3-31-2011 12:00 AM

The characteristics of successful evidence retrieval by nephrologists and the impact of PubMed search filters

Salimah Shariff
University of Western Ontario

Supervisor
Dr. Amit Garg
The University of Western Ontario Joint Supervisor
Dr. Mark Speechley
The University of Western Ontario

Graduate Program in Epidemiology and Biostatistics
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

THE CHARACTERISTICS OF SUCCESSFUL EVIDENCE RETRIEVAL BY NEPHROLOGISTS AND THE IMPACT OF PUBMED SEARCH FILTERS

(Spine title: Evaluating Evidence Retrieval by Nephrologists)

(Thesis format: Monograph)

by

Salimah Shariff

Graduate Program in Epidemiology and Biostatistics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Co-supervisors	Examiners
Dr. Amit Garg	Dr. Allan Donner
Dr. Mark Speechley	Dr. Daniel Hackam
	Dr. Sonny Cejic
	Dr. Scott Brimble

The thesis by

Salimah Shariff

entitled:

THE CHARACTERISTICS OF SUCCESSFUL EVIDENCE RETRIEVAL BY NEPHROLOGISTS AND THE IMPACT OF PUBMED SEARCH FILTERS

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date	Chair of the Thesis Examination Board

ABSTRACT AND KEYWORDS

Background: Clinicians frequently search PubMed to guide patient care. This study investigated the factors that impact successful searching and the utility of PubMed search filters.

Methods: A random sample of nephrologists was surveyed between 2008 and 2010; 160 valid responses were received (72% response rate). One group of 60 respondents was presented with the same two clinical questions, while the other 100 were each presented with a unique clinical question. The clinical questions were based on recently published systematic reviews. Respondents provided the search terms they would type into PubMed to address their clinical question(s). All physician-provided searches were executed in PubMed and outcome measures of sensitivity (proportion of relevant articles found) and precision (proportion of all articles found that are relevant) were calculated. Primary studies included in the reviews served as the reference standards of relevant articles. For the first group of respondents, the associations between the search query or nephrologist characteristics and search outcomes were investigated through multivariable regression modeling. For the second group, three types of filters were applied to the physician-provided searches: one designed to identify high quality studies about treatment (‘methods’), one designed to identify studies with renal content (‘content’) and one designed to limit searches to journals that publish renal evidence (‘journal’). Search outcomes of the non-filtered and filter-aided searches were compared using the Wilcoxon Signed-Rank test.

Results: Multifaceted searching (e.g. using MeSH, limits) improved sensitivity (rate ratio[RR]:2.6; 95% CI:1.4-5.0) and precision (RR:2.0; 95% CI:1.3-3.3). The addition of concept terms decreased sensitivity (RR:0.7; 95% CI:0.5-0.9), while increasing precision (RR:1.6; 95% CI:1.3-2.0). No associations were evident between nephrologist characteristics and sensitivity. However, physicians who previously received training in literature searching produced searches with better precision (RR:2.3; 95% CI:1.4-3.6). The combined use of the ‘methods’ and ‘content’ filters produced the largest improvement in precision with no change in sensitivity, compared to non-filtered searches (median difference:5.5%; 95% CI:2%-12%).

Conclusions: Use of multifaceted searching and filters can improve physician-provided searches in PubMed. Literature training curricula should adopt the findings from this study. Improved search performance has the potential to enhance clinical practice and improve patient care.

Keywords: Evidence-based Medicine, PubMed Searching, Information Retrieval, Search Filters, Sensitivity, Precision

DEDICATION

To my mother, father & brother:

you are my true teachers.

ACKNOWLEDGEMENTS

Supervisory Committee

I would like to thank my advisory committee for sharing their expertise, and guiding me over the past few years: Dr. Amit Garg (supervisor), Dr. Mark Speechley (co-supervisor), Dr. Amardeep Thind, Dr. R. Brian Haynes and Dr. Ann McKibbon. When I first met with Dr. Garg, he cautioned me to accept only if I was up for a challenge; I hope that I successfully rose to that challenge. Amit, thanks for being my greatest advocate. I have learned much from you.

Funding Support

I am grateful for the various sources of funding that have allowed me to focus on my research. My thanks go to the Canadian Institutes for Health Research (CIHR), the Ontario Ministry of Training, Colleges and Universities, and the University of Western Ontario Schulich School of Medicine & Dentistry. The CIHR also supported a part of this thesis in the form of an Operating Grant.

Research Collaborators

This thesis would not have been possible without the coordinated efforts of numerous collaborators at the Kidney Clinical Research Unit & Health Information Research Unit. I would like to thank Arthur Iansavichus, Dr. Nancy Wilczynski, Meaghan Cuerden, Dr. Robert Yang, Dr. Arsh Jain, Dr. Matthew Weir, Darisuz Gozdzik, Lihua Li, Heather Thiessen-Phillbrook, Shayna Bejaimal, and Stephen Woo. Special thanks to Dr. Jessica Sontrop, who I not only consider a role-model and mentor, but a friend. Thanks for listening, advising, and most importantly, encouraging.

Other Individuals

No production is possible without the ‘behind the scenes’ individuals. Thanks to the “girls” of the KCRU for keeping me grounded. Many thanks, also, to the wonderful administrative support provided by Jessica Burgess, Robin Ljunggren, Mai Trac and Theresa Hands over the years.

I have developed some wonderful friendships during my time in London; thanks to you all for being my family and home away from home.

Finally, to my immediate and extended family, thank you for your tireless love and support. To my late aunt, your resilience and optimism through adversity is my constant inspiration.

Shukran Lillah Wal Hamdulillah.

TABLE OF CONTENTS

CERTIFICATE OF EXAMINATION ii

ABSTRACT AND KEYWORDS iii

DEDICATION v

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS viii

LIST OF TABLES xii

LIST OF FIGURES xiv

LIST OF APPENDICES xv

LIST OF ABBREVIATIONS, SYMBOLS, NOMENCLATURE xvii

CHAPTER 1: INTRODUCTION 1

1.1 Definitions of key terminology 2

1.2 Background and overview 5

CHAPTER 2: LITERATURE REVIEW 6

2.1 Principles of evidence-based medicine 7

2.1.1 PICO 7

2.2 Increase in the amount of medical literature 8

2.3 About MEDLINE and PubMed 8

2.4 Current practice of literature acquisition by physicians 9

2.4.1 Common sources used 9

2.4.2 Use of MEDLINE by physicians 9

2.4.3 Choice of success measure 13

2.5 Evidence needs differ for specialists compared to general physicians 13

2.6 Challenges of searching electronic bibliographic databases 14

2.7 A solution to improve PubMed search query performance: filters 15

2.7.1 Filter testing framework 16

2.8 Understanding searching with ‘AND’ and ‘OR’ Boolean logic operators and the tradeoff between sensitivity and precision 17
2.9 Limitations of existing studies evaluating electronic databases 19
2.9.1 Use of pre-specified sets of relevant articles 19
2.9.2 Factors associated with success in searching 19
2.9.3 Evaluation studies of MEDLINE filters on end-user searching 20
2.9.4 Sample size 20
2.9.5 Studies targeted at specialists 20
2.10 Conclusion 21

CHAPTER 3: STUDY OBJECTIVES 22

3.1 Study objectives and hypotheses 23
3.1.1 Objective 1: Determinants of search success 23
3.1.2 Objective 2: Impact of search filters on search query performance 24
3.2 Conceptual model 25

CHAPTER 4: STUDY METHODS 26

4.1 Study design and sampling 27
4.1.1 Assembled sample of treatment questions and relevant articles 27
4.1.2 Surveyed nephrologists 30
4.1.3 Performed searches 36
4.1.4 Compared responders to non-responders 36
4.2 Objective 1: Determinants of search success 37
4.2.1 Study variables 37
4.2.2 Analytic strategy 42
4.2.3 Sample size 44
4.3 Objective 2: Impact of search filters on search query performance 46
4.3.1 Study variables 46
4.3.2 Analytic strategy 47
4.3.3 Sample size 51
4.4 Methods used to minimize potential threats to validity 52

CHAPTER 5: RESULTS **54**
5.1 Survey results 55
5.1.1 Sample 55
5.1.2 Data management and cleaning 56
5.1.3 Respondent characteristics 57
5.1.4 Use of online sources 58
5.2: Objective 1: Determinants of search success 59
5.2.1 Main study variables 59
5.2.2 Relationship between search query characteristics and search success 62
5.2.3 Relationship between nephrologist characteristics and search success 65
5.2.4 Additional analyses 68
5.3 Objective 2: Impact of search filters on search query performance 69
5.3.1 Primary and Secondary analyses: Using unmodified search queries 69
5.3.2 Additional analyses: Using modified search queries 80
CHAPTER 6: DISCUSSION & CONCLUSIONS **81**
6.1 Summary of principal findings 82
6.1.1 Objective 1: Determinants of search success 82
6.1.2 Objective 2: Impact of search filters on search query performance 83
6.2 Contribution of this work to the current literature 84
6.3 Recognized limitations 87
6.3.1 Determining article relevancy 87
6.3.2 Performance metrics 87
6.3.3 Searching is a dynamic process 88
6.3.4 Target audience is nephrologists 89
6.3.5 Self-administered survey 89
6.3.6 Non-responder bias 90
6.3.7 Statistical modeling decisions 91
6.4 Future directions 91
6.5 Study significance 93
6.5.1 Implications for searching physicians 93

6.5.2 Implications for developers of bibliographic databases 93
6.5.3 Implications for literature training workshops 94
6.6 Conclusions 95
CHAPTER 7: BIBLIOGRAPHY **96**
CHAPTER 8: APPENDICES **107**
CHAPTER 9: CURRICULUM VITAE **243**

LIST OF TABLES

CHAPTER 2: LITERATURE REVIEW

Table 2-1: Studies evaluating the impact of PubMed filters on search performance..... 11

Table 2-2: Factors associated with search success 12

Table 2-3: Formulae for calculating search Sensitivity and search Precision................ 15

Table 2-4: Search filter testing framework.. 17

CHAPTER 4: STUDY METHODS

Table 4-1: Summary of survey groups 33

Table 4-2: Methods to overcome missing responses or discrepancies........................... 34

Table 4-3: List of study variables for Objective 1... 40

Table 4-4: References to appropriate appendices for regression models 45

Table 4-5(a-d): Regression models and variables considered for model inclusion........ 45

Table 4-6: Filters available for testing.. 46

Table 4-7: Rules for syntactically improving physician-provided search queries 51

CHAPTER 5: RESULTS

Table 5-1: Characteristics of responding and non-responding nephrologists 55

Table 5-2: Number of missing values or discrepancies for survey responses for Group 1 56

Table 5-3: Respondent demographics and their characteristics separated by the two groups of surveys 57

Table 5-4: Outcome variables for Question 1 (used for models 1-4)............................ 60

Table 5-5: Outcome variables for Question 2 (used for models 5-8)............................ 60

Table 5-6: Primary predictors and potential confounders for use in models 1 & 2 61

Table 5-7: Primary predictors and potential confounders for use in models 5 & 6 61

Table 5-8: Categorical predictors for use in models 3, 5, 7, 8 62

Table 5-9: Continuous predictors for use in models 3, 5, 7, 8 62

Table 5-10: Results from negative binomial regression analysis assessing the effect of search query characteristics on sensitivity for question 1 63

Table 5-11: Results from the negative binomial regression analysis assessing the effect of search query characteristics on sensitivity for question 2 63
Table 5-12: Results from the negative binomial regression analysis assessing the effect of search query characteristics on precision for question 1 64
Table 5-13: Results from the negative binomial regression analysis assessing the effect of search query characteristics on precision for question 2 64
Table 5-14: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on sensitivity for question 1 66
Table 5-15: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on sensitivity for question 2 66
Table 5-16: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on search precision for question 1 67
Table 5-17: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on precision for question 2 67
Table 5-18: Search performance of physician-provided searches and searches aided by filters 71
Table 5-19: Change in search performance between filtered and non-filtered physician-provided searches 73
Table 5-20: Search performance of physician-provided searches and searches aided by filters, when restricted to the top 40 returned results 76
Table 5-21: Change in search performance between filtered and non-filtered physician-provided searches, when restricted to the top 40 returned results 78

LIST OF FIGURES

CHAPTER 2: LITERATURE REVIEW

Figure 2-1: Model for evidence-based clinical decisions 1

Figure 2-2: Performance of a diagnostic tool 1

Figure 2-3: Searching with Boolean logic operators 18

CHAPTER 3: STUDY OBJECTIVES

Figure 3-1: Conceptual model 25

CHAPTER 4: STUDY METHODS

Figure 4-1: Process of including renal therapy reviews 29

CHAPTER 5: RESULTS

Figure 5-1: Response to survey question: "when you search a bibliographic database, how many results do you generally scan per search?" 58

LIST OF APPENDICES

Appendix 1: Snapshot of PubMed Clinical Queries 108
Appendix 2: Procedure for determining whether an article contains renal information 109
Appendix 3: Details of included systematic reviews 115
Appendix 4: Survey of nephrologists 132
Appendix 5: Research ethics approval 139
Appendix 6: Sample-size calculation for Objective 1 - Determinants of search success 140
Appendix 7: MODEL #1; Predictor: Search Query Characteristics; Outcome: Sensitivity; Question 1 144
Appendix 8: MODEL #2; Predictor: Search Query Characteristics; Outcome: Precision; Question 1 152
Appendix 9: MODEL #3; Predictor: Nephrologist Characteristics; Outcome: Sensitivity; Question 1 162
Appendix 10: MODEL #4; Predictor: Nephrologist Characteristics; Outcome: Precision; Question 1 170
Appendix 11: MODEL #5; Predictor: Search Query Characteristics; Outcome: Sensitivity; Question 2 176
Appendix 12: MODEL #6; Predictor: Search Query Characteristics; Outcome: Precision; Question 2 182
Appendix 13: MODEL #7; Predictor: Nephrologist Characteristics; Outcome: Sensitivity; Question 2 190
Appendix 14: MODEL #8; Predictor: Nephrologist Characteristics; Outcome: Precision; Question 1 195
Appendix 15: Pilot data used to calculate standard deviation values 203
Appendix 16: Sample-size calculations for Objective 2 – Impact of search filters on search query performance 207
Appendix 17: IndividualPower SAS macro to evaluate power for multiple comparisons using Dunnett’s method 209

Appendix 18: Sample-size calculations for Objective 2 using Dunnett's method........ 212
Appendix 19: Objective 2 - Histograms of the differences in search sensitivity and precision between filtered and non-filtered searches 213
Appendix 20: SAS code used for Objective 2 225
Appendix 21: Objective 2 - Details of search results (all results) 229
Appendix 22: Objective 2 - Details of search results (when restricting the results to the top 40 citations) 231
Appendix 23: Objective 2 - Details of additional analyses (using modified search queries) 233

LIST OF ABBREVIATIONS, SYMBOLS, NOMENCLATURE

CI – Confidence Interval

EBM – Evidence Based Medicine

NLM – National Library of Medicine

NNR – Number needed to read

PICO – Patient, Problem, Population; Intervention; Comparison, Control; Outcome

MeSH– Medical Subject Heading

PMID – PubMed Unique Identifier

CHAPTER 1: INTRODUCTION

1.1 Definitions of key terminology

Article or Citation - A listing about a publication that includes the title, names of authors, name of journal, date of publication or other publication information which allows the researcher to locate the item.

Boolean - Boolean is a logic system. Using the "AND" operator between terms retrieves documents containing both terms. "OR" retrieves documents containing either term. "NOT" excludes the retrieval of terms from the search.

Broad search filter – A search filter that is designed to maximize the sensitivity of a search for a particular search topic.

Clinical Practice Guideline - A document with the aim of guiding decisions and criteria regarding diagnosis, management, and treatment in specific areas of healthcare. They are often based on an examination of current evidence within the paradigm of evidence-based medicine.

Clinical Queries - Specialized PubMed search filter intended for clinicians. Limits the retrieval to articles that report research conducted with specific methodologies.

Concept Terms - A word or group of words that embody the most specific clinical aspect used in a search query combined with an ‘AND’ operator (implicit or explicit)

Electronic Bibliographic Database – Electronic index to bibliographic journal articles, containing citations, abstracts and often either the full text of the articles, or links to the full text.

Electronic Database - A structured collection of information that can be retrieved via a computer system.

Evidence based medicine - The integration of best research evidence with clinical expertise to aid in the diagnosis and management of patients.

Free text words – When a search term is entered into PubMed as a text word, a search is performed on all fields such as the title, abstract, MeSH terms, MeSH Subheadings, Publication Types.

Limits – When conducting a search, the ability to focus on specific data fields such as study types, age groups and gender, which may result in retrieval that is more relevant.

MEDLINE (Medical Literature Analysis and Retrieval System Online) - The US National Library of Medicine's® (NLM) bibliographic database that contains over 18 million references to journal articles in life sciences with a concentration on biomedicine.

MeSH (Medical Subject Heading) - The NLM controlled vocabulary thesaurus of indexing terms. Search queries entered into PubMed are automatically mapped to MeSH vocabulary when a match is found. By default, PubMed automatically searches the MeSH headings as well as more specific terms beneath the heading in the MeSH hierarchy (known as **explosion**). Many MeSH terms also include **subheadings** (such as prevention, control, analyses).

Multifaceted search features – These features include Boolean logic, truncation, limits, or controlled vocabulary.

Multifaceted search query – A query that includes multifaceted search features.

Narrow search filter - A search filter that is designed to maximize the specificity of a search for a particular search topic.

Number Needed to Read - The average number of non-relevant articles retrieved per relevant article retrieved (the inverse of precision).

Primary studies - Studies used in systematic reviews that meet strict inclusion and exclusion criteria, as stated in each review. Primary studies include original data on samples or individuals.

Precision (also known as 'positive predictive value') - The proportion of relevant articles retrieved of all articles retrieved (the inverse of number needed to read).

PubMed – Free search engine for accessing the MEDLINE database of citations. Includes journal articles, reports and commentaries on life sciences and biomedical topics. Sometimes citations include abstracts and links to full-text articles. Can be accessed from www.pubmed.gov.

Retrieved - Articles/citations returned for a search.

Search - An execution of a search query.

Search Filter – An optimized search query (consisting of a combination of Boolean logic operators, truncations, medical subject headings [MeSH], subject heading explosions, free-floating subheadings, and free text words) that when applied to a search will return a limited subset of the database enriched with relevant material for a specific search topic. Also referred to as a 'hedge'.

Search Query/Search String - The keywords, key phrases, or list of words that are typed into a search box to find citations on a topic of interest.

Sensitivity - The proportion of relevant articles retrieved of all relevant articles. In the information science literature this is also known as 'recall'.

Systematic Review - A literature review focused on a single question which tries to identify, appraise, select and synthesize all high quality research evidence relevant to that question. Most quality indicators of systematic reviews suggest that a systematic review should include the search strategies and processes by which studies or articles were identified and the inclusion and exclusion criteria for article inclusion in addition to the single focused question.

Truncation (also referred to as **wildcard**) **-** The use of a symbol to search only part of a term to retrieve variant endings of that term. The truncation symbol in PubMed is "*". Truncations can be either single or multiple letters.

1.2 Background and overview

Physicians are essential members in the delivery of health services, and are expected to offer their patients care that is based on current, best evidence from health care research. Retrieving health literature is a cornerstone of evidence-based practice[1]. PubMed is the most widely-used repository of health literature by physicians: In 2009, it is estimated that 15% of the 1.3 billion searches in PubMed were conducted by clinicians[2;3]. Unfortunately, many clinicians fail to retrieve relevant articles when they perform a search and at the same time retrieve large numbers of non-relevant articles. In addition, the factors that lead to successful searches are still largely unknown. A proposed solution to improve physician searching is PubMed filters. Filters are pre-tested searches optimized to improve the accuracy of retrieving articles for a given purpose. The use of filters is akin to screening for disease in high-risk populations. By filtering PubMed, rather than searching the entire database, clinician searches are performed on a subset of articles where relevant information is more likely to be found[4;5].

Three types of PubMed filters can be used by physicians when searching in the field of renal medicine. The first filter, ‘content’, was designed to help clinicians find clinical content for renal medicine[2]. When clinicians select this filter, they no longer need to type renal terms into their search; rather their search is executed in a subset of PubMed articles related to renal diseases. The second filter, Clinical Queries (‘methods’), was designed to help clinicians find articles of high methodological quality for clinical questions of therapy, diagnosis, prognosis and etiology[6-10]. For example, when “therapy” is selected, the filter retrieves articles of randomized controlled trials. The third filter, ‘journal’, limits searches to a subset of journals where renal practice evidence has been published[11]. Clinicians can use all three filters alone or in combination when searching PubMed.

