Primary care: balancing health needs, services and technology. Int J Integr Care [Internet]. 2001 1 [cited 2017 Feb 13];1. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1484414/>
9. De Pouvourville Rapport_Pouvourville.pdf.
10. van Weel C, Rosser WW. Improving Health Care Globally: A Critical Review of the Necessity of Family Medicine Research and Recommendations to Build Research Capacity. Ann Fam Med. 2004;2(Suppl 2):s5–16.
11. Pimlott N, Upshur REG. From clinical observation to clinical discovery. Can Fam Physician. 2014 ;60(1):14–6.
12. Collard-Siedlecki C Thèmes et tendances de recherche en médecine générale : analyse de la littérature extraite de Pubmed.
13. Fact SheetMEDLINE® [Internet]. [cited 2017 Feb 16]. Available from: <https://www.nlm.nih.gov/pubs/factsheets/medline.html>
14. PubMed Help [Internet]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.How_do_I_search_PubMed
15. Schuurs M, Griffon N, Kerdelhue G, Foubert Q, Mercier A, Darmoni SJ. Behavior and attitudes of residents and general practitioners in searching for health information: From intention to practice. Int J Med Inf. 2016;89:9–14.
16. Bernard E, Arnould M, Saint-Lary O, Duhot D, Hebbrecht G. Internet use for

information seeking in clinical practice: a cross-sectional survey among French general practitioners. *Int J Med Inf.* 2012;81(7):493–9.

17. Alper BS, Hand JA, Elliott SG, Kinkade S, Hauan MJ, Onion DK, et al. How much effort is needed to keep up with the literature relevant for primary care? *J Med Libr Assoc.* 2004;92(4):429–37.

18. Younger P. Internet-based information-seeking behaviour amongst doctors and nurses: a short review of the literature. *Health Inf Libr J.* 2010;27(1):2–10.

19. Coumou HCH, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. *J Med Libr Assoc.* 2006;94(1):55–60.

20. Ely JW, Osheroff JA, Ebell MH, Chambliss ML, Vinson DC, Stevermer JJ, et al. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ.* 2002 23;324(7339):710.

21. CISMeF. HeTOP [Internet]. [cited 2017 Feb 17]. Available from: <http://www.hetop.eu/hetop/>

22. Jenkins M. Evaluation of methodological search filters--a review. *Health Inf Libr J.* 2004;21(3):148–63.

23. Brown L, Carne A, Bywood P, McIntyre E, Damarell R, Lawrence M, et al. Facilitating access to evidence: Primary Health Care Search Filter. *Health Inf Libr J.* 2014;31(4):293–302.

24. Gill PJ, Roberts NW, Wang KY, Heneghan C. Development of a search filter for identifying studies completed in primary care. *Fam Pract.* 2014 1;31(6):739–45.

25. Pols DHJ, Bramer WM, Bindels PJE, Laar FA van de, Bohnen AM. Development and Validation of Search Filters to Identify Articles on Family Medicine in Online Medical Databases. *Ann Fam Med.* 2015 1;13(4):364–6.

26. Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. *J Med Libr Assoc JMLA.* 2008 96(4):356–61.

27. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc.* 1994;1(6):447–58.

28. Sladek R, Tieman J, Fazekas BS, Abernethy AP, Currow DC. Development of a subject search filter to find information relevant to palliative care in the general medical literature. *J Med Libr Assoc.* 2006 ;94(4):394–401.

29. Wentz R. Visibility of research: FUTON bias. *Lancet Lond Engl.* 2002 19;360(9341):1256.

30. Changing Vocabularies: A Guide to Help Bioethics Searchers Find Relevant Literature in National Library of Medicine Databases Using the Medical Subject Headings

(MeSH) Indexing Vocabulary [Internet]. Available from:
<https://bioethics.georgetown.edu/publications/changingvocab.pdf>

31. Petrova M, Sutcliffe P, Fulford KWM (Bill), Dale J. Search terms and a validated brief search filter to retrieve publications on health-related values in Medline: a word frequency analysis study. *J Am Med Inform Assoc JAMIA*. 2012;19(3):479–88.

32. Analyse des criteres d'evaluations des systemes de recherche d'information [Internet]. Available from: https://www.irit.fr/publis/SIG/2010_ISI_BADSKNDMJ.pdf

33. jauzein. Différentes façons d'exprimer un résultat d'étude épidémiologique — Acces [Internet]. [cited 2017 Feb 18]. Available from: http://acces.ens-lyon.fr/acces/ressources/sante/epidemiologie/niveau_preuve/expression_risque

34. Hoogendam A, de Vries Robbé PF, Stalenhoef AFH, Overbeke AJPM. Evaluation of PubMed filters used for evidence-based searching: validation using relative recall. *J Med Libr Assoc JMLA*. 2009;97(3):186–93.

35. LiSSA [Internet]. Available from: <http://www.lissa.fr/dc/#lang=fr&env=lissa>

FIGURES ET TABLEAUX

Requete lancée sur PubMed	MGoui	MGnon
Article retrouvé	a	b
Article non retrouvé	c	d

Tableau 1. Sensibilité, spécificité, précision et F-score

Sensibilité = $a/(a + c)$; Spécificité = $d/(b + d)$; Précision = $a/(a + b)$;

F-score = $2a/(2a + b + c)$ (27)

Titre du journal	Relevant de la médecine générale				
	MGoui	MGnon	Total	MGoui/TotalMGoui	MGoui/Total
Annals of internal medicine	54	615	669	5%	0,71%
Australian family physician	0	248	248	0%	0,00%
BMJ	417	2796	3213	42%	5,49%
Canadian family physician	55	263	318	6%	0,72%
Family medicine	13	162	175	1%	0,17%
Family practice	34	84	118	3%	0,45%
JAMA	178	1097	1275	18%	2,34%
The New England journal of medicine	243	1343	1586	24%	3,20%
Total	994	6608	7602	100%	13,08%

Tableau 2. Articles relevant de la médecine générale issus de 8 revues pour la constitution du GS