To remedy the challenge of evidence retrieval, this thesis focused on two objectives. The first objective identified the determinants of search success when nephrologists search PubMed for articles on focused renal therapy questions. The second objective examined whether filters (used alone, or in combination) can enhance the performance of searches created by clinicians.

CHAPTER 2: LITERATURE REVIEW

2.1 Principles of evidence-based medicine

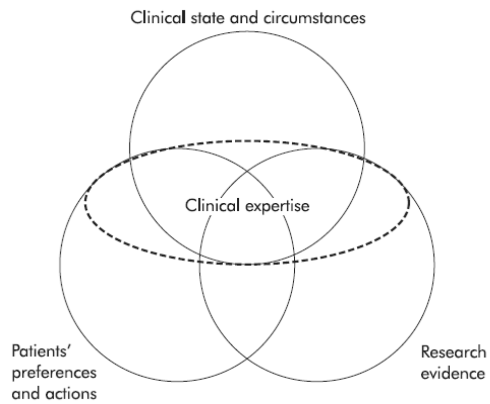


Figure 2-1: Model for evidence-based clinical decisions[1]

Physicians are encouraged to follow the process of evidence-based medicine (EBM) in their medical practice. This paradigm emphasizes the use of research evidence integrated with clinician expertise and patient preferences when practicing medicine[12]. At times, physicians require additional information to assist with patient care. In such situations, for the successful delivery of healthcare, EBM outlines a process of five steps[13]: 1) convert the information need into an answerable clinical question, 2) search for best evidence to answer the question, 3) critically appraise and interpret the evidence, 4) combine the appraisal with clinical expertise and patient preferences to apply the evidence, and 5) evaluate the first four steps and identify methods for improvement. One of the barriers widely documented in the practice of EBM is the difficulty in successfully executing the second step in the process[14-16]. When searching for literature, physicians are encouraged to identify and interpret studies of the highest methodological rigour that are best suited for the question they are trying to answer[13;17].

2.1.1 PICO

To enhance literature searching, physicians are encouraged to formulate their clinical questions using PICO (steps 1 and 2 of the EBM framework) [13;18;19]. This acronym represents the different facets of a clinical question – the **P**atient or **P**roblem or **P**opulation being addressed, the **I**ntervention or exposure being considered, the **C**omparison intervention or **C**ontrol group and when relevant, and the clinical **O**utcomes of interest. By breaking down a clinical question into these different areas, it is believed that physicians can more accurately search for pertinent literature, often-times including each category in PICO to develop a search query.

2.2 Increase in the amount of medical literature

The amount of useful knowledge continues to grow, and is greater than any one practitioner can easily retain. From the years 2000 to 2010, the MEDLINE database grew by 7.5 million citations, to 18.3 million citations[20]. The conclusion that medical professionals have unmet medical information needs is inescapable[14-17;21-23]. At times, physicians are unaware of new clinically relevant information. As a result, physicians report the need for supplementary information for an average of two of three patient encounters in clinical settings[24-26]. Unfortunately, physicians find it difficult to search for answers to clinical questions. Many questions that arise in practice go unanswered[14;16;21;23;27-29]. In its most severe form, this can undermine patient safety and the quality of care a patient receives[30-33].

2.3 About MEDLINE and PubMed

The MEDLINE database was introduced to the medical community in the late 1960s. This service indexes journal articles in life sciences with the concentration in biomedicine. In an effort to improve access, in 1997 the National Library of Medicine (NLM) introduced PubMed (www.pubmed.gov). This portal to MEDLINE is freely available through the World Wide Web. Every article in PubMed is indexed based on its title, abstract, author(s), journal name, language of publication and year of publication. All included articles are also manually annotated by personnel at the NLM who assign Medical Subject Headings to each article (acronym MeSH: a controlled hierarchical vocabulary that covers a wide range of medical and scientific topics).

When a search is performed in PubMed, query terms are first mapped to MeSH terms through a process referred to as ‘query preprocessing’. Following this, the service then performs an actual search attempting to match the extended query to articles in the database. Results are presented to the user in reverse-chronological order from the date they were entered into the database (which translates to roughly the date of publication). PubMed also allows for the use of multifaceted search queries that include Boolean logic operators (‘AND’, ‘OR’, ‘NOT’), the use of truncation, and limiting the search by data fields such as study types, age groups and gender. MEDLINE / PubMed is now the most widely used and accepted repository of medical literature with 1.3 billion searches conducted in 2009[3].

2.4 Current practice of literature acquisition by physicians

2.4.1 Common sources used

A systematic review by Dawes and Sampson identified the primary sources of information used by physicians to access medical knowledge. The sources include books, colleagues, journals and electronic bibliographic databases[34]. Unfortunately, all these sources have their challenges in providing current, best evidence. Many textbooks are outdated by the time they are printed[35]. Colleagues frequently have the same challenge keeping up to date as the physician asking the question[13;36;37]. Best evidence may be widely dispersed across journals that are not typically reviewed. For example, articles relevant to the care of renal patients are published across 466 journals in over 18 different disciplines[11]. For these reasons, physicians are increasingly turning to electronic bibliographic databases, such as MEDLINE, as a way to track down medical information[38-40]. However, a review of information seeking-behaviours by physicians identified two prominent challenges when using MEDLINE: lack of time and limited search skills[14-17]. Outside of clinical practice, health professionals spend, on average, half an hour per search topic to find, read and critically appraise retrieved literature[30;41]. In truth, in practice, physicians only have time to spend an average of 2 minutes or less to find the literature they need[21;42;43].

2.4.2 Use of MEDLINE by physicians

In 2009, approximately 1.3 billion searches were conducted in PubMed. It has been estimated that 15% of PubMed searches were conducted by clinicians (personal communication, U.S NLM staff) [2;3]. Reports documenting clinicians use of MEDLINE vary in their design, types of clinicians considered (medical trainees, healthcare professionals, family physicians, specialists physicians), and the types of MEDLINE interfaces used (e.g., Grateful Med, MEDIS, OVID, PubMed). These studies can be grouped into four categories based on their objectives:

1) *To determine whether searches in MEDLINE are able to identify relevant literature*[30;41;44-53]. Overall, the studies show that physicians vary in their ability to identify relevant literature in MEDLINE. Two of these studies evaluated the impact of PubMed filters on search performance and are summarized in Table 2-1.

2) *To determine whether clinical questions can be answered using MEDLINE searches*[15;41;54-58]. At times, the use of MEDLINE can assist in answering clinical questions. However, there are cases where the use of electronic bibliographic databases causes a physician to change correct answers to those which are incorrect[57;58].

3) *To determine whether training sessions on literature searching can improve search success*[45-49;51;54;59;60] . In general, point estimates indicate that training sessions may improve the ability to answer questions and retrieve relevant literature. A systematic review evaluating the effect of information training skills workshops concluded that there "was limited evidence to show that training improves skills, insufficient evidence to determine the most effective methods of training and limited evidence to show that training improves patient care" [60].

4) To *identify factors associated with success in searching*[41;45;50;55;60-62]. Four definitions of search success have been used in those studies where factors associated with success were examined. These factors are summarized in Table 2-2.

Table 2-1: Studies evaluating the impact of PubMed filters on search performance

Study	Target population	Filters tested	No. of search queries	Reference standard used	Who created the queries	Primary findings
Yousefi-Nooraie et al. 2010[53]	Physicians	Searches were run using the native web-based PubMed interface. Each of the listed limits/filters was added separately to each search query. • PubMed Limits of 'randomized controlled trial' and 'clinical trial'[¥] • Clinical Queries specific filter • Clinical Queries sensitive filter	100 search queries for 100 clinical questions	Primary articles from systematic reviews	Researchers	• CQ-sensitive had highest sensitivity and lowest precision • Use of Intervention name in query significantly improved sensitivity and precision • No. of terms did not impact search outcomes
Schardt et al. 2007[45]	Physicians	Users were randomized to one of three PubMed interfaces: *A.* Available on handheld device, interface included fields for PICO, Age group, Gender & Publication type *B.* Same as *A* with addition of type of question (using the Clinical Queries filter - user chose sensitive or specific) *C.* Web-based, native PubMed	10 search queries for each of 3 clinical questions per protocol (total of 30 queries per protocol)	3 researches selected what was considered relevant from the results retrieved	Interns and residents from an inpatient general medicine rotation	Interfaces *A* and *B* exhibited higher precision for each question than interface *C*

[¥]PubMed syntax for filter: Clinical Trial[ptyp] OR Randomized Controlled Trial[ptyp]

Table 2-2: Factors associated with search success

Definition of Success	Factors Identified
Ability to answer a question correctly[41]	• Questions targeted to therapy, diagnosis, harm or prognosis vs. other types of questions • Knowledge of the correct answer before searching • User experience with MEDLINE • Spatial visualization
Ability to answer questions in a timely fashion[55]	• Use of multifaceted search features • Use of the 'related articles' feature • Spell check
Ability to retrieve relevant literature[45;50;53;60]	• Previous training in literature searching • Search experience • Use of Intervention name in search query • Use of a PICO framework when searching
Likelihood of viewing an abstract[61;62]	• Query consisted of 4 or more terms • Search that retrieved less than 161 articles

2.4.3 Choice of success measure

While a number of definitions of search 'success' have been used in evaluating the use of MEDLINE, the most prominent and important outcome for physicians is the ability to identify an answer, rooted in evidence, for a clinical question of interest. Intuitively, this outcome of successfully answering a question is dependent on two steps[13]. First, physicians need to be able to identify 'evidence' by retrieving articles that are both scientifically sound and relevant to the health problem they are trying to solve. Physicians then need to be able to critically appraise and interpret the evidence to successfully determine the answer to their question. When a physician is unable to answer a question correctly or is unsatisfied with the search, it may be due to a failure at either step. Thus, to better understand the process of literature searching by physicians, this staged program of research focused on the first step: the ability to retrieve clinically and scientifically relevant literature.

2.5 Evidence needs differ for specialists compared to general physicians

The principles of evidence-based medicine strongly recommend the use of systematic reviews and guidelines to answer clinical questions[63]. However, most clinical questions have not been the subject of a systematic review. In 2003 the Cochrane Collaboration had published fewer than 3200 systematic reviews, but estimated that about 10,000 systematic reviews were needed just to cover questions of therapy[64]. Not only are more new reviews needed, an increasing number of reviews must be regularly updated if they are to continue to be useful. The use of guidelines in medicine has been met with increased criticism, as some guidelines were influenced by large corporations and based on less rigorous evidence, including expert opinion[65-67]. For example, in nephrology, guideline use is a controversial issue[67-72] with one publication speculating that they have done more harm than good[67]. Similar to systematic reviews, guidelines become obsolete unless they are continuously updated[73].

A survey of 2400 randomly selected US physicians conducted in 2003 identified that a larger proportion of specialists perform literature searches compared to general practitioners (74% vs. 61%)[74]. While reviews and guidelines may be acceptable for general medical practitioners (and advocated for evidence based practice), their use is frequently seen as inadequate for specialists. Specialists are 'experts' in their field and other medical practitioners rely on them as a definitive source of knowledge[75]. Specialists need to be well versed in the best and most up-to-date primary evidence in their area of expertise. In addition to their role as educators, specialists are also expected to apply this knowledge in a tailored way to meet each individual patient's circumstances, rather than using a 'cookbook' type approach.

2.6 Challenges of searching electronic bibliographic databases

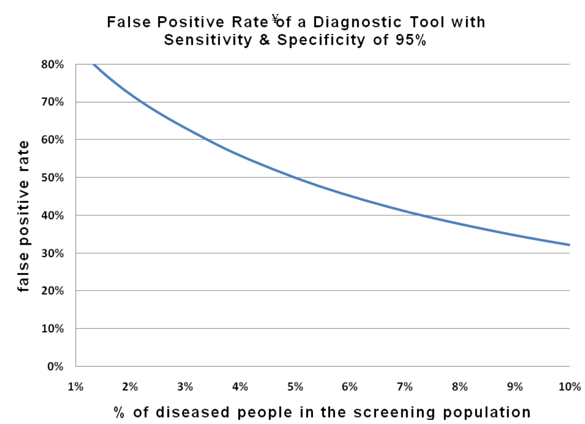


Figure 2-2: Performance of a diagnostic tool

¥False Positive Rate: Proportion of individuals with a positive test who do not have the disease = (number of false positives) / (true positive + false positive)

Searching for relevant articles amongst large quantities of literature is akin to screening for rare diseases in populations. Even with an excellent screening tool with high sensitivity (ability to produce a positive test among people with disease) and high specificity (ability to produce a negative test among people without disease), screening a population in which the number of diseased individuals is low will result in identifying many false positives (a positive test for people without disease) (Figure 2-2). To curtail such findings, in clinical practice, screening of this nature is conducted on high-risk groups and not the entire population. For example, mammograms and colonoscopy procedures are often limited to higher risk individuals over the age of 50. Using lessons learned from clinical practice, a potential solution to improve performance of searches is to search portions of the bibliographic databases where relevant material is more likely to be present. A promising way to achieve this is to use filters, which 'select' potentially relevant content and 'weed-out' unwanted information, leaving a higher concentration of relevant articles for searching.

2.7 A solution to improve PubMed search query performance: filters

The two most prominent metrics to assess the retrieval of information through searching are sensitivity and precision[76] (also called recall and 'positive predictive value'). These measures concentrate on evaluating the proportion of relevant items found. Relevance has loosely been defined as a retrieved document satisfying an information need as specified by a search query[77]. Sensitivity refers to the proportion of all relevant articles that are retrieved, while precision refers to the proportion of articles retrieved by the search that are deemed relevant (Table 2-3).

Table 2-3: Formulae for calculating search Sensitivity and search Precision

		Relevant article	*Non-relevant article*
Search Query	*Articles found*	a (True Positive)	b (False Positive)
	Articles not found	c (False Negative)	d (True Negative)

Sensitivity: a/(a+c), proportion of relevant articles found by the search query
Precision: a/(a+b), proportion of articles found by the search query that are relevant

In an attempt to improve these two metrics for clinical users, PubMed filters have been developed to enhance searching[5-10;34;78-91]. By selecting a filter for use, a clinical user would no longer be searching the entire PubMed repository; rather they would be searching within a set of articles enriched for what they were looking for. Filters are, in essence, search strings optimized to retrieve all articles in PubMed for a given purpose (different purposes described below). To develop a filter, terms are combined in various ways and formats using a systematic approach, and performance is measured[5]. The terms make special use of features provided by PubMed, such as Boolean logic operators, truncations, MeSH terms, subject heading explosions, free-floating subheadings and free text words. Sometimes over a million PubMed filters have been tested to find the one that optimizes performance for a given purpose. Often the filters are developed in subsets of the whole database to provide a more reasonable work space[5-10;34;78-91]. When the filter development process is complete, often two forms of the filters are presented: a 'broad' filter and a 'narrow' filter. The broad filters are designed to use a more inclusive approach to find relevant articles (i.e. to increase sensitivity), whereas narrow filters optimize the exclusion of non-relevant articles, thereby increasing precision.

Three types of PubMed filters have been previously developed that can be used to improve searches for renal medical practice evidence[2;7;11]: ‘methods’, ‘content’ and ‘journals’ filters. Testing was done by comparing filter performance against a hand search where assistants rated the relevance of each article. The first type of filter identifies articles of high methodological rigor (for the prevention or treatment of health disorders, diagnostic tests, prognosis, etiology of disease and so on), independent of any clinical discipline[7] (‘methods’ filter). The best performing methods filters, including the most sensitive filter and most specific filter, are a part of the PubMed interface, and can be accessed through the Clinical Queries section (Appendix 1). The second type of filter identifies articles relevant to the practice of renal medicine[2] (‘content’ filter). Broad and narrow filters have been recently developed for this purpose. The third type of filter identifies a list of journals where renal practice evidence has been published[11] (renal ‘journal’ filter). Of the 5375 journals in PubMed[92], 451 have published at least one article directly relevant to the care of renal patients. Each of these filters reduces the PubMed database to sets of articles where information of interest is most likely to be present. For example, applying one of the renal ‘content’ filters to PubMed reduces the number of citations from over 20 million citations to 466,319 (when applied January 12, 2011). Given their theoretical promise, these PubMed filters now require further evaluation to determine if they can enhance physician searching.

2.7.1 Filter testing framework

A testing framework using key recommendations from reviews of electronic search databases and search filters is presented in Table 2-4 [5;93].To date, researchers have developed, optimized and validated the identified filters in closed, experimental environments (stage one and two). The next stage is to determine if these PubMed filters improve real physician searches (stage three). This is a focus of this thesis. The impact of filters used in combination can be investigated and has never been attempted before. Physician information management has the impact to improve if filters can maximize the number of relevant articles retrieved (increase sensitivity), and minimize the number of non-relevant articles retrieved (increase precision). If the filters operate well at this stage of evaluation, then future studies will be justified to test the effect of these filters on physician knowledge, medical decisions and patient outcomes.

Table 2-4: Search filter testing framework.

Development	Stage one	Promising search filters are developed through a rigorous process of combining terms in various ways. The relevance of each article in a set of articles is defined by some reference standard. The ability of a filter to restrict the set of articles to those that are relevant is then considered.
Validation	Stage two	Promising filters are independently evaluated on a second, distinct, set of articles to ensure equivalent performance in replication.
Physician search query performance	Stage three	Determine whether search filters improve end-user search query performance (i.e. sensitivity and precision).
Physician knowledge	Stage four	Determine whether search filters improve physician knowledge.
Medical decisions or care	Stage five	Determine whether the acquired knowledge changes medical decision making or processes of care.
Patient outcomes	Stage six	Determine whether patient outcomes are improved.

2.8 Understanding searching with 'AND' and 'OR' Boolean logic operators and the tradeoff between sensitivity and precision

Most searching situations can be conceptualized in the form of Venn diagrams as demonstrated in Figure 2-3. When terms in a search query are combined with an 'AND' operator, this results in limiting the results to articles that include all concepts. On the other hand, when terms are included with an 'OR' operator, this produces results containing articles that include any one of the concepts. A search with an 'AND' term can be thought of narrowing the results, while those that include an 'OR' term broaden the search results. Consequently, a broader search has the potential to reveal more relevant articles (and thus increases search sensitivity), but often also results in more non-relevant articles being retrieved (and thus decreases search precision). Alternately, a narrower search has the potential to retrieve fewer non-relevant articles (and thus increases search precision), however this might also miss some relevant articles (and thus decreases search sensitivity). This inverse relationship between sensitivity and precision is known as the sensitivity-precision tradeoff.[94;95] It should be noted that the addition of ‘AND’ or ‘OR’ operators do not always result in a true inverse relationship, at times, precision can be increased with no change to sensitivity and vice versa. This is the goal of searching, to maximize both precision and sensitivity.

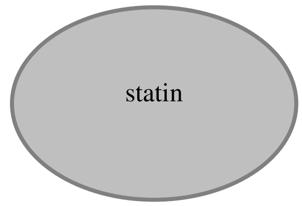
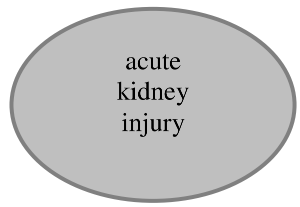
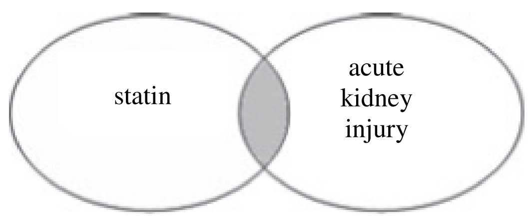
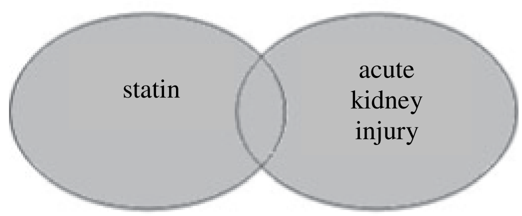
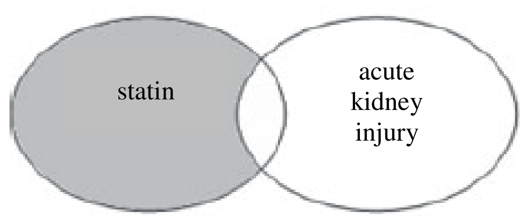
Boolean logic operator	**Search query**	**Venn diagram**	**Description**
No operator	statin		All articles on statin use
No operator	acute kidney injury		All articles on acute kidney injury
AND	statin AND acute kidney injury		All articles on statin use in acute kidney injury
OR	statin OR acute kidney injury		All articles on statin use, plus all articles on acute kidney injury
NOT	statin NOT acute kidney injury		All articles on statin use, except those on statin use in acute kidney injury

Figure 2-3: Searching with Boolean logic operators

2.9 Limitations of existing studies evaluating electronic databases

2.9.1 Use of pre-specified sets of relevant articles

The relevance-based measures of electronic database effectiveness, sensitivity and precision, are well established and have been used widely in evaluating bibliographic databases[76]. To calculate these metrics, a definition of relevance is required for pertinent literature. Most often, a defined set of relevant articles is not specified prior to conducting an experiment. In such cases, relevant articles are identified after all query results have been retrieved and evaluators are required to make judgments of relevance for each identified article. While evaluators sometimes define criteria for relevance and use kappa statistics to quantify the similarity of their relevance judgments, relevance continues to remain a subjective measure[96]. Additionally, this method cannot be used to quantify whether relevant material was missed by the searches (the search metric sensitivity cannot be calculated). To avoid these difficulties, a pre-specified set of relevant documents needs to identified prior to testing the retrieval of electronic bibliographic databases[97;98].

2.9.2 Factors associated with success in searching

As outlined in Table 2-2, factors that impact the performance of searches have been considered in previous studies, albeit in a limited way. These factors are related to the characteristics of: a) the clinical questions searched for, b) the searching physician, and c) the search query. The documented studies were conducted on various MEDLINE interfaces, some of which are obsolete, and used numerous definitions of success. The three studies which identified factors associated with the ability to retrieve relevant literature also varied on their choice of relevant literature[41;45;50;55;60-62]. In fact, researchers and librarians are unclear on the best practices for searching as is apparent when reviewing the content of literature training sessions[60]. A systematic study is needed to determine which factors can positively impact the ability of physicians to retrieve relevant literature. Better knowledge of these factors would help to train physicians to search literature more effectively.

2.9.3 Evaluation studies of MEDLINE filters on end-user searching

While numerous search filters in MEDLINE have been developed and tested[5-10;34;78-87;89;90], only two studies have attempted to determine how these filters actually impact end-user (physician) searching[45;53]. The first study compared the use of the 'methods' filter, Clinical Queries, on three clinical questions and found that on average the use of the filter resulted in an increased precision[45] (Table 2-1). This study used real searches created by physicians to test the filters. The second study did not utilize real physician searches. Instead researchers created 100 search queries for 100 clinical questions and tested the Clinical Queries sensitive and specific filters, and the Limit option available in PubMed for 'randomized controlled trials' and 'clinical trials'. As this study did not test searching unaided by filters, it is unclear whether the filters indeed improved on the original search queries. Studies are now needed to test whether the use of search filters using searches created by clinicians can improve their ability to retrieve relevant articles.

2.9.4 Sample size

When evaluating electronic databases it is strongly recommended that a reasonable number of search topics be used to arrive at valid conclusions about the ability of the database in finding relevant material (using only 1 search query per topic)[70]. Evaluation studies available in the medical literature are often performed using a single clinical question or a convenient sample of questions, and sample size calculations are seldom provided. As with all research, a broad sample of searches is warranted to maximize generalizations.

2.9.5 Studies targeted at specialists

Although specialists use bibliographic databases more often than general practitioners, most available information on searching practices by physicians has originated from studies targeted at primary physicians, medical residents and medical students[17]. The literature searching abilities of specialists, such as nephrologists, has not been evaluated.

2.10 Conclusion

Physicians continue to search PubMed for answers to clinical questions. Better knowledge is expected to improve the delivery of care. However, how physicians search PubMed for their clinical questions and the factors associated with successful searches remain largely unknown. If the potential of large electronic bibliographic databases to maximize health are to be realized, they must be used to quickly retrieve articles that are both scientifically sound and directly relevant to the health problem physicians are trying to solve, without missing key studies or retrieving excessive non-relevant studies. The use of filters when searching appears to be promising in improving the identification of relevant literature. These filters now require further evaluation with real physician searches. Thus, the focus of this thesis was twofold; first, to identify determinants of search success when nephrologists search PubMed and, second, to examine whether filters can enhance clinician search performance.

CHAPTER 3: STUDY OBJECTIVES

3.1 Study objectives and hypotheses

The purpose of this thesis was to establish the current performance of search queries created by nephrologists (clinicians who care for patients with kidney disorders) in retrieving relevant scientific articles for the treatment of renal patients when searching PubMed, and to investigate the utility of search filters to improve this retrieval. The two measures of search success used throughout this thesis are sensitivity and precision.

3.1.1 Objective 1: Determinants of search success

To determine whether there is a relationship between *search query characteristics* or *nephrologist characteristics* and the ability to identify relevant articles in PubMed for renal treatment questions.

<u>Specific Questions</u>

Search query characteristics

1. Does the **use of multifaceted search features** improve search success?
2. Does increasing the **number of concept terms** improve search success?

Nephrologist characteristics

3. Does **increased experience in literature searching** improve search success?
4. Does **having received previous training in literature searching** improve search success?

<u>Hypotheses</u>*:* Search queries that use of multifaceted search features will improve search sensitivity and precision. Queries that include a greater number of concept terms will exhibit an increased precision but also a lower sensitivity compared to queries with fewer concept terms. When compared to their colleagues, nephrologists who search more often or who have previously received training in literature searching will have enhanced search success.

3.1.2 Objective 2: Impact of search filters on search query performance

To determine whether the addition of PubMed search filters to nephrologist-provided search queries improves the retrieval of relevant articles for renal-treatment questions compared to non-filtered queries. Three types of filters, 'content', 'methods' and 'journal', will be tested, alone and in all combinations, for a total of 17 different filter combinations.

Specific Questions

1. Which filter combinations improve search sensitivity the most?
2. Which filter combinations improve search precision the most?
3. Which filter combinations optimize both search sensitivity and precision?

Hypotheses: The addition of filters will improve a nephrologist's search, compared to a non-filtered search. A combination of all three types of filters, 'content', 'methods' and 'journal', will produce the largest improvement in search sensitivity and precision.

3.2 Conceptual model

The hypothesized conceptual model for the relationship between search query and nephrologist characteristics and their ability to retrieve clinically relevant literature for a specific clinical question using PubMed is depicted in Figure 3-1[41;99].

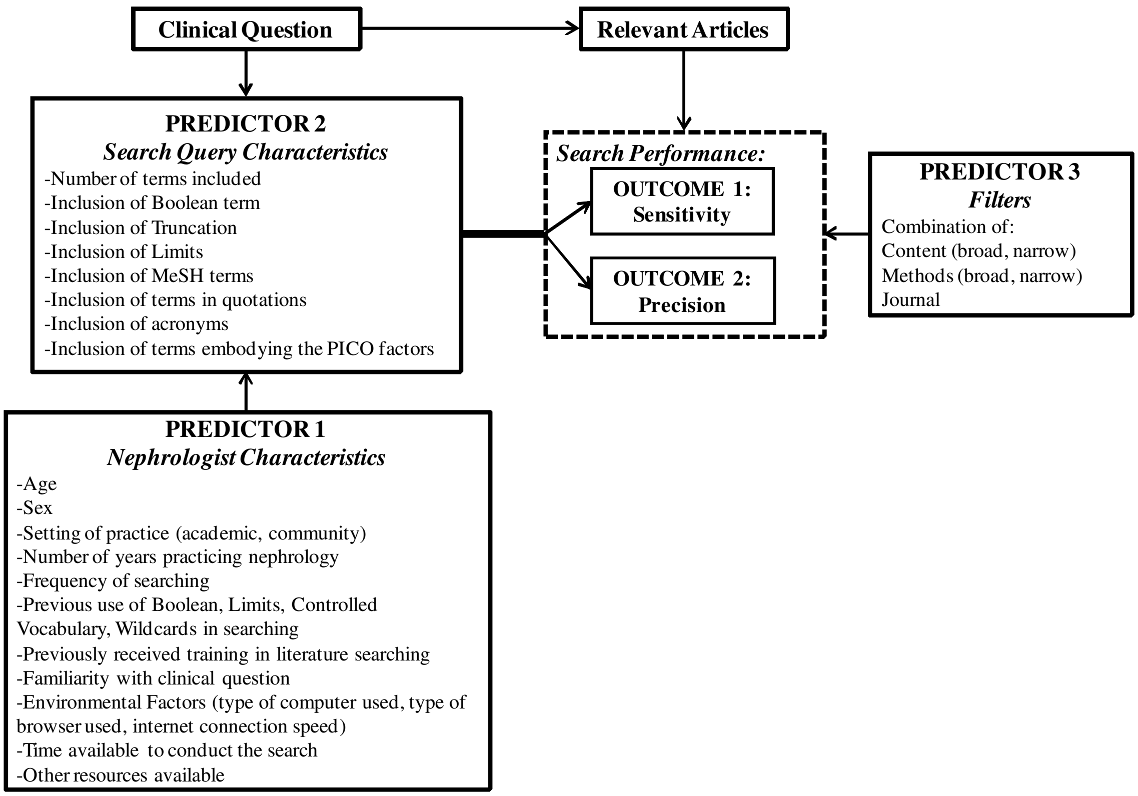


Figure 3-1: Conceptual model

CHAPTER 4: STUDY METHODS

4.1 Study design and sampling

Each study objective employed a cross-sectional study design. The initial data were collected from a survey of nephrologists in Canada. The thesis methods are outlined in three steps:

STEP 1. Assembled a series of treatment questions, to which there were known sets of relevant articles.

STEP 2. Surveyed nephrologists in Canada. Participants were asked to provide search queries that they would use in PubMed to address renal-treatment clinical questions.

STEP 3. Executed each nephrologist-provided search query in PubMed.

4.1.1 Assembled sample of treatment questions and relevant articles

<u>A set of treatment questions</u>

To gather a sample of real search queries used by physicians, Canadian nephrologists were surveyed. Each nephrologist was presented with one or two renal-treatment questions and was requested to provide a search query to address the question(s) (details of the survey are provided in Section 4.1.2). Thus, to maximize the generalizability of the search queries received, it was vital that the questions be directly applicable to the main study group: nephrologists. To assemble a representative set of renal treatment questions, the objectives of recently published renal systematic reviews were selected as they targeted questions in patient care where uncertainty exists. The EvidenceUpdates (http://plus.mcmaster.ca/evidenceupdates) service was used to identify systematic reviews. This service pre-screens and identifies recently published systematic reviews and meta-analyses from over 130 journals that meet strict methodological criteria and have a high potential for clinical relevance[100]. The following criteria is used by the service to identify reviews: "the clinical topic being reviewed must be clearly stated; there must be a description of how the evidence on this topic was tracked down, from what sources, and with what inclusion and exclusion criteria"[100]. Only questions of therapy were selected so that the impact of two 'methods' filters, which were optimized for this purpose (required for Objective 2; see Section 4.3) [6;7], could be tested. A standardized checklist was used by two nephrologists to independently confirm whether each review was pertinent to the treatment of renal patients (Appendix 2). This method previously resulted in a kappa (agreement beyond chance) of 0.98[2]. The assessors further determined whether each review asked a focused treatment question with one main objective.

Inclusion Criteria for systematic reviews (treatment question):

1. Answers one treatment question for renal patients as identified in the Objectives section.
2. Includes a statement of the search strategy used, including years searched.
3. Lists all primary articles used, including evaluation of the methodological quality.
4. Includes two or more primary articles indexed in PubMed.

The primary objectives from the Introduction section of each review were used to compose the treatment questions. Each objective was transformed into a question, using the wording recorded in the review (see example below). Furthermore, for each systematic review and primary articles included in the review the following information was collected: article title, journal name, all authors, publication year and PubMed unique identifier (PMID), if available. All information was entered into an Excel™ spreadsheet.

Example:
Objective: We aimed to assess whether prophylactic use of acetylcysteine reduces incidence of contrast nephropathy in patients with renal insufficiency[101].
Clinical Question: Does prophylactic use of acetylcysteine reduce the incidence of contrast nephropathy in patients with renal insufficiency?

At the initiation of the study, it was anticipated that 100 systematic reviews that satisfied the inclusion criteria would be collected (and acted as the sample size for Objective 2, see section 4.3.3). The EvidenceUpdates service was searched monthly until 100 eligible reviews were compiled in November 2009. In total, 207 reviews from EvidenceUpdates were found by selecting the option to view all reviews for the discipline of nephrology. Details of the review process leading to 100 included reviews are diagrammed in Figure 4-1. The 100 reviews included an average of 16 primary articles that were indexed in PubMed (ranging from 2 to 68; median=12), and together covered a variety of renal topics [acute kidney injury (n=24); chronic kidney disease (n=22); dialysis (n=22); renal transplantation (n=20); glomerular diseases (n=11); other (n=1)]. Of the 100 reviews, 84% included only randomized controlled trials, while 16% included both randomized and non-randomized trials. Further details of the systematic reviews and their objectives are presented in Appendix 3.

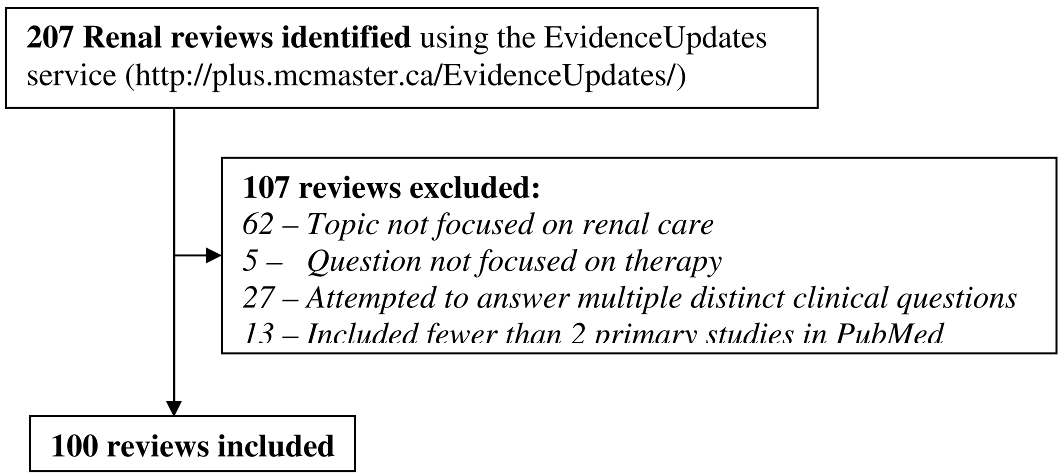


Figure 4-1: Process of including renal therapy reviews

<u>Relevant articles</u>

The purpose of performing a search in PubMed is to identify relevant articles to answer a question of interest. For this thesis, as a means of measuring the performance of a search query, a set of relevant articles was required for each question, also known as a reference standard. Instead of using a subjective measure of relevance to identify important articles, the primary articles included in each review were considered as sets of relevant articles. Thus, for each clinical question, the primary articles from the corresponding systematic reviews served as the set of relevant articles. Articles not indexed in PubMed were excluded. To determine if an article was available in PubMed, the PubMed single citation tool was used (http://www.ncbi.nlm.nih.gov/entrez/query/static/citmatch.html) to search for each article. This involved searching various combinations of the article's title (both English and non-English), the authors' names, journal title, page numbers and the year published. All links to candidate matches were selected to confirm a true match and the PMID was recorded. If a primary article could not be found in PubMed, further searches were performed by a second assessor to confirm the article's absence from the database.

The second assessor also confirmed that each collected PMID corresponded to the proper extracted citation.

4.1.2 Surveyed nephrologists

Survey design

The survey was designed to query nephrologists about their information-gathering practices (Appendix 4). The survey design followed the Dillman tailored design method[102] and used recommendations from a resource that targets the design and conduct of self-administered surveys for clinicians[103]. The survey included questions on demographics (age and gender), years of nephrology practice, practice location, and general literature searching preferences and practices. Participants were also queried on their frequency of use of online information sources as well as their use of Boolean logic operators (‘AND’, ‘OR’, ‘NOT’), controlled vocabularies (e.g. MeSH terms in PubMed), search limits (e.g. language, publication type), PubMed Clinical Queries feature and truncation symbols (inclusion of multiple endings achieved through use of the * symbol in PubMed – e.g. nephro*) when searching bibliographic resources. In addition, each participant was provided with treatment questions and were requested to provide the search queries they would use in PubMed to address the questions. To minimize respondent burden, each nephrologist was provided a maximum of two treatment questions. Pilot testing for validity and usability of the survey was conducted by four nephrologists, two research physicians trained in internal medicine, three individuals trained in library sciences, and three non-medical graduate students. Ethics approval was received from the Office of Research Ethics at the University of Western Ontario (Appendix 5).

The survey was made available via fax, mail or online. The online version of the survey was housed on the Kidney Clinical Research Unit (London Health Sciences Centre, London, ON) server. To access a survey for completion, physicians were provided with a personalized web link (URL).

Administering the survey

Using the Royal College of Physicians and Surgeons of Canada[104], Provincial Colleges of Physicians and Surgeons[105] and the Canadian Medical Directory[106] online databases, a list of 519 practicing nephrologists in Canada was identified in 2007. The sampling frame consisted of nephrologists practicing in nine of the 10 Canadian provinces and included both academic (practicing in a centre with a fellowship training program) and community-based nephrologists. Nephrologists' full name, mailing address and phone number were recorded. Designation date and year of graduation from medical school, fax number and email address were also recorded when available. Eligible participants for the survey included English-speaking, practicing nephrologists in Canada. Consent to participate in the survey was indicated by completing and returning the survey as stated in the letter of information and consent.

The survey respondents were divided into two groups (the information from Group 1 was used for Objective 1 of this thesis, and the information from Group 2 was used for Objective 2). Group 1 included 60 nephrologists (see Section 4.2.3 for sample size calculations), all of whom received two identical clinical questions. The two questions were randomly selected from the final set of 100 systematic reviews (see Section 4.1.1). Each question was selected from two strata; strata one consisted of reviews with more than the median number of included articles (median=12 articles) and strata two consisted of reviews with the median number or fewer included articles. The first question was based on a review that consisted of 49 included studies. Physicians were presented with the following question: "How effective and safe are statins for renal and cardiovascular outcomes in each stage of chronic kidney disease (pre-dialysis, dialysis, and transplantation)?" The second review included 4 primary studies and physicians were presented with the following question: "What are the benefits and harms of continuous ambulatory peritoneal dialysis (CAPD) versus automated peritoneal dialysis (APD) for end-stage renal disease?"

For Group 2, responses were collected from 100 nephrologists (see Section 4.3.3 for sample size calculations), where each physician received one unique clinical question, randomly selected from the 100 included systematic reviews (Appendix 3). All other questions in the survey remained identical between Groups 1 and 2.

The survey was conducted by two research team members: the author (SS) and an undergraduate research assistant. The methods of simple random sampling, and the tailored design method outlined by Dillman[102] were applied to the survey administration. Specifically, nephrologists were selected from the sampling frame using a random number generator; one nephrologist was selected at a time. Nephrologists were initially invited to participate in the survey by email (if available) or phone. Interested nephrologists were provided with a copy of the survey using the modality of their choice (fax, mail or a web link) that included a letter of information describing the study objectives, consent to participate and assurance that all responses would be kept confidential. Each survey was coded to track for non-responders. Participants who did not submit a completed survey within three weeks were sent a follow-up correspondence with a copy of the survey. Another follow-up correspondence via telephone or email (if available) occurred approximately three weeks after the previous follow-up attempt where the participant was once again sent a copy of the survey via the method of their preference. A nephrologist was considered a non-responder if the survey was not completed within nine weeks from the initial date of physician contact. The survey was then re-administered to a new nephrologist. Upon receipt of a completed survey, if a respondent did not provide a search query to the clinical question(s) presented in the survey (question 11), the same survey was re-administered to a new participant. This was done because the response to this question formed the basis of the analysis for this thesis. The survey administration was continued until 60 responses (with question 11 completed) were received for Group 1 and 100 responses for Group 2. Upon receipt of a completed survey, the responses were entered into a spreadsheet. To minimize data entry error, a second reviewer verified the correctness of the entries. All electronically completed surveys were printed. All completed surveys (print and electronic) were archived.

Table 4-1: Summary of survey groups

Group	**Sample Size**	**Number and type of Clinical Question(s)**	**Target Objective**
Group 1	60	*Two* identical questions posed to all participants	Objective 1
Group 2	100	*One* unique question posed to each participant	Objective 2

Dealing with missing or invalid survey responses

As the survey was self administered by nephrologists, the occasional question was left unanswered leading to missing data or invalid responses. As stated earlier, only surveys with completed search queries (responses to question 11) were included in the study. In total, complete data for 14 variables were central to the analysis. Methods to overcome data discrepancies for these responses are listed in Table 4-2.