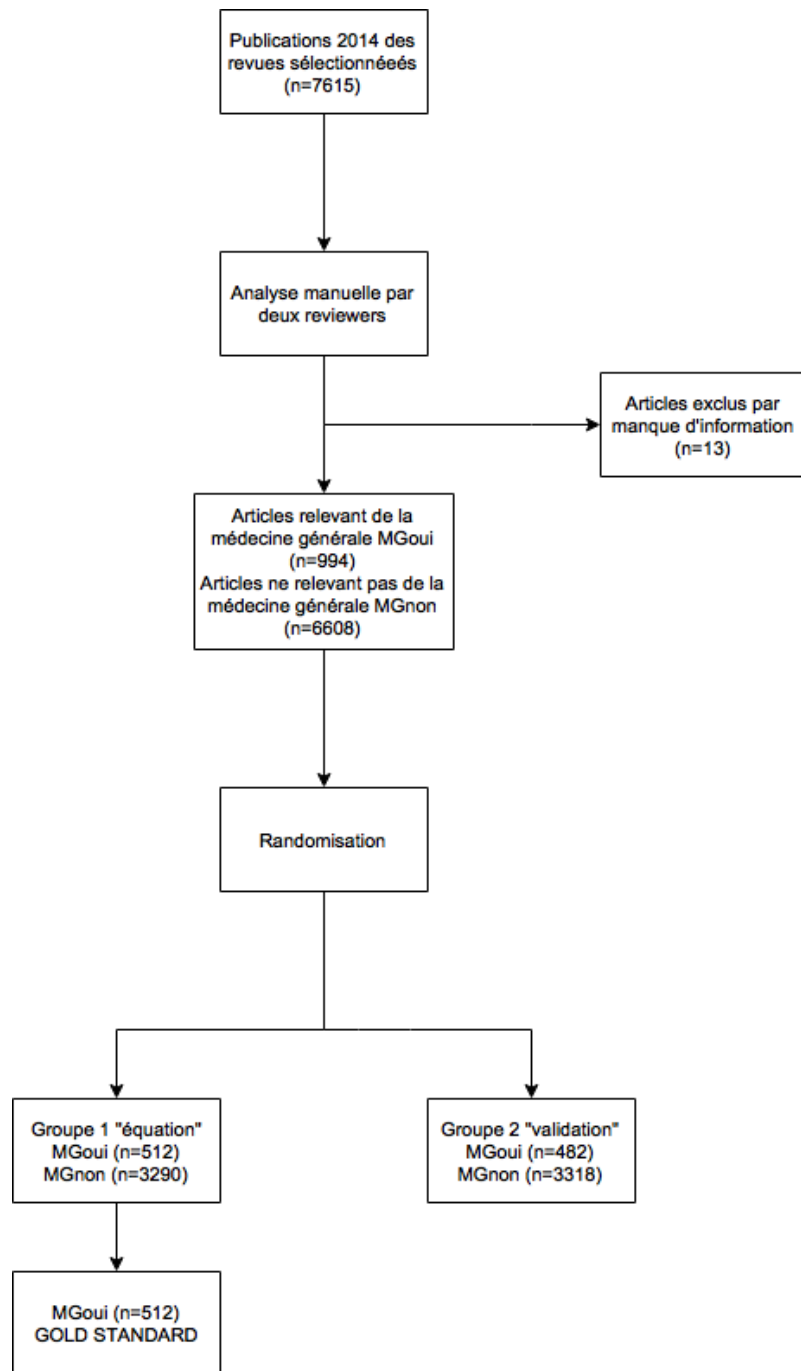


Figure 3. Constitution du Gold Standard

Termes MeSH	Effectif
tabagisme	44
maladies cardiovasculaires	30
grossesse	29
médecine d'État	26
hypertension artérielle	26
médecins	24
antihypertenseurs	24
médecine de famille	22
tumeurs du sein	21
infarctus du myocarde	19
politique de santé	19
relations médecin-patient	18
soins de santé primaires	17
prestations des soins de santé	16
antiviraux	16
diabète de type 2	15
marqueurs biologiques	14
anticorps monoclonaux humanisés	14
antibactériens	14
évaluation des risques	13
dépistage systématique	13
assurance maladie	13
grippe humaine	13
médecine générale	13
diagnostic différentiel	13
prévalence	12
rôle médical	12
medicare (USA)	12
hypoglycémiants	12
promotion de la santé	12
maladie d'Alzheimer	12

Tableau 3. 30 termes MeSH les plus fréquents issus du GS

Terme MeSH	Sensibilité	Précision	F-score	Spécificité	RR
maladies	16,02%	16,17%	16,09%	87,08%	1,72
cardiovasculaires					
prestations des soins de santé	16,99%	13,43%	15,00%	82,95%	3,20
tabagisme	5,47%	40,00%	9,62%	98,72%	7,59
soins de santé primaires	6,25%	16,93%	9,13%	95,23%	7,10
grossesse	5,66%	20,71%	8,90%	96,63%	2,04
médecins	6,25%	15,38%	8,89%	94,65%	11,76
assurance maladie	4,88%	21,74%	7,97%	97,26%	10,51
politique de santé	4,69%	22,02%	7,73%	97,42%	9,29
médecine générale	5,66%	10,94%	7,46%	92,83%	27,52
infarctus du myocarde	3,32%	27,42%	5,92%	98,63%	5,56
médecine de famille	3,52%	13,04%	5,54%	96,35%	43,59
relations médecin-patient	3,32%	15,04%	5,44%	97,08%	12,45
hypertension artérielle	2,93%	34,09%	5,40%	99,12%	3,60

Tableau 4. Sensibilité, précision, F-score, spécificité et risque relatif pour les 30 termes MeSH les plus fréquent du GS