Table 4-2: Methods to overcome missing responses or discrepancies

Question Number	Variable label	Measurement	Methods to overcome missing responses or discrepancies
2	Use of PubMed in the past year	Binary: Yes; No	*Missing value:* if responses to other parts of this question were provided, then mark no; otherwise leave as missing value
2	Use of PubMed Clinical Queries in the past year	Binary: Yes; No	*Missing value:* if responses to other parts of this question were provided, then mark no; otherwise leave as missing value
3	Previous training in literature searching	Binary: Yes; No	*Missing or illegible value:* leave as missing value
5	Frequency of searching	Quantitative (≥0)	*Missing or illegible value:* leave as missing value *Range of values provided ex. 2-10:* Select mid-point of range (ex. 6).
7	Number of results scanned	Quantitative (≥0)	*Missing or illegible value*: leave as missing value *Rage of values provided:* Select upper limit of range
11	Search query or queries	Text	Surveys with no search query provided were excluded from the analysis
12	Used Boolean logic operators	Binary: Yes; No	*Missing value:* if responses to other parts of this question were provided, then mark no; otherwise leave as missing value
12	Used Limits	Binary: Yes; No	*Missing value:* if responses to other parts of this question were provided, then mark no; otherwise leave as missing value
12	Used Controlled Vocabulary (e.g. MeSH)	Binary: Yes; No	*Missing value:* if responses to other parts of this question were provided, then mark no; otherwise leave as missing value
12	Used Truncation/ Wildcards	Binary: Yes; No	*Missing value:* if responses to other parts of this question were provided, then mark no; otherwise leave as missing value

Continued on following page…

Question Number	Variable label	Measurement	Methods to overcome missing responses or discrepancies
13	Academic Setting	Binary: Yes; No	*Missing value or illegible value:* contact the physician's place of work to determine if their institution is a centre with a nephrology fellowship training program
14	Years Practicing	Quantitative (1-40)	*Missing value or illegible value*: refer to the Directory of Fellows to determine the date the physician completed their nephrology fellowship and accordingly calculate the difference in the date to the year the survey was received
15	Sex	Binary: Yes; No	*Missing value:* refer to the Directory of Fellows or Canadian Medical Directory to ascertain sex
16	Age	Quantitative (25-75)	*Missing value:* leave blank

4.1.3 Performed searches

The primary analyses required performing 120 searches for Objective 1 (60 search queries x 2 clinical questions) and 1800 searches for Objective 2 (100 search queries x 18 different searches each; 1 non-filtered search and 17 filtered aided searches). Details of the searches are provided in the respective objectives' methods sections (Sections 4.2.1 & 4.3.1). When searching PubMed, it takes approximately one to five minutes to execute a search and download the results. Thus, performing 1920 searches would require a minimum of 32 hours to complete. Manual searching is also prone to human error as searches need to be copied and pasted into PubMed and appropriate limits be specified. PubMed is cognizant of research that requires batch searches and as such has developed the Entrez Programming Utilities (http://eutils.ncbi.nlm.nih.gov/; also known as eUtils). The eUtils are designed to be used within programs that automate the searching process by directly interfacing with PubMed to execute searches. Programs that use eUtils can also download information about the searches they execute, such as the number of total results retrieved and details about all the articles found (PMID, article titles, authors etc.). Accordingly, instead of manually performing the searches for this thesis, a program was developed using the Perl programming language (http://www.perl.org/) that used eUtils to automate all searching. Before using this program for the thesis, the process was tested and it was confirmed that the results retrieved through eUtils matched those retrieved using the PubMed interface.

4.1.4 Compared responders to non-responders

To elucidate potential systematic non-response and aid with conclusions of generalizability an analysis of the baseline characteristics (province of clinical practice and gender) of non-responding physicians, compared to physicians from whom responses were received was conducted.

4.2 Objective 1: Determinants of search success

4.2.1 Study variables

Objective 1 made use of the survey results from Group 1, where each participant was provided with the same two renal-treatment questions. The only modifications made to the physician searches were the addition of appropriate PubMed searching syntax for the MeSH terms and limits as indicated by the physicians on the survey. The Perl program discussed in section 4.1.3 was used to execute the queries received for both questions in PubMed. Each executed search was restricted to the search dates provided in the methods section of the systematic review from which the clinical question originated. The variables captured from the survey and used in the analysis are provided in Table 4-3.

<u>Outcome: Search query performance</u>

For each search executed in PubMed the total number of articles found and the number of relevant articles found were collected. To determine the latter, the PMIDs of the retrieved articles were compared to the PMIDs of the articles identified from the systematic review corresponding to the clinical question. Using the collected information, two outcome measures per search in PubMed were calculated: sensitivity and precision. These measures are summarized in Table 2-3. Sensitivity and precision are widely used and acceptable measures in evaluating search performances[76].

<u>Predictor variables</u>

The purpose of this objective was to characterize the relationships between nephrologist characteristics or search query characteristics (labeled Predictor 1 and Predictor 2, respectively, in Figure 3-1) and the outcome of a search. All measurable factors previously identified as predictors of search success were examined (Section 2.4.2). While previous studies identified various definitions of ‘success’, this analysis determined whether the identified factors were associated with the current success/outcome definitions (sensitivity and precision). The main predictors of interest were: 1) number of concept terms included in the search query, 2) the use of multifaceted search features in the search query, 3) the nephrologist’s experience with literature searching as indicated by the frequency of searches performed in a bibliographic database per month and 4) whether the nephrologist had previously received training in literature searching. All predictors were obtained from the survey results. The first two predictors were determined by analyzing the search queries that were provided by the participants. The measurement of the primary predictors are explained below and summarized in Table 4-3. Other potential predictors included the inclusion of each of the PICO factors in a search query, the use of acronyms or quotations in the query, the number of years of nephrology practice, age of the nephrologist, sex of the nephrologist, whether the nephrologist had used Boolean logic operators, limits, truncation or MeSH searching in the past, and whether the nephrologist worked in an academic setting.

Explanation of predictor: number of concept terms included in the search query

As explained in the Introduction (Section 2.3) PubMed performs a process of query pre-processing before executing a search. Two procedures used in this pre-processing are the mapping of concept terms to appropriate MeSH terms and, adding 'AND' between concepts. For example, in the search "statin AND acute renal failure", PubMed would map the concept term 'statin' to the MeSH term "hydroxymethylglutaryl-coa reductase inhibitors"[MeSH Terms] and the concept term 'acute renal failure' to the MeSH term "acute kidney injury"[MeSH Terms]. Thus, a **concept term** was defined as a word or group of words that embody a clinical aspect used in a search query combined with an 'AND' operator (either implicit or explicit) (see explanation of the use of 'AND' operator in Section 2.8). To identify concept terms from each physician-provided search query, the queries were reviewed in duplicate by two research team members: the author (SS) and a nephrologist. In all cases, the <u>most specific</u> clinical concept was selected. For example, 'acute renal failure' also includes the concept 'renal failure'. However, 'acute renal failure' is the more specific concept and thus was considered as one concept term. Discrepancies were resolved through consensus.

Explanation of predictor: use of multifaceted search features

Multifaceted search features include either the use of a Boolean logic operator (‘AND’, ‘OR’ or ‘NOT’), controlled vocabulary (MeSH), truncation/wildcard, PubMed specified Limit or Search Filter (such as Clinical Queries) in a search query.

For the purposes of this analysis, the Boolean logic operator ‘AND’ was only considered as a multifaceted search feature when it was used within a concept term (see definition of concept terms on previous page). By default, PubMed automatically adds an ‘AND’ operator between each concept in a query. Therefore, the explicit addition of an ‘AND’ operator in a search query is redundant and does not change the results from a search[107]. For example, the search queries “statin AND acute renal failure” and “statin acute renal failure” retrieve identical results in PubMed (the same 188 results were retrieved when tested on January 12, 2011). Thus, the first search query would not meet the definition of a multifaceted search feature. However, the search query “statin AND acute AND renal failure” would meet the definition of a multifaceted search feature as the ‘AND’ appears within the concept “acute renal failure”. This search found 190 results, but only 174 overlapped with the previous two queries.

Explanation of predictor: nephrologists experience with literature searching

The experience of a nephrologist with literature searching was quantified as the average number of times a physician searched a bibliographic database each month. This predictor is referred to as **frequency of searching** and was self-reported by the respondents when they answered survey question #5: “On average, how many times per month do you search a bibliographic database for medical literature?”

Explanation of predictor: previous training in literature searching

Previous training in literature searching was self-reported by each respondent when they answered survey question #3: “Have you previously received training in literature searching? *Examples of training include Searching Skills Workshops, Library Training Sessions, PubMed Tutorials”.*

Table 4-3: List of study variables for Objective 1
Variables in bold are predictors of interest for Objective 1: Determinants of search success.

Category	Variable	Measurement
OUTCOME: Search Query Performance	Sensitivity/Recall	Continuous 0.0-1.0: The number of relevant citations identified (primary studies included in systematic review) compared to the total number of relevant citations
	Precision/Positive predictive value	Continuous 0.0-1.0: The number of relevant citations identified divided by the total number of citations retrieved by a search
PREDICTOR 1: Nephrologist Characteristics *All characteristics were self reported on the survey*	**Frequency of searching:** the average number of times per month a physician uses a bibliographic database to search for medical literature	Quantitative: (≥0)
	Previous training in literature searching	Binary: Yes; No
	Years Practicing: the number of years a nephrologist has been working since completing their nephrology training	Quantitative: 1-40
	Academic Setting: whether a nephrologist worked in an institution with a nephrology training program	Binary: Yes; No
	Has previously used Boolean logic operators, Limits, Controlled Vocabulary or Truncation/Wildcards: an answer of 'Yes' to any one of the questions	Binary: Yes; No
	Sex	Categorical: M; F
	Age	Quantitative: 25-75

Continued on following page…

Category	Variable	Measurement
PREDICTOR 2: Search Query Characteristics *All characteristics were determined by examining the search queries provided by the physicians*	**Number of Concept Terms included in the Query**	Quantitative (≥1)
	Multifaceted Search Feature: Boolean logic, Limits, Truncation MeSH or Search Filter; Boolean logic operator 'AND' only considered when it appeared within a 'concept'	Binary: Yes; No
	Use of an acronym: whether the search query included an acronym (e.g. CKD for 'chronic kidney disease')	Binary: Yes; No
	Use of quotations: whether the search query included words enclosed in double quotation marks	Binary: Yes; No
	Use of PICO factors: whether the search query included any terms related to the PICO factors (Patient, Intervention, Comparison, Outcome)	Binary: Yes; No

Potential confounders

An examination of the literature did not identify any potential confounders for inclusion in this analysis. Instead, the method of conceptual model evaluation[108] was used to identify potential confounders through the assessment of the model depicted in Figure 3-1. Analyses for the two predictors, 'Nephrologist Characteristics' and 'Search Query Characteristics' were conducted separately; see Figure 3-1 and Section 4.2.2.

Nephrologist Characteristics: All measured covariates were considered on an individual basis for inclusion as confounders. Identified confounding variables were controlled for in the analysis phase.

Search Query Characteristics: An evaluation of the conceptual model (Figure 3-1) identified two factors, 'Clinical Question' and 'Relevant Articles', which may act as confounders in the relationship between the search query characteristics and the outcomes. One method to control for confounding at the design phase is by restriction or selection. This method operates on the principle that a variable cannot exhibit a confounding effect if it cannot vary within subjects[109]. Thus, the study was designed such that the same two renal-treatment questions were provided to all the participants of Group 1 and analysis was conducted separately for each clinical question (Section 4.2.2). This countered the potential confounding effect that varied clinical questions (and their corresponding sets of relevant articles) may have incurred on the analysis. All measured covariates for 'Search Query Characteristics' were also considered on an individual basis for inclusion as confounders. Variables identified as confounders were controlled for in the analysis phase.

4.2.2 Analytic strategy

The regression modeling strategy outlined by Kleinbaum et al. was used to target the specific questions in Objective 1[110]. Regression modeling was used to identify the association between the two categories of predictors, search query characteristics and nephrologist characteristics, on search success (sensitivity and precision). Upon examination of the conceptual model (Figure 3-1) it can be seen that while nephrologists create search queries to be used in PubMed, the characteristics of the query immediately impact the performance of the search. In other words, the search query characteristics act as an intermediate variable between the nephrologist characteristics and the search outcomes. Thus, separate regression models were developed for the two forms of predictors, as embedding the search query characteristics into an assessment of the physician characteristics on the search outcomes could lead to an over-adjustment and inaccurate results[111]. A different set of models was designed for each clinical question (Table 4-5), for a total of eight models. Sixty (60) observations were included in each model, one for each surveyed nephrologist. Descriptions of analysis follow.

Initial analysis

Initial univariate exploratory analyses were conducted for the all covariates and outcome variables (covariates include the main predictors and potential confounders). This analysis consisted of an examination of descriptive statistics, frequency distributions, missing data and outliers. Continuous data were summarized by the mean and standard deviation or median and interquartile range as appropriate.

Selection of potential confounders

All measured covariates were considered for selection as potential confounders. Confounders were chosen for inclusion in the models using collapsibility testing; a variable was included in the model if its addition changed the regression coefficient of the primary predictor by approximately 10% or more[109]. An examination of collinearity between the primary predictors and each covariate was also considered[112].

Multivariable analysis

As stated earlier, four regression models were developed for each clinical question. The dependent and independent variables are depicted in Table 4-5. For each model, a multivariable linear regression model was initially fitted and the assumptions of normality, homoscedasticity and linearity were assessed[112-114]. In cases where assumptions did not hold, attempts were made to remedy any deviations as appropriate. If deviations persisted, a new modeling procedure was selected. Regression diagnostics were performed on the final models (residual analyses, assessment of outliers and assessment of collinearity). Model building and regression diagnostics are detailed separately for each of the eight models in Appendices 7-14. Table 4-4 provides a reference to the appropriate Appendix for each model.

4.2.3 Sample size

Pilot data generated for Objective 2 (Appendix 15) was used to guide the sample size calculations for Objective 1. For all calculations, power was specified at 80% with a significance level of 0.05. Predictors were dichotomized for the purposes of sample size estimation and details of the calculations are presented in Appendix 6. The estimates computed represent the minimum number of subjects needed to detect a minimum mean difference of 15% in sensitivity between predictor groups. This resulted in a total sample size of 60 responses. The sample size was also able to detect a minimum difference of 4% in precision.

Table 4-4: References to appropriate appendices for regression models

Model Number	Predictor Type	Outcome	Clinical Question	Appendix Number
1	Search query characteristics	Sensitivity	1	7
2	Search query characteristics	Precision	1	8
3	Nephrologist characteristics	Sensitivity	1	9
4	Nephrologist characteristics	Precision	1	10
5	Search query characteristics	Sensitivity	2	11
6	Search query characteristics	Precision	2	12
7	Nephrologist characteristics	Sensitivity	2	13
8	Nephrologist characteristics	Precision	2	14

Table 4-5(a-d): Regression models and variables considered for model inclusion

(a) Predictor Type: Search Query Characteristics

Outcome	Primary Predictors	Potential Confounders
Sensitivity	• Number of concept terms • Use of multifaceted search features	• Inclusion of terms for each of the PICO factors • Inclusion of acronym term • Use of quotations

(b) Predictor Type: Search Query Characteristics

Outcome	Primary Predictors	Potential Confounders
Precision	• Number of concept terms • Use of multifaceted search features	• Inclusion of terms for each of the PICO factors • Inclusion of acronym term • Use of quotations

(c) Predictor Type: Nephrologist Characteristics

Outcome	Primary Predictors	Potential Confounders
Sensitivity	• Frequency of searching • Previous training in literature searching	• Years practicing nephrology • Practice in an academic setting • Previously used multifaceted searching • Sex • Age

(d) Predictor Type: Nephrologist Characteristics

Outcome	Primary Predictors	Potential Confounders
Precision	• Frequency of searching • Previous training in literature searching	• Years practicing nephrology • Practice in an academic setting • Previously used multifaceted searching • Sex • Age

4.3 Objective 2: Impact of search filters on search query performance

4.3.1 Study variables

The survey data from Group 2 (see Table 4-1) was used to assess filter performance (Objective 2). For each response, 18 different searches were performed. The first search consisted of a query provided by a physician, with no filters applied. The next 17 searches combined the search query provided by a physician with at least one type of filter ('methods', 'content' or 'journal') (Table 4-6). The 18 searches can be tabulated by considering there are three options for each of the methods and content filters (no filter, narrow filter, broad filter) and two options for the journal filter (no filter vs. filter), for a total of 3 (methods) x 3 (content) x 3 (journal) = 18 different searches, or one unaided search and searches aided with each of 17 different filter combinations. The only modifications made to the physician searches was the addition of appropriate PubMed searching syntax for the MeSH terms and limits as indicated by the physicians on the survey. Each executed search was restricted to the search dates provided in the methods section of the systematic review from which the clinical question originated.

Table 4-6: Filters available for testing

Category	Available Filters	Special Instructions
Journal[11]	Renal Journal Subset	
Methods[7] (therapy)	Broad Narrow	Remove all methods terms from physician-generated search query
Content[2]	Broad Narrow	Remove all renal content terms from physician-generated search query

The advantage of using filters for specific subject areas ('methods' or renal 'content') is that some terms need not be entered in the search query; rather, the filters act as a substitute for these terms. For example, instead of adding the term 'clinical trial' to a search query, a user can simply select the 'methods' filters for evaluations of treatments, which would filter PubMed to those studies using best methods for questions of therapy (i.e. randomized clinical trials). Thus, when the 'methods' and/or renal 'content' filters were added to physician-provided searches, methods and/or renal content terms were removed from the physician search queries. To do this, each search query was reviewed independently and in duplicate by two assessors trained in epidemiology and medicine. Discrepancies in decisions to remove terms by the assessors were resolved by consensus.

> Example of removing methods and content terms as appropriate:
> *Clinical Question:* What are the benefits of intradermal compared to intramuscular hepatitis B vaccination in chronic kidney disease?
> *Search query provided by a physician (unaided):* hepatitis b vaccination dialysis randomized trial
> *Query aided by methods filter:* hepatitis b vaccination dialysis ~~randomized trial~~ AND < methods filter>
> *Query aided by content filter:* hepatitis b vaccination ~~dialysis~~ randomized trial AND <content filter>
> *Query aided by methods & content filters:* hepatitis b vaccination ~~dialysis~~ ~~randomized trial~~ AND <methods filter> AND <content filter>

The Perl program discussed in Section 4.1.3 was used to execute all searches in PubMed. For each search, the program downloaded the list of PMIDs corresponding to the search results in the exact order they would have been displayed to a user of the PubMed interface. In addition, for each search, the total number of articles found and the number of relevant articles found were recorded. To determine the latter, the PMIDs of the retrieved articles were compared to the PMIDs of the relevant articles identified from the systematic review corresponding to the specified clinical question. This process was also automated using a separate Perl program.

4.3.2 Analytic strategy

The performance of each search (sensitivity and precision) was measured by comparing the articles found by the search to a set of relevant articles as defined in Table 2-3. The differences in sensitivity and differences in precision were then calculated between every physician-provided search query, and the query when each of 17 filter-combinations was applied.

<u>Primary analysis</u>

The primary analysis compared the application of filters to physician-provided search queries and considered the full set of results returned by PubMed. A paired design was employed where outcomes of the search queries (both sensitivity and precision) were compared with and without the use of filters. The analysis determined whether any of the 17 filter-combinations (applied to physician-provided queries) resulted in significant improvements in outcomes, compared to non-filtered searches. To account for the large number of significance tests required for this study (17 tests for sensitivity, 17 tests for precision, total 34 tests), the application of a multiple comparisons procedure was adopted to reduce the risk of type I error. Two appropriate methods are Dunnett and Bonferroni[115]. The Dunnett method is used for pairwise comparisons of interventions to a control group; for this study, the mean values of the physician-provided search queries (control, unaided by filters) would be compared to the same queries with each of the 17 filter-combinations applied (intervention). The second method, Bonferroni, is a conservative method where the alpha value for the hypothesis test is adjusted by dividing the initial alpha by the number of comparisons; in this case the alpha value would be adjusted to 0.0015 (0.05 ÷ 34)[116]. Unfortunately, the Dunnett's method that is available in most statistical packages does not allow for paired analysis[115]. However, if the Bonferroni method is applied, a paired analysis can be used. To compare the two methods, a power calculation was conducted for the Dunnett's method using the SAS Statistical Package version 9.2 (SAS Institute Inc., Cary, NC, and U.S.A.) (without a paired option; see SAS code in Appendix 17) to match the sample size calculations discussed next (Section 4.3.3). This resulted in a sample size of 179 survey responses (Appendix 18), which is much greater than the sample size of 100 that was required with the Bonferroni alpha adjustment. Given the conservative nature of the Bonferroni method and the large difference in survey responses that would have been required with the use of an unpaired Dunnett's test, the Bonferroni-adjusted method was applied for this analysis.

Initially, a 2-sided one-sample (paired) t-test was selected to determine if a difference existed between non-filtered searches and filter-aided searches. The Null Hypothesis, H_0, was defined as a mean difference in sensitivity or precision between non-filtered searches and filter-aided searches equal to zero. The Alternate Hypothesis, H_1, was defined as a mean difference in sensitivity or precision not equal to zero.

Before continuing with the t-test, the assumptions of the test were examined[112;117]. The use of the t-test requires that a) each search query and the subsequent outcomes of the search are independent, b) the outcomes are approximately normally distributed and c) the variances of the filtered and non-filtered searches are approximately equal. The first assumption, independence, was assured by the sampling method as each physician was independently provided with one unique clinical question for which they provided a search query. For the second assumption, histograms were used to visually assess normality. If a measure is normal, a histogram would be expected to exhibit a bell-shaped pattern with 95% of the points appearing within two standard deviations from zero. The histograms for the outcomes of sensitivity and precision are diagrammed in Appendix 19. While some histograms appeared bell-shaped, many were skewed and exhibit peakedness due to the large number of zeros, suggesting the results did not follow a normal distribution.

As the assumptions of the t-test did not hold for this analysis, the non-parametric paired Wilcoxon Signed-Rank test was used instead. While this test does not stipulate assumptions of normality, it requires that the observations be independent (which was previously confirmed) and the outcome measures come from a symmetric distribution. There is no information in the literature to suggest that the differences in search sensitivity and precision do not follow a symmetric distribution. However, as a confirmatory analysis the Sign test, which does not require any distributional assumptions, was also conducted. A similar pattern of results observed with the Sign test and Wilcoxon Signed-Rank test would suggest a robustness of the data and a true association if one is observed. The non-parametric tests compare paired samples at the ordinal level of measurement (‘greater than’, ‘less than’, ‘equal to’). They test for equality of the medians between two samples, instead of testing for equality of the means between the two samples, as is done for the t-test. Thus, when testing the differences between the filtered and non-filtered searches, confidence intervals (CI) were produced for the median value. This was done using the SAS Statistical Package version 9.2 (SAS Institute Inc., Cary, NC, and U.S.A.), which follows the method of Hahn and Meeker[118].

Secondary analyses

Restricting the results set to the first 40 citations

Results from the survey (Section 5.1.4) indicated that 80% of the participants do not review beyond 40 search results (equivalent to two default search pages in PubMed). Therefore, a secondary analysis was conducted, while restricting the PubMed search results to the first 40 citations. Accordingly, search sensitivity and precision were calculated while considering only the first 40 retrieved results.

Additional analyses

Modifying physician-provided searches

An analysis of the physician-provided searches indicated that in many cases physicians provided misspelled words or acronyms that were not recognized by PubMed and consequently no results or very few results were returned by the database. In total, eight searches resulted in no relevant articles being found, of which 6 (75%) cases included a misspelled word or acronym. In addition, in four of these cases, the results from the filter-aided searches appeared to perform much better that the non-filtered searches as the use of the filters replaced the misspelled search terms or acronyms. Searching is a dynamic process where a failed search is often tried again. It is likely that if a physician recognizes that a misspelled word or acronym resulted in a poor search, they may try the search again after modifying the search. In fact, an analysis of PubMed query logs over a one month period in 2008[62] indicated that when queries did not return any results, 82% of the users searched again with a modified or new search query. Also, in 41% of searches, users performed subsequent searches even when their initial search returned results.

In an effort to mimic how a physician might improve a search (s)he was not satisfied with, a nephrologist, the author (SS) and a medical librarian developed and pre-specified rules by which to modify the physician-provided search queries. The rules are listed in Table 4-7. All modifications were carried out in duplicate by the author (SS) and the medical librarian, and discrepancies were resolved by consensus. The modified search queries were then executed in PubMed and the primary and secondary analyses for the new results were repeated.

Table 4-7: Rules for syntactically improving physician-provided search queries

1. Update MeSH terms indicated as exploded terms and add PubMed syntax for limits described
2. Correct spelling errors
3. Capitalize Boolean logic operators ('AND', 'OR' or 'NOT')
4. Remove commas ',' periods '.' semi-colons ';' and apostrophes "'"
5. Replace '/' with an 'OR' term
6. Replace 'and/or' with an 'OR' term
7. Replace '+' with an 'AND' term
8. Remove preposition and article terms (e.g. 'in,' 'by,' 'at,' 'for,' 'from,' 'a,' 'the')
9. Expand short forms or acronyms and include the original term with an 'OR' term

4.3.3 Sample size

A sample of 100 reviews was identified for study inclusion (see Section 4.1.1). Pilot data (Appendix 15) was used to estimate a standard deviation of 0.28 for the difference in sensitivity, and a standard deviation of 0.14 for the difference in precision. Given a sample of 100 clinical question responses (with each nephrologist receiving a single unique question), power of 80%, an alpha of 0.0015 (to adjust for multiple comparisons; see Section 4.3.2) for a 2-sided paired t-test, this study had the ability to detect a minimum mean difference of 11.5% in sensitivity and a mean difference of 6% in precision. A review of the literature did not elucidate a minimal important difference in sensitivity or precision. However, upon consultation with clinicians and information scientists, these values represented a reasonable benefit to warrant the on-going effort to incorporate the filters into mainstream use. See Appendix 16 for sample size calculation performed using the using SAS Statistical Package version 9.2 (SAS Institute Inc., Cary, NC, and U.S.A.).

4.4 Methods used to minimize potential threats to validity

This thesis adapted methodology originating from the field of information retrieval. Several attempts were made to control for the following biases identified in previous studies of search engine evaluation[77;96]:

***Suggestion**:* To ensure precision in estimates, a sufficiently large number of search topics must be utilized to produce meaningful evaluations of search engine effectiveness.
Solution: Recently published systematic reviews in nephrology were used to assemble a large variety of clinical questions and identify corresponding sets of relevant articles.

***Suggestion**:* To ensure external validity, search topics should be motivated by the genuine information needs of the target users.
Solution: Nephrology systematic reviews were used to gather search topics. Systematic reviews target answerable questions for which uncertainty exists and are of interest to nephrologists.

***Suggestion**:* To ensure external validity, search queries used to evaluate the retrieval quality should be derived by individuals in the target population.
Solution: Through the use of a survey, nephrologists were asked to create search queries that they would use to search for literature to answer a focused clinical question.

***Suggestion**:* To ensure overall applicability, relevance judgments must be made in relation to the target population.
Solution: Primary articles included in systematic reviews were used to identify relevant literature. Through this procedure, the thesis engaged in widely accepted principles of EBM to identify the most important articles to retrieve in a search. In addition, selected systematic reviews detailed reliable and comprehensive methods of assembling relevant articles for a focused therapy question. This helped ensure all sound evidence was accounted for, minimizing subjectivity in the selection of relevant studies.

Furthermore, several methods to avoid bias and maximize generalizations were used:

1) To avoid misclassification of the outcome the dates for which information was compiled in each review was recorded and subsequently all searches were limited to the appropriate start and end dates. Date restriction was used to exclude articles, both relevant and non-relevant, not considered in the systematic review process. In addition, only primary studies that were indexed in PubMed were included.
2) By ensuring that each included systematic review targeted one objective, the study further minimized misclassification bias by ensuring that all included articles in the review were truly relevant for the corresponding treatment question.
3) Selection bias was minimized through the use of random, rather than convenience sampling, to select Canadian nephrologists for survey participation. This ensured that a large variety of nephrologists with varied search abilities participated in the study. Further, for Objective 2, clinical questions were randomly assigned to each nephrologist ensuring that, on average, physicians had equal familiarity with the topic presented. In addition, the characteristics of non-responding physicians were compared to physicians for whom responses were received to identify potential systematic non-response.
4) The survey employed the tailored design method to maximize response rates[102].
5) For Objective 2, when testing the impact of filter usage, the alpha level of significance was adjusted to avoid detecting spurious associations (type I errors) through multiple statistical comparisons.
6) Analysis in Objective 2 employed a paired design to ensure equivalence in potential biases between the non-filtered and filter-aided searches.
7) To increase generalizability, the analysis for Objective 1 was performed separately on each of the two clinical questions. Two different questions with different numbers of relevant articles (4 vs. 49) were selected. While restricting the analysis to a single clinical question was required to limit confounding, the use of a single question would have limited the generalizability of the findings. The analysis of the second question was used to ensure replication of the findings and confirm that the results were not dependent on the question asked or the number of relevant articles available for the clinical question.

CHAPTER 5: RESULTS

5.1 Survey results

5.1.1 Sample

The surveys were administered between January 2008 and October 2010. The survey for Group 1 (used for Objective 1: Determinants of search success) commenced in November 2009, after a sample of 100 reviews was achieved. Two clinical questions were randomly selected and included in the survey (see details in Section 4.1.2). In total, 173 responses from 267 randomly selected nephrologists were received. After excluding 27 known ineligible respondents (not practicing nephrology, moved out of the country, retired or deceased), an overall response rate of 72.1% (173/240) was achieved.

5.1.4 Comparing responders to non-responders

Baseline characteristics comparing respondents and non-respondents are presented in Table 5-1. Respondents and non-respondents differed slightly on gender with a larger proportion of males completing the survey, compared to females. In addition, response patterns differed by the nephrologists' province of practice; nephrologists from Quebec exhibited the poorest response rate (50%) while nephrologists from Manitoba exhibited the best response rate (100%).

Table 5-1: Characteristics of responding and non-responding nephrologists

Characteristic of nephrologist	No. of nephrologists surveyed	No. and proportion of non-responders (n=67)	No. and proportion of responders (n=173)
Gender			
Male	177	43 (24%)	134 (76%)
Female	63	24 (38%)	39 (62%)
Province of practice			
Alberta	27	7 (26%)	20 (74%)
British Columbia	19	3 (16%)	16 (84%)
Manitoba	8	0 (0%)	8 (100%)
New Brunswick	9	3 (33%)	6 (67%)
Newfoundland	6	1 (17%)	5 (83%)
Nova Scotia	8	2 (25%)	6 (75%)
Ontario	96	18 (19%)	78 (81%)
Quebec	64	32 (50%)	32 (50%)
Saskatchewan	3	1 (33%)	2 (67%)

5.1.2 Data management and cleaning

Survey: Group 1 (used for Objective 1: Determinants of search success)

The survey for Group 1 included two identical clinical questions (see Section 4.1.2 for details). Survey responses with missing values or data discrepancies are presented in Table 5-2. These include only the main study variables for Objective 1. Methods outlined in Table 4-2 were used to handle missing values and other discrepancies. One survey was excluded from the analysis (for a total sample of 60 responses) and three observations persisted with missing values after the data management step.

Table 5-2: Number of missing values or discrepancies for survey responses for Group 1

Question Number	Question description	Number and of missing values or discrepancies (n=61) and resolution
3	Previous training in literature searching	0
5	Frequency of searching	Invalid response - illegible: 1 (left as a missing value) Discrepancy – range provided: 3 (selected mid-point of range)
11	Search queries	Missing value: 1 (survey excluded from analysis)
12	Used Boolean logic operators	0
12	Used Limits	Missing value: 1 (since a response was provided for other parts of this question, response changed to ‘no’)
12	Used Controlled Vocabulary (e.g. MeSH)	0
12	Used Truncation/ Wildcards	0
13	Academic Setting	0
14	Years Practicing	Missing value: 1 (calculated using the designation date from the Directory of Fellows)
15	Sex	Missing value: 1 (verified gender with the Canadian Medical Directory)
16	Age	Missing value: 3 (left as missing values)

Survey: Group 2 (used for Objective 2: Impact of search filters on search query performance)

The survey for Group 2 included one unique clinical question, randomly selected from the 100 included systematic reviews (see Section 4.1.2 for details). Results from Group 2 required minimum data management or cleaning as only one survey response was required (question 11). In total, 12 surveys were excluded from further analysis: five surveys with a response to question 11 were missing or illegible and seven surveys were received after a physician was deemed a non-responder.

5.1.3 Respondent characteristics

Respondent characteristics for the 160 eligible survey results are presented in Table 5-3. The average age of respondents was 46 years (range 33–74), having practiced nephrology an average of 14 years (range 1–45). Approximately 77% of the respondents were male and 63% practiced in an academic setting. Respondent characteristics did not differ greatly between Groups 1 and 2.

Table 5-3: Respondent demographics and their characteristics separated by the two groups of surveys

Characteristic	**Overall** n=160	**Group 1** n=60	**Group 2** n=100
Age[¥], years; mean (SD)	46 (10)	46 (10)	47 (11)
Male (%)	77	82	74
Years practicing nephrology; mean (SD)	14 (10)	15 (10)	14 (10)
Practice in an academic setting (%)	63	62	63

Abbreviations: SD, Standard deviation.
[¥]Invalid/missing responses: Overall (9; 6%), Group 1 (3; 5%), Group 2 (6; 6%)

5.1.4 Use of online sources

Question one of the survey asked respondents to indicate whether they had used PubMed or Clinical Queries in the past year to guide the treatment of a patient; 92% indicated using PubMed, while 21% had also used Clinical Queries.

Question seven of the survey asked the respondents to indicate the number of results they scanned, in general, per search in a bibliographic database (see Figure 5-1). This result was categorized into groups of 20 results (since by default PubMed presents 20 citations on one page of results). Exactly 80% of the respondents indicated that they do not scan more than 40 results, which equates to two default search pages in PubMed.

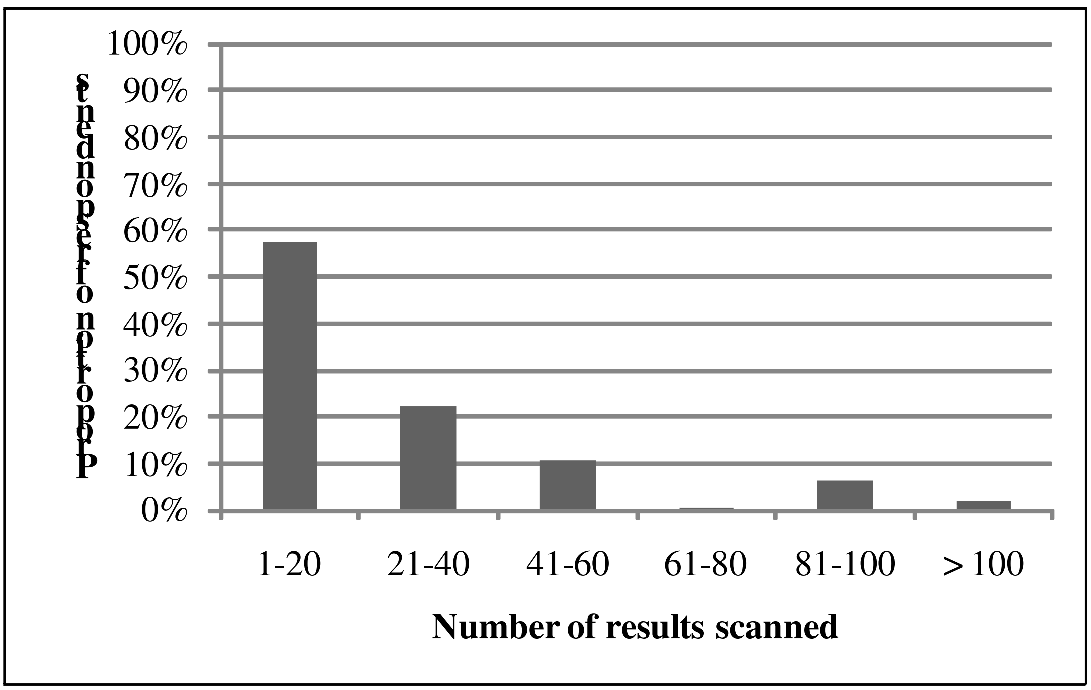


Figure 5-1: Response to survey question: "when you search a bibliographic database, how many results do you generally scan per search?"

5.2: Objective 1: Determinants of search success

5.2.1 Main study variables

Survey responses for Group 1 (where each nephrologist was presented with the same two clinical questions) were used for Objective 1. Descriptive statistics for the main study variables (outcomes, primary predictors and potential covariates) for all 60 eligible surveys are presented in Tables 5-4 to 5-9

The search sensitivity and precision values varied from the search queries provided for the two clinical questions (Tables 5-4 and 5-5). Six queries provided for the first clinical question resulted in no articles being found and thus an invalid precision value was ascertained (division by zero). This occurred in four instances for the second question.

An examination of the search queries for the first clinical question (Table 5-6), revealed that most or all nephrologists included a **p**atient term (such as ‘chronic kidney disease’ or ‘renal insufficiency’) and **i**ntervention term (such as ‘statin’), and did not include a **c**omparison term or quotations in their search queries. An **o**utcome term (such as ‘safety’ or ‘efficacy’) was used in less than half the search queries (43%). The majority of queries (88%) consisted of 2-4 concept terms, 25% included an acronym and 15% included the use of multifaceted search features.

Unlike the search queries provided for question one, most queries provided for question two included a **c**omparison term (for this question this included any terms referring to ‘automated peritoneal dialysis’) (Table 5-7). The remaining features of the queries for both questions were similar. Most or all nephrologists included a **p**atient term and an **i**ntervention term in their query, while an **o**utcome term was used in 40% of the search queries. A quotation was only used in one query, while the majority of search queries (93%) consisted of 2-4 concept terms, 52% included an acronym and 12% included the use of multifaceted search features.

Characteristics of the nephrologist respondents are summarized in Tables 5-8 and 5-9. On average, the respondents had searched a bibliographic database seven times a month, 100% had used advanced search features in the past and 43% had previously received training in literature searching.

<u>Outcome variables</u>

Table 5-4: Outcome variables for Question 1 (used for models 1-4)

Variable	Mean	Median	No. of zero values	Min/Max	Standard deviation	Number of invalid points
Sensitivity	0.16	0.06	16	0.00/0.82	0.20	0
Precision	0.11	0.05	10	0.00/0.75	0.15	6

Table 5-5: Outcome variables for Question 2 (used for models 5-8)

Variable	Mean	Median	No. of zero values	Min/Max	Standard deviation	Number of invalid points
Sensitivity	0.29	0.25	23	0.00/1.00	0.28	0
Precision	0.01	0.001	19	0.00/0.13	0.02	4

Covariates for use in models predicting the effect of search query characteristics on search success

Table 5-6: Primary predictors and potential confounders for use in models 1 & 2

Variable	Frequency
Number of concept terms	
1	0
2	22
3	17
4	14
5	4
6	3
Use of multifaceted searching	
No	51
Yes	9
Use of patient term	
No	0
Yes	60
Use of intervention term	
No	1
Yes	59
Use of comparison term	
No	60
Yes	0
Use of outcome term	
No	34
Yes	26
Use of acronym	
No	45
Yes	15
Use of quotation	
No	59
Yes	1

Table 5-7: Primary predictors and potential confounders for use in models 5 & 6

Variable	Frequency
Number of concept terms	
1	3
2	23
3	20
4	13
5	1
Use of multifaceted searching	
No	53
Yes	7
Use of patient term	
No	0
Yes	60
Use of intervention term	
No	0
Yes	60
Use of comparison term	
No	3
Yes	57
Use of outcome term	
No	36
Yes	24
Use of acronym	
No	29
Yes	31
Use of quotation	
No	59
Yes	1

<u>Covariates for use in models predicting the effect of nephrologist characteristics on search success</u>

Table 5-8: Categorical predictors for use in models 3, 5, 7, 8

Variable	Frequency
Previous training in literature searching	
No	34
Yes	26
Practicing in an academic setting	
No	23
Yes	37
Previously used advanced searching	
No	0
Yes	60
Sex	
Female	11
Male	49

Table 5-9: Continuous predictors for use in models 3, 5, 7, 8

Variable	Mean (median)	Min	Max	Standard deviation	Number of missing values
Frequency of searching (number of times per month)	7 (5)	0	30	6.34	1
Years practicing nephrology	15 (11.5)	2	45	9.91	0
Age	46 (44)	33	74	9.89	3

5.2.2 Relationship between search query characteristics and search success

Models for search sensitivity and precision were developed using linear, Poisson or negative binomial regression, as appropriate (see Appendices 7, 8, 11, 12 for details on model building). The latter two modeling techniques provided estimates for the rate ratio for each predictor. A rate ratio can be interpreted as a relative (between level) increase or decrease of the response variable (sensitivity or precision) by a specific factor with every one unit increase of a predictor, when adjusting for other predictors in the model. Models were developed separately for each measure (sensitivity and precision) and for each clinical question.

Sensitivity

Rate ratios for the final models for each of the two questions are provided in Tables 5-10 and 5-11. For each additional concept term used in a search query, sensitivity decreased by 30% (RR for Question 1: 0.7; 95% CI: 0.5 to 0.9), while the use of multifaceted search features improved sensitivity approximately two-fold. Analysis for question one also suggested that the use of outcome terms or acronyms in a search query can decrease sensitivity. This was statistically significant in the analysis for question one, but not for question two, although the point estimates confirmed the direction of the association for the use of an outcome term.