Requête lancée sur PubMed	groupe 1 "développement"				groupe 2 "validation"			
	Sensibilité	Précision	F-score	Spécificité	Sensibilité	Précision	F-score	Spécificité
1. médecins[MH] OU assurance maladie[MH] OU médecine générale[MH] OU médecine de famille[MH] OU relations médecins-patient[MH]	18,16%	15,15%	16,52%	84,16%	13,28%	11,49%	12,32%	85%
2. médecins[MH] OU assurance maladie[MH] OU tabagisme[MH] OU soins de santé primaires[MH] OU médecine générale[MH] OU politique de santé[MH] OU médecine de famille[MH] OU relations médecins-patient[MH]	28,91%	17,19%	21,56%	78,33%	22,61%	13,16%	16,64%	78%
3. maladies cardiovasculaires[MH] OU prestations de soins de santé[MH] OU tabagisme[MH] OU soins de santé primaires[MH] OU grossesse[MH] OU médecins[MH] OU assurance maladie[MH] OU politique de santé[MH] OU médecine générale[MH] OU infarctus du myocarde[MH] OU médecine de famille[MH] OU relations médecin-patient[MH] OU hypertension artérielle[MH]	52,15%	15,79%	24,24%	56,72%	49,38%	13,59%	21,32%	54,40%

Tableau 5. Validation interne et externe des équations de recherche

Requête lancée sur PubMed	groupe 2 "validation"			
	Sensibilité	Précision	F-score	Spécificité
1. médecins[MH] OU assurance maladie[MH] OU médecine générale[MH] OU médecine de famille[MH] OU relations médecins-patient[MH]	13,28%	11,49%	12,32%	85%
2. médecins[MH] OU assurance maladie[MH] OU tabagisme[MH] OU soins de santé primaires[MH] OU médecine générale[MH] OU politique de santé[MH] OU médecine de famille[MH] OU relations médecins-patient[MH]	22,61%	13,16%	16,64%	78%
3. maladies cardiovasculaires[MH] OU prestations de soins de santé[MH] OU tabagisme[MH] OU soins de santé primaires[MH] OU grossesse[MH] OU médecins[MH] OU assurance maladie[MH] OU politique de santé[MH] OU médecine générale[MH] OU infarctus du myocarde[MH] OU médecine de famille[MH] OU relations médecin-patient[MH] OU hypertension artérielle[MH]	49,38%	13,59%	21,32%	54,40%
Pols & al. FILTRE SENSIBLE ("family"[all elds] OR physician*[all elds] OR practice*[tw] OR "primary care"[all elds] OR "Primary Health Care"[mh] OR primary[tw] OR general pract*[tiab] OR gp[tiab] OR gps[tiab])	29,88%	11,56%	16,67%	67%
Pols & al. FILTRE SPECIFIQUE ("Primary Health Care"[mh] OR "primary care"[all elds] OR "Physicians, Family"[mh] OR general pract*[all elds] OR "family"[ad] OR family pract*[all elds] OR family physician*[tw])	13,49%	11,11%	12,18%	84%
Brown & al. (((Primary care[tw] OR General practi*[tw] OR Primary health care[tw] OR Community mental health services[mh:noexp] OR Family practice[mh:noexp] OR Home care services[mh:noexp] OR Family physicians[mh:noexp] OR Community health services[mh:noexp] OR Community health nursing[mh:noexp] OR Community pharmacy services[mh:noexp] OR Community health workers[mh:noexp] OR Preventive health services[mh:noexp]) AND Medline[sb]) OR ((Primary care[tiab] OR General practi*[tiab] OR Primary health*[tiab] OR Community mental health*[tiab] OR Family practice[tiab] OR Family medicine[tiab] OR Family physician*[tiab] OR Home care[tiab] OR Home based[tiab] OR Home health*[tiab] OR Community health*[tiab] OR Community nurs*[tiab] OR health visit*[tiab] OR Community pharmac*[tiab] OR Preventive care[tiab] OR Prevention program*[tiab] OR Preventive service*[tiab] OR Preventive health[tiab] OR Health promotion[tiab]) NOT Medline[sb]) AND English[la])	11,00%	11,45%	11,22%	88%
Gill \$ al. FILTRE SENSIBLE Clinic*[tiab] OR practi*[tiab] OR primary[tiab] OR physician*[tiab] OR refer*[tiab] OR visit*[tiab] OR outpatient*[tiab] OR consult*[tiab] OR family[tiab] OR communit*[tiab] OR ambulatory[tiab] OR centre*[tiab] OR centre*[tiab] OR of office[tiab]	22,61%	12,11%	15,77%	76%
Gill \$ al. FILTRE SPECIFIQUE general practi*[tiab] OR primary health*[tiab]	2,28%	9,57%	3,69%	97%