Table 5-10: Results from negative binomial regression analysis assessing the effect of search query characteristics on sensitivity for question 1[a]

Variable	Estimate of rate ratio[b] (95% Confidence intervals)	p-value
Number of concept terms[c]	0.69 (0.53 to 0.89)	0.005
Use of multifaceted search features[d] (referent group: No)	2.64 (1.39 to 5.00)	0.003
Outcome term[e] used in search (referent group: No)	0.21 (0.12 to 0.39)	<0.001
Acronym used in search (referent group: No)	0.19 (0.09 to 0.36)	<0.001

a:Question 1: "How effective and safe are statins for renal and cardiovascular outcomes in each stage of chronic kidney disease (pre-dialysis, dialysis, and transplantation)?"
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression
c: A concept refers to a word or group of words that embody one clinical aspect
d: Use of multifaceted search features include the use of a Boolean logic operators ('AND', 'OR', 'NOT'), controlled vocabulary (MeSH), truncation/wildcard, or limits
e: Outcome term refers a word of phrase referring to the relevant outcomes of an intervention (e.g. morbidity, mortality, complications)

Table 5-11: Results from the negative binomial regression analysis assessing the effect of search query characteristics on sensitivity for question 2[a]

Variable	Estimate of rate ratio[b] (95% Confidence intervals)	p-value
Number of concept terms[c]	0.61 (0.43 to 0.85)	0.004
Use of multifaceted search features[d] (referent group: No)	2.27 (1.21 to 4.25)	0.011
Outcome term used in search[e] (referent group: No)	0.64 (0.34 to 1.22)	0.176

a:Question 2: "What are the benefits and harms of continuous ambulatory peritoneal dialysis (CAPD) versus automated peritoneal dialysis (APD) for end-stage renal disease?".
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression
c: A concept refers to a word or group of words that embody one clinical aspect
d: Use of multifaceted search features include the use of a Boolean logic operators ('AND', 'OR', 'NOT'), controlled vocabulary (MeSH), truncation/wildcard, or limits
e: Outcome term refers a word of phrase referring to the relevant outcomes of an intervention (e.g. morbidity, mortality, complications)

Precision

Rate ratios for the final models for each of the two clinical questions are provided in Tables 5-12 and 5-13. The increase in the number of concept terms used in a search query was significantly associated with increased search precision (RR for Question 1: 1.6; 95% CI: 1.3 to 2.0), while the use of multifaceted search features also appeared to improve precision. The latter measure was statistically significant for question one (RR: 2.0; 95% CI: 1.3 to 3.3) but not for question two, although the point estimates for question two confirm the direction and magnitude of the association (RR: 2.7; 95% CI: 0.7 to 10.9).

Table 5-12: Results from the negative binomial regression analysis assessing the effect of search query characteristics on precision for question 1[a]

Variable	Estimate of rate ratio[b] (95% Confidence Interval)	p-value
Number of concept terms[c]	1.63 (1.29 to 2.04)	<0.001
Use of multifaceted search features[d] (referent group: No)	2.01 (1.25 to 3.26)	0.004

a:Question 1: "How effective and safe are statins for renal and cardiovascular outcomes in each stage of chronic kidney disease (pre-dialysis, dialysis, and transplantation)?"
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression
c: A concept refers to a word or group of words that embody one clinical aspect
d: Use of multifaceted search features include the use of a Boolean logic operators ('AND', 'OR', 'NOT'), controlled vocabulary (MeSH), truncation/wildcard, or limits

Table 5-13: Results from the negative binomial regression analysis assessing the effect of search query characteristics on precision for question 2[a]

Variable	Estimate or rate ratio[b] (95% Confidence intervals)	p-value
Number of concept terms[c]	2.21 (1.24 to 3.91)	0.007
Use of multifaceted search features[d] (referent group: No)	2.71 (0.67 to 10.90)	0.159

a:Question 2: "What are the benefits and harms of continuous ambulatory peritoneal dialysis (CAPD) versus automated peritoneal dialysis (APD) for end-stage renal disease?".
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression adjusting for the number of outcome terms used
c: A concept refers to a word or group of words that embody one clinical aspect
d: Use of multifaceted search features include the use of a Boolean logic operators ('AND', 'OR', 'NOT'), controlled vocabulary (MeSH), truncation/wildcard, or limits

5.2.3 Relationship between nephrologist characteristics and search success

Models for search sensitivity and precision were developed using linear, Poisson or negative binomial regression, where appropriate (see Appendices 9, 10, 13, 14 for details on model building). The latter two modeling techniques provided estimates for the rate ratio. A rate ratio can be interpreted as a relative (between level) increase or decrease of the response variable (sensitivity or precision) by a specific factor with every one unit increase of a predictor, when adjusting for other predictors in the model. Models were developed separately for each measure (sensitivity and precision) and for each clinical question.

<u>Sensitivity</u>

The results from the multivariable regression testing the relationship between nephrologist characteristics and search sensitivity are presented in Tables 5-14 and 5-15. No associations between nephrologist characteristics and search sensitivity were evident. Attempts were made to analyze the data using different techniques to ensure that the results were not due to poor model specification (details provided in Appendix 9 & 13). However, the absence of an association persisted; all analyses provided effect measure estimates close to unity, with p-values greater than 0.2.

Table 5-14: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on sensitivity for question 1[a]

Variable	Estimate of rate ratio[b] (95% Confidence Interval)	p-value
Frequency of searching[c]	1.03 (0.97 to 1.09)	0.334
Previous training in literature searching (referent group: No)	0.99 (0.44 to 2.21)	0.983

a:Question 1: "How effective and safe are statins for renal and cardiovascular outcomes in each stage of chronic kidney disease (pre-dialysis, dialysis, and transplantation)?"
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression
c: Frequency of searching was self reported as the number of times per month physicians search a bibliographic database

Table 5-15: Results from the Poisson regression analysis assessing the effect of nephrologist characteristics on sensitivity for question 2[a]

Variable	Estimate of rate ratio[b] (95% Confidence intervals)	p-value
Frequency of searching[c]	1.01 (0.97 to 1.05)	0.604
Previous training in literature searching (referent group: No)	1.07 (0.65 to 1.73)	0.814

a:Question 2: "What are the benefits and harms of continuous ambulatory peritoneal dialysis (CAPD) versus automated peritoneal dialysis (APD) for end-stage renal disease?".
b: Rate ratios and 95% confidence intervals were estimated from the Poisson regression
c: Frequency of searching was self reported as the number of times per month physicians search a bibliographic database

Precision

The results from the multivariable regression testing the relationship between nephrologist characteristics and search precision are presented in Tables 5-16 and 5-17. The analysis of question one suggests that previous training in literature searching is an independent predictor of improved precision (RR: 2.3; 95% CI 1.4 to 3.6). This association, however, was not confirmed in the analysis for the second question; although the point estimate supported the magnitude and direction of the association (RR: 2.5; 95% CI: 0.8 to 7.5).

Table 5-16: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on search precision for question 1[a]

Variable	Estimate of rate ratio[b] (95% Confidence Interval)	p-value
Frequency of searching[c]	1.02 (0.99 to 1.05)	0.118
Previous training in literature searching (referent group: No)	2.27 (1.43 to 3.62)	<0.001

a:Question 1: "How effective and safe are statins for renal and cardiovascular outcomes in each stage of chronic kidney disease (pre-dialysis, dialysis, and transplantation)?"
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression adjusting for years of nephrology practice and sex
c: Frequency of searching was self reported as the number of times per month physicians search a bibliographic database

Table 5-17: Results from the negative binomial regression analysis assessing the effect of nephrologist characteristics on precision for question 2[a]

Variable	Estimate rate ratio[b] (95% Confidence intervals)	p-value
Frequency of searching[c]	0.98 (0.89 to 1.08)	0.683
Previous training in literature searching (referent group: No)	2.46 (0.80 to 7.50)	0.114

a:Question 2: "What are the benefits and harms of continuous ambulatory peritoneal dialysis (CAPD) versus automated peritoneal dialysis (APD) for end-stage renal disease?".
b: Rate ratios and 95% confidence intervals were estimated from the negative binomial regression adjusting for age
c: Frequency of searching was self reported as the number of times per month physicians search a bibliographic database

5.2.4 Additional analyses

For each model, diagnostics were assessed and additional sensitivity analyses were conducted. Diagnostics included assessment of residuals, collinearity and outliers. Assessments of model fit through residual analyses for all models are presented in Appendices 7-14. There was minimal collinearity between predictors used in all the models as indicated by r-values of less than 0.6[112].

Sensitivity analyses included removing outliers, categorizing predictor variables and imputing missing values, such as assigning a zero precision to the invalid values. No sensitivity analyses substantively change the estimates of the models presented here and in no cases changed the direction of an effect.

5.3 Objective 2: Impact of search filters on search query performance

Survey responses for Group 2 (where each nephrologist was presented with a unique clinical question, randomly selected from the 100 eligible reviews) were used for Objective 2. SAS code for this analysis is presented in Appendix 20. The Wilcoxon Signed-Rank statistic was used to evaluate differences between the physician-provided searches and the filter aided searches. A Bonferroni adjusted alpha value of 0.0015 was adopted for this analysis; significant p-values are indicated in bold in the tables to follow. As a confirmatory analysis, the Sign test was also calculated for the differences in sensitivity and precision (see Section 4.3.2). Results indicated minimal differences in the Wilcoxon Signed-Rank and Sign tests, but in some cases the Sign test appeared to be less conservative (declaring a significant value in favour of the filtered search which was not significant with the Wilcoxon Signed-Rank test).

5.3.1 Primary and Secondary analyses: Using unmodified search queries

Appendix 3 provides details of the 100 search questions used in the survey. As stated in the methods (Section 4.3.1), for the primary analysis, the only adjustment made to the physician-provided searches was the addition of appropriate PubMed searching syntax for the MeSH terms and limits in cases where a physician responded they used such features in their search. This was done for five search queries received.

Primary analysis: Analyzing all returned citations

For each physician-provided search query, 18 searches were executed in PubMed (one physician-provided search unaided by filters and 17 filter-aided searches). The mean and median sensitivity and precision of the 18 different searches are presented in Table 5-18 (see Appendix 21 for further details). Descriptively, physician-provided search queries exhibited a median sensitivity of 25% (half the search queries retrieved over 25% of the relevant articles) and a median precision of 1% (1 in 100 articles retrieved by the searches were considered relevant). After applying the filters, median sensitivity ranged from 14% to 54% and median precision ranged from 1% to 9%.

Table 5-19 presents the mean and median differences in sensitivity and precision between the physician-provided searches and the filter-aided searches. When considering the filters alone, sensitivity was most improved after applying the renal 'content' broad filter, while precision significantly decreased. Precision was most improved after applying the 'methods' narrow filter, while sensitivity significantly decreased. The 'methods' narrow filter and the 'content' narrow filter produced the best combined improvement: a 5.5% median improvement in precision (99% CI: 2% to 12%) and sensitivity remained unchanged. Expressing this improvement in precision another way, the ratio of relevant to non-relevant articles improved from 1 in 100 with the non-filtered search to 1 in 12 when both filters were used in combination. No filters produced significant simultaneous improvements in both sensitivity and precision. The addition of the journal filter did not produce noteworthy improvements over the 'methods' and 'content' filters.

Table 5-18: Search performance of physician-provided searches and searches aided by filters

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
Physician-provided search (unaided)					P	5.3%	1.1%
					S	37.5%	25.0%
x					P	5.5%	1.5%
					S	36.7%	25.0%
	x				P	22.5%	8.8%
					S	31.5%	18.6%
		x			P	4.2%	0.8%
					S	50.2%	53.6%
			x		P	5.4%	1.0%
					S	48.0%	48.5%
x		x			P	4.4%	1.2%
					S	49.5%	50.0%
x			x		P	5.6%	1.7%
					S	47.3%	45.8%
	x	x			P	18.1%	6.4%
					S	42.4%	33.3%
	x		x		P	20.3%	8.5%
					S	40.7%	33.3%

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
				x	P	7.7%	1.3%
					S	34.0%	20.0%
x				x	P	7.9%	1.8%
					S	33.1%	20.0%
	x			x	P	22.8%	8.1%
					S	28.9%	14.3%
		x		x	P	5.8%	1.2%
					S	45.4%	40.8%
			x	x	P	7.0%	1.6%
					S	43.5%	33.3%
x		x		x	P	6.0%	1.7%
					S	44.7%	36.7%
x			x	x	P	7.3%	2.3%
					S	42.8%	33.3%
	x	x		x	P	19.1%	6.6%
					S	39.0%	29.3%
	x		x	x	P	20.2%	7.6%
					S	37.5%	28.6%

Table 5-19: Change in search performance between filtered and non-filtered physician-provided searches

Methods Filter		Content Filter		Journal Filter	Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)					
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean difference***	***% of queries improvement seen***	***Median difference (99% CI)***	***p-value Wilcoxon***	***p-value Sign test***
x					P	0.2%	66	0.18 (0.03 to 0.52)	**<0.0001**	**<0.0001**
					S	-0.9%	1	0.00 (0.00 to 0.00)	0.0469	0.1250
	x				P	17.2%	65	6.49 (0.82 to 14.42)	**<0.0001**	**<0.0001**
					S	-6.0%	0	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
		x			P	-1.1%	23	-0.02 (-0.19 to 0.00)	**0.0002**	**0.0008**
					S	12.7%	32	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
			x		P	0.1%	42	0.00 (-0.01 to 0.02)	0.6142	0.7376
					S	10.4%	29	0.00 (0.00 to 0.00)	**0.0001**	0.0017
x		x			P	-0.9%	43	0.00 (0.00 to 0.42)	0.8679	0.5764
					S	12.0%	32	0.00 (0.00 to 0.00)	**<0.0001**	**0.0001**
x			x		P	0.3%	58	0.04 (0.00 to 0.42)	0.0191	**0.0002**
					S	9.8%	29	0.00 (0.00 to 0.00)	**0.0002**	0.0115
	x	x			P	12.8%	71	4.39 (0.96 to 10.11)	**<0.0001**	**<0.0001**
					S	4.9%	30	0.00 (0.00 to 0.00)	0.1149	0.4101
	x		x		P	15.0%	72	5.56 (1.97 to 12.36)	**<0.0001**	**<0.0001**
					S	3.1%	28	0.00 (0.00 to 0.00)	0.3848	0.8919

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)					
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean difference***	***% of queries improvement seen***	***Median difference (99% CI)***	***P-value Wilcoxon***	***p-value Sign test***
				x	P	2.4%	66	0.32 (0.02 to 0.83)	**<0.0001**	**<0.0001**
					S	-3.6%	0	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
x				x	P	2.6%	66	0.66 (0.07 to 1.69)	**<0.0001**	**<0.0001**
					S	-4.4%	0	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
	x			x	P	17.5%	65	6.29 (0.97 to 15.27)	**<0.0001**	**<0.0001**
					S	-8.6%	0	0.00 (-0.05 to 0.00)	**<0.0001**	**<0.0001**
		x		x	P	0.4%	51	0.00 (0.00 to 0.27)	0.1669	0.0352
					S	7.9%	30	0.00 (0.00 to 0.00)	0.0227	0.3317
			x	x	P	1.7%	58	0.04 (0.00 to 0.55)	0.0061	**0.0002**
					S	6.0%	27	0.00 (0.00 to 0.00)	0.1193	0.8899
x		x		x	P	0.7%	59	0.11 (0.00 to 0.66)	0.0057	**0.0001**
					S	7.2%	30	0.00 (0.00 to 0.00)	0.0366	0.4101
x			x	x	P	2.0%	62	0.35 (0.00 to 1.20)	**0.0004**	**<0.0001**
					S	5.3%	27	0.00 (0.00 to 0.00)	0.1794	1.0000
	x	x		x	P	13.7%	70	4.86 (0.32 to 10.96)	**<0.0001**	**<0.0001**
					S	1.5%	26	0.00 (0.00 to 0.00)	0.9762	0.4350
	x		x	x	P	14.9%	69	4.96 (0.53 to 11.85)	**<0.0001**	**<0.0001**
					S	0.0%	24	0.00 (0.00 to 0.00)	0.6304	0.1925

Secondary analysis: Analyzing the top 40 returned citations

The mean and median sensitivity and precision values of the 18 different searches, when restricting the results to the top 40 citations are presented in Table 5-20 (see Appendix 22). Descriptively, physician provided search queries exhibited a median sensitivity of 0% (half the search queries retrieved 0 relevant articles within the first 40 citations) and thus, a median precision of 0%. After applying the filters, median sensitivity ranged from 0% to 14% and median precision ranged from 0% to 13%.

Table 5-21 presents the mean and median differences in sensitivity and precision between the physician-provided searches and the filter aided searches when restricted to the top 40 returned citations. When considering the filters used alone, sensitivity and precision was maximally improved after applying the 'methods' narrow filter (sensitivity median difference: 0%, 99% CI: 0% to 11%; precision median difference 10%, 99% CI: 3% to 17%). The combined use of the 'methods' narrow filter and the 'content' narrow filter produced the greatest combined improvement; a 0% median improvement in sensitivity (99% CI: 0% to 17%; p-value <0.001) and an 8% median improvement in precision (99% CI: 3% to 13%; p-value <0.001).

Table 5-20: Search performance of physician-provided searches and searches aided by filters, when restricted to the top 40 returned results

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
Physician-provided search (unaided)					P	4.6%	0.0%
					S	9.3%	0.0%
x					P	5.1%	0.0%
					S	10.3%	0.0%
	x				P	22.9%	12.5%
					S	22.7%	13.7%
		x			P	3.9%	0.0%
					S	8.3%	0.0%
			x		P	5.1%	0.0%
					S	9.8%	0.0%
x		x			P	4.0%	0.0%
					S	9.7%	0.0%
x			x		P	5.3%	0.0%
					S	10.9%	0.0%
	x	x			P	18.4%	6.3%
					S	23.4%	11.6%
	x		x		P	20.5%	7.5%
					S	25.1%	14.3%

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
				x	P	7.0%	0.0%
					S	9.8%	0.0%
x				x	P	7.7%	0.0%
					S	11.1%	0.0%
	x			x	P	23.2%	10.0%
					S	21.6%	11.1%
		x		x	P	5.6%	0.0%
					S	10.6%	0.0%
			x	x	P	6.8%	0.0%
					S	11.0%	0.0%
x		x		x	P	5.8%	0.0%
					S	11.5%	0.0%
x			x	x	P	7.3%	2.5%
					S	12.6%	3.8%
	x	x		x	P	19.3%	7.5%
					S	22.9%	11.6%
	x		x	x	P	20.5%	7.6%
					S	23.7%	12.3%

Table 5-21: Change in search performance between filtered and non-filtered physician-provided searches, when restricted to the top 40 returned results

Methods Filter		**Content Filter**		**Journal Filter**	**Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)**					
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean difference***	***% of queries improvement seen***	***Median difference (99% CI)***	***p-value Wilcoxon***	***p-value Sign test***
x					P	0.5%	27	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
					S	0.9%	11	0.00 (0.00 to 0.00)	**0.0010**	**0.0010**
	x				P	18.4%	63	10.00 (2.50 to 17.39)	**<0.0001**	**<0.0001**
					S	13.3%	42	0.00 (0.00 to 11.11)	**<0.0001**	**<0.0001**
		x			P	-0.7%	7	0.00 (0.00 to 0.00)	0.0333	0.0125
					S	-1.0%	7	0.00 (0.00 to 0.00)	0.1310	0.0931
			x		P	0.6%	16	0.00 (0.00 to 0.00)	0.8475	1.0000
					S	0.5%	12	0.00 (0.00 to 0.00)	0.8899	1.0000
x		x			P	-0.5%	16	0.00 (0.00 to 0.00)	0.6802	1.0000
					S	0.4%	11	0.00 (0.00 to 0.00)	0.7632	0.8238
x			x		P	0.8%	25	0.00 (0.00 to 0.00)	0.1856	0.1081
					S	1.6%	18	0.00 (0.00 to 0.00)	0.2206	0.1849
	x	x			P	13.8%	63	5.00 (2.50 to 11.11)	**<0.0001**	**<0.0001**
					S	14.1%	48	0.00 (0.00 to 11.11)	**<0.0001**	**<0.0001**
	x		x		P	16.0%	65	7.50 (2.5 to 12.50)	**<0.0001**	**<0.0001**
					S	15.7%	49	0.00 (0.00 to 16.67)	**<0.0001**	**<0.0001**

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)					
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean difference***	***% of queries improvement seen***	***Median difference (99% CI)***	***P-value Wilcoxon***	***p-value Sign test***
				x	P	2.4%	32	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
					S	0.5%	16	0.00 (0.00 to 0.00)	0.0568	0.0118
x				x	P	3.1%	40	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
					S	1.8%	22	0.00 (0.00 to 0.00)	0.0033	**0.0002**
	x			x	P	18.7%	63	8.39 (2.50 to 17.65)	**<0.0001**	**<0.0001**
					S	12.3%	43	0.00 (0.00 to 9.09)	**<0.0001**	**<0.0001**
		x		x	P	1.0%	27	0.00 (0.00 to 0.00)	0.0819	0.0237
					S	1.2%	18	0.00 (0.00 to 0.00)	0.2223	0.1221
			x	x	P	2.3%	31	0.00 (0.00 to 0.00)	0.0063	0.0054
					S	1.7%	21	0.00 (0.00 to 0.00)	0.1677	0.0708
x		x		x	P	1.3%	32	0.00 (0.00 to 0.00)	0.0065	0.0037
					S	2.2%	25	0.00 (0.00 to 0.00)	0.0219	0.0046
x			x	x	P	2.7%	37	0.00 (0.00 to 0.00)	**0.0001**	**0.0001**
					S	3.3%	29	0.00 (0.00 to 0.00)	0.0041	**0.0001**
	x	x		x	P	14.8%	62	6.61 (0.00 to 12.50)	**<0.0001**	**<0.0001**
					S	13.5%	48	0.00 (0.00 to 14.29)	**<0.0001**	**<0.0001**
	x		x	x	P	15.9%	64	7.50 (2.5 to 12.50)	**<0.0001**	**<0.0001**
					S	14.4%	50	0.00 (0.00 to 14.29)	**<0.0001**	**<0.0001**

5.3.2 Additional analyses: Using modified search queries

For this analysis, the physician-provided search queries were modified using the pre-specified rules listed in Table 4-7. This was done as it was observed that some physician-provided search queries included spelling errors and other discrepancies. Details of the analysis when considering all returned citations and only the top 40 citations are presented in Appendix 23. The use of modified search queries produced the same patterns of results as those previously observed for the unmodified queries (Sections 5.3.1). While modifications improved the search performance of the initial physician searches (median sensitivity of 42% for modified searches vs. 25% for unmodified searches; median precision of 1% for modified searches vs. 2% for unmodified searches) the impact of the best-performing filter-combinations remained similar. The combination of the 'methods' narrow and 'content' narrow filters produced the best improvements in sensitivity and precision.

CHAPTER 6: DISCUSSION & CONCLUSIONS

6.1 Summary of principal findings

Searching for evidence is a key step in the practice of EBM. Physicians regularly have clinical questions for which they do not know the answer. Questions can be answered through the use of PubMed; however, physicians find literature searching challenging and often lack the time and skills to efficiently identify relevant articles in a timely manner. Thus, helping health professionals keep up-to-date with the latest advances has the potential to improve the transfer of research into practice, medical decision-making, health care delivery and patient outcomes. To remedy the challenge of literature searching, this thesis sought to ascertain the current state of searching by nephrologists and identify methods by which to improve this process.

6.1.1 Objective 1: Determinants of search success

Using search queries provided by 60 nephrologists for two distinct clinical questions, this study determined whether there is an association between search query or nephrologist characteristics and the search outcomes of sensitivity and precision. Analyses were conducted separately for each characteristic type.

The search query characteristics of interest included the number of concept terms specified in a search query and the use of multifaceted search features (e.g. use of MeSH terms, Boolean logic operators, search limits). Other factors of the search query considered for their potential confounding effects included the use terms embodying any of the PICO (Patient, Intervention, Control, Outcome) aspects, the use of an acronym or quotations.

Analyses for both clinical questions indicated that the addition of each concept term to a search query (combined with an implicit or explicit ‘AND’ operator), decreases search sensitivity and increases search precision. In contrast, the use of multifaceted search features improves both sensitivity and precision. The relationship for precision was only statistically significant for question one, although, the point estimate for the second question supported both the magnitude and direction of this association.

When considering the effects of nephrologist characteristics on search success, the primary factors of interest were the frequency of searching, as indicated by the number of times per month a physician searches a bibliographic database, and whether a physician had received previous training in literature searching. Other characteristics considered for their potential confounding effects included age, sex, the number of years a physician had been practicing nephrology, whether the physician practiced in an academic setting or whether the physician had previously used advanced search features when searching. All values were self reported by the physicians.

The relationship between nephrologist characteristics and the search outcomes were less clear than for the search query characteristics. No relationship was evident between the characteristics and the outcome of sensitivity. However, nephrologists who had previously received training in literature searching exhibited improved precision when compared to their counterparts. This observed relationship was statistically significant for question one, but not for question two; although, the direction and magnitude of the point estimate for question two was supportive of the association.

6.1.2 Objective 2: Impact of search filters on search query performance

To address the second objective, the use of PubMed search filters was investigated to determine whether they improve searches provided by physicians. Three types of PubMed filters were tested: 'methods', 'content' and 'journal'. Each filter was applied to 100 search queries provided by nephrologists that targeted 100 unique clinical questions. The search outcomes, sensitivity and precision, were compared between the filter-aided searches and non-filtered searches. The results indicated that PubMed filters can improve search performance. When evaluating all results provided by PubMed, the combinations of the 'methods' narrow filter (Clinical Queries therapy filter) and 'content' narrow filter (Nephrology filter) produced the greatest improvement in search performance. While sensitivity (the number of relevant articles retrieved) remained relatively constant, the precision (proportion of articles retrieved that were relevant) improved noticeably. No filters simultaneously improved both sensitivity and precision.

Results from the survey indicated that 80% of respondents do not scan more than two search pages in PubMed when conducting a search (equivalent to the first 40 results retrieved). Thus, a secondary analysis was performed while restricting the PubMed results to the first 40 articles. Analyses showed that the use of the ‘methods’ narrow filter alone maximally improved both search sensitivity and precision.

6.2 Contribution of this work to the current literature

To the best of our knowledge this is the first study to examine, both conceptually and empirically, the association between physician characteristics or search query characteristics and search success in PubMed (measured as search sensitivity and precision). Using the method of conceptual model evaluation[108], we identified two methodological considerations that have not been addressed in the past. First, an examination of the conceptual model revealed a potential mediating effect between nephrologist characteristics and search outcomes; while nephrologists create search queries to be used in PubMed, it is the characteristics of the query that immediately impact the performance of the search, not the physician characteristics. Accordingly, the search query characteristics may act as a mediator between the physician characteristics and the search outcomes. Including both factors in a regression model, consequently, can lead to an over-adjustment and inaccurate results[111]. Second, the model evaluation indicated that two factors a) the clinical question being addressed and b) its associated set of relevant articles may act as confounders in the relationship between the search query characteristics (or physician characteristics) and the outcomes of searching. Studies can consider ways to control for this confounding effect either at the design phase or the analysis phase

To mitigate the potential for meditation, we conducted analyses separately for each of the characteristic types. Further, to prevent the confounding effects of the nature of different clinical questions on the outcomes of search success, 60 nephrologists were provided with the same two clinical questions and analyses were performed separately for each clinical question. To support the generalizability of our findings, the results proved to be robust and were, for most part, consistent across the two clinical questions.

Our findings, conversely, were less consistent with two previous studies that attempted to evaluate whether sensitivity and precision varied in relation to physician characteristics or search query characteristics[50;53]. These studies, however, were limited by the confounding nature of the clinical questions tested[53], and small number of queries used[50]. Also, the one study evaluating the effect of experience with searching was conducted in 1986[50], long before the advent of the Internet and the popularity of online literature searching. Nonetheless, our findings support the results seen in other studies that evaluated other definitions of search success (e.g. likelihood of viewing an abstract, answering a question correctly).

We found that the increase in the number of concept terms used in a search query is associated with a reduced sensitivity, but improved precision. This finding would be expected and follows from the mechanism of searching with Boolean logic operators. As was outlined in the introduction section (Section 2.8), the use of an ‘AND’ operator in a search query produces a narrow search, whereby in most situations precision increases and sensitivity often decreases. Thus, with each addition of a concept term to a search the changes in sensitivity and precision are further magnified.

While previous studies have considered evaluating whether the use of terms corresponding to the PICO factors impacts searching[45;53], none have considered the impact of acronyms in a query. While this factor was not of primary interest to this study, our results provide an indication that the use of acronyms can severely reduce the sensitivity of a search. Such an effect may occur when an acronym term is not mapped to a MeSH term during the query pre-processing stage and instead the term is included as a text word in the search. This process would result in PubMed only finding articles that include the specific acronyms in the abstract or title and would preclude relevant articles that do not include the acronym. For example, one acronym often used in the searches received for the first clinical question was ‘ckd’. This acronym, which refers to the condition “chronic kidney disease”, is not recognized by PubMed. However, the acronym ‘capd’ that many clinicians used for the second clinical question is recognized by PubMed and is mapped to "peritoneal dialysis, continuous ambulatory"[MeSH

Terms]. This might explain why the inclusion of an acronym term was found to significantly impact the searches for the first question, but not the second question.

We also demonstrated that the use of an outcome term (such as “outcome”, “mortality”, “safety” or “efficacy”) also reduces search sensitivity. This effect occurs in a similar manner as the inclusion of an acronym as only articles that mention such terms in their abstract or title will be found by PubMed.

While not confirmed from the analyses for both searches, there is an indication that previous training in literature searching can produce more precise searches. This finding may speak to the positive effect training has on physicians.

To our knowledge, this is also the first study to develop and consider a testing framework for evaluating PubMed search filters for use by clinician searchers (Table 2-4). To date, researchers have developed, optimized and validated search filters in closed, experimental environments (stage one and two in the staged program of research, see Table 2-4). This study moves beyond developing filters[90] to testing their functionality in the “real world” context of physician searching. This has only been attempted three times in the past[44;45;53], and ours is the first study to test the utility of ‘methods’, ‘content’ and ‘journal’ filters in combination. The three previous evaluations compared the use of the Clinical Queries ‘methods’ filters to standard searches in PubMed[45] and Google Scholar[44] and to the use of search limits in PubMed[53]. While the first two studies found that filters improved search precision, the conclusions are tempered by the small number of searches conducted and limited number of clinical topics tested. The second study also used searches developed by the researchers, which may generalize less well to searches conducted by physicians in a busy clinical setting. In addition, while the third study used a large variety of searches and clinical topics (100 each), this study also used searches developed by the researchers and did not compare the filtered searches to non-filtered searches.

6.3 Recognized limitations

6.3.1 Determining article relevancy

There is no perfect, easily applied measure to determine whether an article is relevant to a focused clinical question. This, in fact, is a challenge for most search evaluation studies. And since the choice of a reference standard directly impacts generalizability of sensitivity and precision, the choice deserves special consideration. This study chose to use primary articles identified in systematic reviews as an external measure of relevance. All other articles were viewed as non-relevant. The reviews selected came from the EvidenceUpdates service, which pre-screens and identifies systematic reviews and meta-analyses from over 130 journals that meet strict methodological criteria and have a high potential for clinical relevance[100]. These reviews thus provide a clinically important problem and include a comprehensive search for relevant studies to address the problem. It is recognized, however, that some practitioners may consider additional articles, such as commentaries, narrative reviews, case reports, and animal studies as relevant when searching. However, by using systematic reviews to define relevance, this study engaged in the widely accepted principles of the hierarchy of evidence to identify the most important primary articles to retrieve in a search. Such a method has also previously been used in other searching studies [34;44;52;53;91;98;119].

6.3.2 Performance metrics

The study used sensitivity and precision as metrics to determine how well reference sets of relevant articles are retrieved. Some have claimed that these are misleading surrogate outcomes and that other more relevant outcomes would be desired[77]; for example, assessing whether a search can provide a physician with the ability to come up with the correct answer (better knowledge), whether this will change medical decisions or processes of care, and whether this can improve patient outcomes. As described in the background (Sections 2.4.3 & 2.7.1), the current study represents a key milestone in a staged program of research, to guide the development and execution of future studies (Table 2-4).

6.3.3 Searching is a dynamic process

This study used the search queries provided by physicians to ascertain search sensitivity and precision. It is most likely that these queries were the initial searches the physicians would attempt. Evidence that physicians provided untested search queries was apparent as some queries included misspelled terms and some retrieved no articles in PubMed. In truth, searching is a dynamic process; an unsuccessful search is often tried again using different terms. An analysis of PubMed query logs over a one month period in 2008[62] indicated that when queries did not return any results, 82% of the users searched again with a modified or new search query. Also, in 41% of searches, users performed subsequent searches even when their initial search returned results.

To combat this limitation, other research frameworks were considered, such as surveillance of local nephrologists using PubMed filters in practice or in a laboratory setting. However, those frameworks also have their limitations. For reasons of feasibility, the study thus obtained the *initial* search queries provided by a random sample of nephrologists practicing in academic and non-academic settings across Canada.

It is unlikely that the use of these searches impacted the internal validity of the evaluations presented here. However, this may temper the generalizability of the findings to only the initial searches created by specialist physicians, nephrologists in particular. In an effort to mimic how a physician might improve a search (s)he was not satisfied with, for Objective 2, the search queries were modified using pre-specified rules and were re-evaluated. The analyses indicated that while modifications improved the initial non-filtered searches, the conclusions of filter impact were unchanged. Consequently, this study has shown that the use of filters can improve search query performance of initial searches; this may potentially obviate the need for additional searches by physicians, saving time and reducing frustration.

6.3.4 Target audience is nephrologists

The thesis focused on nephrologists for five reasons: 1) The purpose was to test the application of the 'content' filters that were designed to identify articles relevant for the care of renal patients; 2) Nephrologists as specialists are interested in identifying and reviewing primary studies for focused questions in renal medicine; 3) The systematic reviews identified through the EvidenceUpdates database are primarily targeted at physicians; 4) A list of nephrologists in Canada was compiled; 5) This study received support from strong knowledge translation partners in nephrology (Canadian Society of Nephrology, Kidney Foundation of Canada).

Although nephrologists acted as the primary study group, the same principles and procedures can serve other health care providers, patients, managers and policy makers who need to be informed about best evidence-informed care for renal disorders. Having shown that the filters make an impact, it is hoped that this study serves as a proof-of-concept and that future research continues with evaluating filter use in other medical disciplines.

6.3.5 Self-administered survey

All data for this thesis was initially captured through a self-completed questionnaire. As with all surveys, the study is limited by the correctness of the responses provided by the physicians. The primary predictors for the first objective included a) the characteristics of the search query provided, which may not be an accurate representation of the queries physicians use in practice, b) frequency of searching, which may suffer from reporting or recall errors and c) whether physicians received previous training in literature searching, which may also suffer from reporting errors .

The survey presented physicians with an artificial, though plausible, searching situation by providing them with a clinical question and requested that they develop a search query. As stated earlier, searching is a dynamic process and it is unlikely that physicians spent a large amount of time developing the search queries they provided on the survey. Thus, these searches may not truly represent their searching abilities. Evidence that physicians provided untested search queries was apparent as some queries included misspelled terms and some retrieved no articles in PubMed. This discrepancy between what physicians provided and what they may do in practice may have contributed to the negative result observed for the relationship between the nephrologist characteristics and the outcomes of the searches. There is thus a concern that errors in the responses may have affected the internal validity of the results when identifying nephrologist characteristics associated with successful searches. The misrepresented search queries, however would not have affected the internal validity of the analyses testing the relationship between search query characteristics and the outcomes. Instead, as discussed earlier (see Section 6.3.3), this may impede the generalizability of the results.

In addition, the errors present in the responses for frequency of searching and previous training in literature searching may also have hindered the internal validity of the regression analyses when testing the associations between nephrologist characteristics and search outcomes. Finally, by using a self-reported, cross-sectional survey, the study was unable to capture other potential factors that may influence search ability, such as an understanding of physicians' time constraints or cognitive abilities. This omission may have led to residual confounding in the analyses.

6.3.6 Non-responder bias

A comparison of responders to non-responders identified some differences. Responders were more likely to be male than female and varied response patterns were observed from the different provinces. The lowest response rate was received from Quebec, however, this is unlikely the true response rate as some physicians may have been French-speaking and were, in fact, ineligible for the survey (the survey was provided in English only). In addition, while our sample included a large proportion of males (77%), this was comparable to our sampling frame (72% male).

While the differences between responders and non-responders should not impact the internal validity of this study, they may temper the generalizability of the findings, specifically when analyzing the impact of nephrologist characteristics on search success, as the respondents may be more proficient at searching, compared to non-responders.

6.3.7 Statistical modeling decisions

As is a concern in all analysis, there is a possibility that the regression models used for Objective 1 were mis-specified leading to erroneous conclusions. However, numerous attempts were made to ensure this was not the case. For all models, descriptive statistics were analyzed for all covariates to identify potential data entry errors, model assumptions were assessed (e.g. independence, linearity, homoscedasticity), appropriate diagnostics were executed and evaluated, collinearity of all included covariates was checked, and sensitivity analyses were performed while imputing missing values and adding and removing candidate outliers. In addition, in most cases, the effects observed were confirmed between the analyses for question one and two, providing evidence of the robustness of the findings.

One of the eight models, however, (model 7: the association between nephrologist characteristics and search sensitivity) continued to exhibit poor model specification after following the procedures outlined above (see details in Appendix 13).

6.4 Future directions

This thesis successfully builds on findings of previous studies on the topic of literature searching by addressing some of their methodological limitations. Although, to more accurately capture physician characteristics and their search queries, a longitudinal surveillance-type (versus cross-sectional) study would present a better methodological approach. Such a design could address the limitations identified in this study (namely, reporting and recall errors, issues of data quality and residual confounding). With such a study, physicians could be observed as they performed literature searches in PubMed, thus establishing a better understanding of the searches they would use in practice. A design observing physicians over time would also allow other physician characteristics to be quantified, such as physicians' cognitive and spatial visualization abilities[43] and would provide an indication of the effects of time-dependent factors, such as increasing searching skills, and changes in caseload and case-mix. Unfortunately, the costs and time to complete such a study precluded this design herein.

This study rigorously considered one form of search success, search sensitivity and precision. As described in the background (Section 2.4.3), the most prominent and important outcome of a search for physicians is the ability to identify an answer, rooted in evidence, for a clinical question of interest. This outcome, however, is further downstream from the initial search executed in a bibliographic database and is dependent on first, being able to retrieve relevant literature and other explanatory factors, such as the ability of clinicians to critically appraise and interpret evidence. Evaluating the ability to answer clinical questions without first establishing physicians' ability to retrieve pertinent literature would therefore be premature. Having now established the factors and methods that impact search sensitivity and precision, future research can follow with evaluating other forms of search success while considering dependent explanatory factors.

For the first objective, this study evaluated search queries provided for two clinical questions. The questions differed on the renal populations of interest (chronic kidney disease vs. peritoneal dialysis) and the number of associated relevant articles (49 vs. 4). While the analysis revealed many similar predictive characteristics (number of concept terms, multifaceted searching and training in literature) other predictive factors differed, mainly the use of an acronym term. This suggests that the content area of the question may influence the searching factors that can lead to successful searching. Future studies should explore this by testing various questions and other subject areas.

To date, numerous search filters aimed at improving PubMed searching have been developed and tested for various clinical disciplines. While these filters show promise in test environments, studies that test their use in improving end-user searching are limited. This research has established that the use of PubMed filters, when applied to physician-provided searches for renal evidence, can improve search success. This study can now serve as a proof of concept for testing filter use in other subject areas (such as cardiology) and for other audiences.

For reasons of feasibility, our research focused on questions of therapy as most systematic reviews pertain to prevention and treatment. As more systematic reviews for diagnosis, prognosis and etiology are published, future research should expand this study to include other clinical study areas. Finally, searching for primary literature in PubMed is only one part of the knowledge acquisition process for physicians. Future studies should evaluate other sources of information such as synopses and syntheses.

6.5 Study significance

6.5.1 Implications for searching physicians

The results of this study support physicians' use of the 'methods' (Clinical Queries filter) already available for use through PubMed. Currently this feature is infrequently used. Among the survey respondents, only 21% indicate using Clinical Queries, despite 92% using PubMed in the past year to assist in patient treatment. The use of the Clinical Queries filters is strongly recommended by proponents of EBM and is often taught at literature searching courses for clinicians. A simple search in Google using the terms "PubMed Clinical Queries" reveals numerous resources from around the world emphasizing the utility of the search filters for clinician use. These filters can be used by all physicians, regardless of their clinical specialty. While this study examined questions related to renal medicine (as it also tested the Nephrology 'content' filter), the superior performance of the Clinical Queries filters used alone suggests that the results may be generalizable to other medical disciplines, not just nephrology.

6.5.2 Implications for developers of bibliographic databases

As stated earlier, the Clinical Queries filters have the potential to improve everyday clinical searching. While these filters are already available in PubMed, access to them is difficult and few clinicians consequently use the features. Currently, to use the filters, searchers must access the PubMed home page to link out to a different interface that provides access to the Clinical Queries feature; this is suggestive of a disjoint between the use of filters and PubMed searching. In addition, queries created from the Clinical Queries interface do not have the opportunity for sophistication (such as the use of Limits) afforded to search queries developed from the PubMed home page. Given the potential impact of these filters, this study promotes a call-to-action for PubMed to make these filters more easily accessible to the lay user. This could be achieved by adding a search filters section into the 'Limits', rather than providing a different search interface.