Tableau 6. Comparaison des filtres de recherche

Résumé

Introduction : Les médecins généralistes rencontrent des difficultés lorsqu'il s'agit de rechercher une information pertinente dans MEDLINE via PubMed. L'objectif de cette étude était d'élaborer un filtre de recherche en médecine générale sur PubMed afin de faciliter la recherche d'information pertinente pour la médecine générale.

Méthode : Nous avons construit un Gold Standard (GS) de référence en médecine générale à partir d'une analyse manuelle d'un échantillon de références indexées en 2014 issues de 8 revues. Une analyse de fréquence des termes MeSH et une analyse statistique ont permis d'identifier des candidats pour la constitution du filtre de recherche. Des combinaisons ont été testées sur le GS et validées dans un deuxième groupe de références. Les résultats ont ensuite été comparés aux filtres existants.

Résultats : 7602 articles ont été analysés manuellement. Le GS était composé de 512 références pertinentes pour la médecine générale. 6099 termes MeSH différents ont été extraits du groupe 1. 13 MeSH avaient un F-score supérieur à 5% ont été sélectionnés parmi les 30 plus fréquents. L'équation 1, composée des 5 termes ayant un risque relatif MEDLINE (RR) supérieur à 10, obtenait un rappel à 13,28%, une précision à 11,49%, un F-score à 12,32% et une spécificité à 85%. L'équation 2, composée des 8 termes ayant un RR supérieur à 7, obtenait un rappel à 22,61%, une précision à 13,16%, un F-score à 16,64% et une spécificité à 78%. Les filtres de notre étude avait des scores similaires aux filtres existants.

Conclusion : Nous avons élaboré deux filtres de recherche en médecine générale sur PubMed, l'un spécifique, l'autre plus sensible. En pratique, il semble difficile d'obtenir un filtre optimal et utilisable par le généraliste. Son utilisation pourrait dépasser le champ de la médecine générale notamment dans le domaine de l'informatique médicale.

Mots-clés : Medical Subject Headings, Médecine générale, Médecine factuelle, mémorisation et recherche des informations

Abstract

Introduction : Searching for relevant information through MEDLINE via PubMed can be challenging for general practitioners. This study aimed to elaborate a general practice search filter for PubMed to facilitate access to relevant information.

Methods : We conducted a search in MEDLINE via PubMed for articles published in 2014 in 8 journals. We manually identified a Gold Standard (GS) set of references relevant to general practice. The GS was used to identify frequently occurring MeSH terms through semantic analysis. Search strategies were developed from selected MeSH terms and internally validated. We evaluated the filters performance against primary care et family medicine search filters from other studies.

Results : 7602 were manually reviewed and 512 articles relevant for general practice from group 1 were identified as the GS. 6099 MeSH terms were extracted from the GS. MeSH terms with a F-score above 5% were selected. Within these 13 MeSH terms, we selected those with a relative risk above 10 for the development of a search filter. For the 5 MeSH terms search strategy : specificity peaked at 85% with 12.32% F-score, 13.28% sensitivity and 11.49% precision. A second search filter was developed with MeSH terms which presented a relative risk above 7. The second search filter peaked at 78% for the specificity and 22.61% for the sensitivity, 13.16% for the precision and a 16.64% F-score. Comparisons with filters from others studies showed minimal difference.

Conclusion : We created search strategies for PubMed to retrieve relevant information for general practice in order to ease online literature searching. However we reckon our filter was not optimal for clinicians use nor researchers in family medicine. Perhaps it could be useful in the field of health informatics

Keywords : Medical Subject Headings, General Practice, EBM, Information storage and retrieval