In addition, for individuals searching for renal clinical evidence, this study demonstrates that the combination of Clinical Queries 'methods' and Nephrology 'content' filters maximize precision when considering all search results. Thus, the Nephrology filters need to be integrated directly into PubMed to be of most use to these searchers. Prior to initiating the thesis, Ms. Ione Auston, a librarian at the NLM who oversaw the integration of the Clinical Queries filters into the PubMed interface, was contacted. Ms. Auston indicated an interest in this research project. She will now be contacted again and provided the results from this research.

6.5.3 Implications for literature training workshops

Curricula of literature training sessions to date are varied in content and are not rooted in empirical evidence of PubMed searching[60]. Instead, they follow from the understanding of database logic and information retrieval technology[120]. This study presents four conclusions that can enhance searching and thus be promoted in training workshops. First, this study demonstrates that increasing the number of concept terms included in a search query decreases sensitivity, but increases precision (as would be expected). Second, multifaceted searching can improve both sensitivity and precision. Contrary to what is currently being taught, the inclusion of the outcome term may severely impede sensitivity. This study also suggests that the use of acronyms not recognized by PubMed may also reduce sensitivity. Finally, this study confirms that, when used properly, combinations of the 'methods' and 'content' filters maximize precision of searches without decreasing sensitivity and for quick clinical searches, the 'methods' Clinical Queries filter alone can maximize both sensitivity and precision within the first 40 retrieved results.

6.6 Conclusions

Given the large volume of searches conducted by physicians, the use of search filters can assist clinicians worldwide to search more effectively, in less time and with less frustration. For quick clinical searches the Clinical Queries filter, available for all to use, can improve physician searches. Searching can also be improved by incorporating multifaceted search features such as MeSH terms, search limits or the Boolean 'OR' operator to search queries. Findings from this study may have important clinical implications as efficient retrieval of the best available evidence is used to inform clinical care protocols, clinical decision making for patient care and medical education. Educational strategies should adopt the conclusions drawn from this study in teaching search query development and emphasize the use of search filters in everyday clinical practice.

CHAPTER 7: BIBLIOGRAPHY

(1) Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. *ACP J Club* 2002;136:A11-A14.

(2) Garg AX, Iansavichus AV, Wilczynski NL et al. Filtering Medline for a clinical discipline: diagnostic test assessment framework. *BMJ* 2009;339:b3435.

(3) Key MEDLINE Indicators. *National Library of Medicine (US)* [serial online] 2009; Accessed May 10, 2010.

(4) 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;1:447-458.

(5) Jenkins M. Evaluation of methodological search filters-a review. *Health Information and Libraries Journal* 2004;21:148-163.

(6) PubMed Clinical Queries. *National Library of Medicine (US)* [serial online] 2008; Accessed May 10, 2010.

(7) Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.

(8) Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ* 2004;328:1040.

(9) Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Med* 2004;2:23.

(10) Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. *AMIA Annu Symp Proc* 2003;719-723.

(11) Garg AX, Iansavichus AV, Kastner M et al. Lost in publication: Half of all renal practice evidence is published in non-renal journals. *Kidney International* 2006;70:1995.

(12) Dickinson HD. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. Evidence-Based Medicine Working Group. *JAMA* 1992;268:2420-2425.

(13) Straus SE, Richardson WS, Glasziou P, Haynes RB. *Evidence-based medicine: how to practice and teach EBM*. Third ed. Churchill Livingstone, 2005.

(14) Ely JW, Osheroff JA, Ebell MH et al. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ* 2002;324:710.

(15) Chambliss ML, Conley J. Answering clinical questions. *J Fam Pract* 1996;43:140-144.

(16) Norlin C, Sharp AL, Firth SD. Unanswered questions prompted during pediatric primary care visits. *Ambul Pediatr* 2007;7:396-400.

(17) Davies K. The information-seeking behaviour of doctors: a review of the evidence. *Health Info Libr J* 2007;24:78-94.

(18) Geddes J. Asking structured and focused clinical questions: essential first step of evidence-based practice. *Evidence Based Mental Health* 1999;2:35.

(19) Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995;123:A12-A13.

(20) Detailed Indexing Statistics: 1965-2010. *National Library of Medicine (US)* [serial online] 2010; Accessed May 10, 2010.

(21) Ely JW, Osheroff JA, Ebell MH et al. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999;319:358-361.

(22) Currie LM, Graham M, Allen M. Clinical information needs in context: an observational study of clinicians while using a clinical information system. *AMIA Annu Symp Proc* 2003;2003:190-194.

(23) Gorman PN, Helfand M. Information Seeking in Primary Care: How Physicians Choose Which Clinical Questions to Pursue and Which to Leave Unanswered. *Medical Decision Making* 1995;15:113.

(24) Gorman PN, Yao P, Seshadri V. Finding the answers in primary care: information seeking by rural and nonrural clinicians. *Medinfo* 2004;11:1133-1137.

(25) Covell DG, Uman GC, Manning PR. Information needs in office practice: are they being met? *Ann Intern Med* 1985;103:596-599.

(26) Green ML, Ciampi MA, Ellis PJ. Residents medical information needs in clinic: are they being met? *Am J Med* 2000;109:218-223.

(27) Gorman P. Does the medical literature contain the evidence to answer the questions of primary care physicians? Preliminary findings of a study. *Proc Annu Symp Comput Appl Med Care* 1993;571-575.

(28) Gonzalez-Gonzalez AI, Dawes M, Sanchez-Mateos J et al. Information needs and information-seeking behavior of primary care physicians. *Ann Fam Med* 2007;5:345-352.

(29) Ely JW, Osheroff JA, Chambliss ML, Ebell MH, Rosenbaum ME. Answering physicians' clinical questions: obstacles and potential solutions. *J Am Med Inform Assoc* 2005;12:217-224.

(30) Gorman PN, Ash J, Wykoff L. Can primary care physicians' questions be answered using the medical journal literature? *Bull Med Libr Assoc* 1994;82:140-146.

(31) Alper BS, White DS, Ge B. Physicians answer more clinical questions and change clinical decisions more often with synthesized evidence: a randomized trial in primary care. *Ann Fam Med* 2005;3:507-513.

(32) Del Mar CB, Silagy CA, Glasziou PP et al. Feasibility of an evidence-based literature search service for general practitioners. *Med J Aust* 2001;175:134-137.

(33) Lindberg DA, Siegel ER, Rapp BA, Wallingford KT, Wilson SR. Use of MEDLINE by physicians for clinical problem solving. *JAMA* 1993;269:3124-3129.

(34) Sampson M, Zhang L, Morrison A et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Med Res Methodol* 2006;6:33.

(35) Weatherall DJ, Ledingham JG, Warrell DA. On dinosaurs and medical textbooks. *Lancet* 1995;346:4-5.

(36) Oxman AD, Guyatt GH. The science of reviewing research. *Ann N Y Acad Sci* 1993;703:125-133.

(37) Schaafsma F, Verbeek J, Hulshof C, van DF. Caution required when relying on a colleague's advice; a comparison between professional advice and evidence from the literature. *BMC Health Serv Res* 2005;5:59.

(38) Two-thirds of European Physicians Agree the Internet is Essential to Their Practices. 12-12-2005. PRNewswire.

(39) Bennett NL, Casebeer LL, Kristofco RE, Strasser SM. Physicians'Internet information-seeking behaviors. *J Contin Educ Health Prof* 2004;24:31-38.

(40) Masters K. For what purpose and reasons do doctors use the Internet: A systematic review. *Int J Med Inform* 2008;77:4-16.

(41) Hersh WR, Crabtree MK, Hickam DH et al. Factors Associated with Success in Searching MEDLINE and Applying Evidence to Answer Clinical Questions. *J Am Med Inform Assoc* 2002;9:283.

(42) Alper BS, Stevermer JJ, White DS, Ewigman BG. Answering family physicians clinical questions using electronic medical databases. *J Fam Pract* 2001;50:960-965.

(43) Sackett DL, Straus SE. Finding and applying evidence during clinical rounds: the "evidence cart". *JAMA* 1998;280:1336-1338.

(44) Anders ME, Evans DP. Comparison of PubMed and Google Scholar literature searches. *Respir Care* 2010;55:578-583.

(45) Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak* 2007;7:16.

(46) Gruppen LD, Rana GK, Arndt TS. A controlled comparison study of the efficacy of training medical students in evidence-based medicine literature searching skills. *Acad Med* 2005;80:940-944.

(47) Rosenberg WM, Deeks J, Lusher A, Snowball R, Dooley G, Sackett D. Improving searching skills and evidence retrieval. *J R Coll Physicians Lond* 1998;32:557-563.

(48) Erickson S, Warner ER. The impact of an individual tutorial session on MEDLINE use among obstetrics and gynaecology residents in an academic training programme: a randomized trial. *Med Educ* 1998;32:269-273.

(49) McKibbon KA, Haynes RB, Johnston ME, Walker CJ. A study to enhance clinical end-user MEDLINE search skills: design and baseline findings. *Proc Annu Symp Comput Appl Med Care* 1991;73-77.

(50) McKibbon KA, Haynes RB, Dilks CJ et al. How good are clinical MEDLINE searches? A comparative study of clinical end-user and librarian searches. *Comput Biomed Res* 1990;23:583-593.

(51) Schaafsma F, Hugenholtz N, de BA, Smits P, Hulshof C, van DF. Enhancing evidence-based advice of occupational health physicians. *Scand J Work Environ Health* 2007;33:368-378.

(52) Freeman MK, Lauderdale SA, Kendrach MG, Woolley TW. Google Scholar versus PubMed in locating primary literature to answer drug-related questions. *Ann Pharmacother* 2009;43:478-484.

(53) Yousefi-Nooraie R, Irani S, Mortaz-Hedjri S, Shakiba B. Comparison of the efficacy of three PubMed search filters in finding randomized controlled trials to answer clinical questions. *J Eval Clin Pract* 2010.

(54) Schaafsma F, Hulshof C, de BA, van DF. Effectiveness and efficiency of a literature search strategy to answer questions on the etiology of occupational diseases: a controlled trial. *Int Arch Occup Environ Health* 2007;80:239-247.

(55) Demner-Fushman D, Hauser SE, Humphrey SM, Ford GM, Jacobs JL, Thoma GR. MEDLINE as a source of just-in-time answers to clinical questions. *AMIA Annu Symp Proc* 2006;190-194.

(56) Schwartz K, Northrup J, Israel N, Crowell K, Lauder N, Neale AV. Use of on-line evidence-based resources at the point of care. *Fam Med* 2003;35:251-256.

(57) McKibbon KA, Fridsma DB. Effectiveness of clinician-selected electronic information resources for answering primary care physicians' information needs. *J Am Med Inform Assoc* 2006;13:653-659.

(58) Westbrook JI, Coiera EW, Gosling AS. Do online information retrieval systems help experienced clinicians answer clinical questions? *J Am Med Inform Assoc* 2005;12:315-321.

(59) Garg A, Turtle KM. Effectiveness of training health professionals in literature search skills using electronic health databases--a critical appraisal. *Health Info Libr J* 2003;20:33-41.

(60) Brettle A. Information skills training: a systematic review of the literature. *Health Info Libr J* 2003;20 Suppl 1:3-9.

(61) Hoogendam A, Stalenhoef AF, Robbe PF, Overbeke AJ. Analysis of queries sent to PubMed at the point of care: observation of search behaviour in a medical teaching hospital. *BMC Med Inform Decis Mak* 2008;8:42.

(62) Islamaj DR, Murray GC, Neveol A, Lu Z. Understanding PubMed user search behavior through log analysis. *Database (Oxford)* 2009;2009:bap018.

(63) Guyatt GH, Haynes RB, Jaeschke RZ et al. Users' Guides to the Medical Literature: XXV. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group. *JAMA* 2000;284:1290-1296.

(64) The Cochrane Library. Product Descriptions: Record Counts. *The Cochrane Collaboration* [serial online] 2009; Accessed February 20, 2009.

(65) Grol R, Cluzeau FA, Burgers JS. Clinical practice guidelines: towards better quality guidelines and increased international collaboration. *Br J Cancer* 2003;89 Suppl 1:S4-S8.

(66) Haycox A, Bagust A, Walley T. Clinical guidelines-the hidden costs. *BMJ* 1999;318:391-393.

(67) Amerling R, Winchester JF, Ronco C. Guidelines have done more harm than good. *Blood Purif* 2008;26:73-76.

(68) Levin A. Practice guidelines do improve patient outcomes: association or causation? *Blood Purif* 2008;26:67-72.

(69) Coyne DW. Practice recommendations based on low, very low, and missing evidence. *Clin J Am Soc Nephrol* 2007;2:11-12.

(70) Narins RG, Bennett WM. Patient care guidelines: problems and solutions. *Clin J Am Soc Nephrol* 2007;2:1-2.

(71) Amerling R, Winchester JF, Ronco C. Guidelines for guidelines. *Blood Purif* 2007;25:36-38.

(72) Uhlig K, Balk EM, Lau J, Levey AS. Clinical practice guidelines in nephrology--for worse or for better. *Nephrol Dial Transplant* 2006;21:1145-1153.

(73) Shekelle PG, Ortiz E, Rhodes S et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? *JAMA* 2001;286:1461-1467.

(74) 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.

(75) Cullen R. The medical specialist: information gateway or gatekeeper for the family practitioner. *Bull Med Libr Assoc* 1997;85:348-355.

(76) Hull D. Using statistical testing in the evaluation of retrieval experiments. *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval* 1993;329-338.

(77) Hersh WR. *Information Retrieval: A Health and Biomedical Perspective*. Third ed. Springer, 2008.

(78) Bardia A, Wahner-Roedler DL, Erwin PL, Sood A. Search strategies for retrieving complementary and alternative medicine clinical trials in oncology. *Integr Cancer Ther* 2006;5:202-205.

(79) Deville WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;53:65-69.

(80) Fraser C, Murray A, Burr J. Identifying observational studies of surgical interventions in MEDLINE and EMBASE. *BMC Med Res Methodol* 2006;6:41.

(81) Kirpalani H, Schmidt B, McKibbon KA, Haynes RB, Sinclair JC. Searching MEDLINE for randomized clinical trials involving care of the newborn. *Pediatrics* 1989;83:543-546.

(82) Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002;31:150-153.

(83) Royle P, Bain L, Waugh N. Systematic reviews of epidemiology in diabetes: finding the evidence. *BMC Med Res Methodol* 2005;5:2.

(84) Royle PL, Waugh NR. Making literature searches easier: a rapid and sensitive search filter for retrieving randomized controlled trials from PubMed. *Diabet Med* 2007;24:308-311.

(85) Shaw RL, Booth A, Sutton AJ et al. Finding qualitative research: an evaluation of search strategies. *BMC Med Res Methodol* 2004;4:5.

(86) Watson RJ, Richardson PH. Identifying randomized controlled trials of cognitive therapy for depression: comparing the efficiency of Embase, Medline and PsycINFO bibliographic databases. *Br J Med Psychol* 1999;72:535-542.

(87) Watson RJ, Richardson PH. Accessing the literature on outcome studies in group psychotherapy: the sensitivity and precision of Medline and PsycINFO bibliographic database searching. *Br J Med Psychol* 1999;72:127-134.

(88) Wilczynski NL, Haynes RB, Lavis JN, Ramkissoonsingh R, Arnold-Oatley AE. Optimal search strategies for detecting health services research studies in MEDLINE. *CMAJ* 2004;171:1179-1185.

(89) Zhang L, Ajiferuke I, Sampson M. Optimizing search strategies to identify randomized controlled trials in MEDLINE. *BMC Med Res Methodol* 2006;6:23.

(90) InterTASC Information Specialists' Sub-Group: Search Filter Resource. *Centre for Reviews and Dissemination, University of York* [serial online] 2010; Accessed September 8, 2010.

(91) Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *J Clin Epidemiol* 2005;58:560-567.

(92) Number of Titles Currently Indexed for Index Medicus and MEDLINE on PubMed. *National Library of Medicine (US)* [serial online] 2009; Accessed June 1, 2009.

(93) Hersh WR, Hickam DH. How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review. *JAMA* 1998;280:1347-1352.

(94) Buckland M, Gey F. The relationship between recall and precision. *Journal of the American Society for Information Science* 1994;45:12-19.

(95) Gordon M, Kochen M. Recall-precision trade-off: A derivation. *Journal of the American Society for Information Science* 1989;40:145-151.

(96) Gordon M, Pathak P. Finding Information on the World Wide Web: The Retrieval Effectiveness of Search Engines. *Inform Process Manag* 1999;35:141-180.

(97) Oppenheim C, Morris A, McKnight C, Lowley S. The evaluation of WWW search engines. *J Doc* 2000;56:190-211.

(98) Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials (Cochrane Methodology Review). *The Cochrane Library* 2003;2.

(99) Fidel R, Soergel D. Factors affecting online bibliographic retrieval: a conceptual framework for research. *J Am Soc Inf Sci* 1983;34.

(100) Haynes RB. bmjupdates+, a new FREE service for evidence-based clinical practice. *Evidence-Based Medicine* 2005;10:35.

(101) Birck R, Krzossok S, Markowetz F, Schnulle P, van der Woude FJ, Braun C. Acetylcysteine for prevention of contrast nephropathy: meta-analysis. *Lancet* 2003;362:598-603.

(102) Dillman DA, NetLibrary I. *Mail and Internet surveys: the tailored design method*. Second ed. Wiley New York: 2007.

(103) Burns KE, Duffett M, Kho ME et al. A guide for the design and conduct of self-administered surveys of clinicians. *CMAJ* 2008;179:245-252.

(104) The Royal College of Physicians and Surgeons of Canada: Directory of Fellows. *National Library of Medicine (US)* [serial online] 2008; Accessed May 10, 2008.

(105) College of Physicians and Surgeons - Provincial Offices. *National Library of Medicine (US)* [serial online] 2008; Accessed May 10, 2008.

(106) MD Select: Canadian Medical Directory. *Don Mills (ON): Southam Information Products, Ltd* [serial online] 2008; Accessed May 10, 2008.

(107) PubMed Help. *National Library of Medicine* [serial online] 2010; Accessed January 3, 2011.

(108) Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37-48.

(109) Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Third ed. Philadelphia, USA: Lippincott Williams & Wilkins, 2008.

(110) Kleinbaum DG, Kupper LL, Muller KE, Nizam A. *Applied Regression Analysis and Other Multivariable Methods. 3rd ed.* Belmont: Duxbury Press, 1998.

(111) Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. Second ed. Sudbury, Massachusetts: Jones and Bartlett Publishers, 2007.

(112) Vittinghoff E, Glidden DV. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. New York: Springer Science+Business Media, Inc., 2005.

(113) Hilbe J. *Negative binomial regression*. Cambridge: Cambridge University Press New York, 2007.

(114) Agresti A. *Categorical data analysis*. 2 ed. Hoboken: John Wiley & Sons, Inc., 2002.

(115) Westfall PH, Tobias RD, Rom D, Wolfinger RD, Hochberg Y. *Multiple Comparisons and Multiple Tests Using SAS*. Cary, NC: SAS Institute Inc., 1999.

(116) Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995;310:170.

(117) D'Agostino RB, Sullivan LM, Beiser AS. Introductory applied biostatistics. *Brooks Cole* 2005.

(118) Hahn GJ, Meeker WQ. *Statistical intervals: a guide for practitioners*. Wiley New York, 1991.

(119) Slobogean GP, Verma A, Giustini D, Slobogean BL, Mulpuri K. MEDLINE, EMBASE, and Cochrane index most primary studies but not abstracts included in orthopedic meta-analyses. *J Clin Epidemiol* 2009;62:1261-1267.

(120) Zhou W, Smalheiser NR, Yu C. A tutorial on information retrieval: basic terms and concepts. *J Biomed Discov Collab* 2006;1:2.

(121) Sandhu S, Wiebe N, Fried LF, Tonelli M. Statins for improving renal outcomes: a meta-analysis. *J Am Soc Nephrol* 2006;17:2006-2016.

(122) Fabrizi F, Dixit V, Magnini M, Elli A, Martin P. Meta-analysis: intradermal vs. intramuscular vaccination against hepatitis B virus in patients with chronic kidney disease. *Aliment Pharmacol Ther* 2006;24:497-506.

(123) Landoni G, Biondi-Zoccai GG, Tumlin JA et al. Beneficial impact of fenoldopam in critically ill patients with or at risk for acute renal failure: a meta-analysis of randomized clinical trials. *Am J Kidney Dis* 2007;49:56-68.

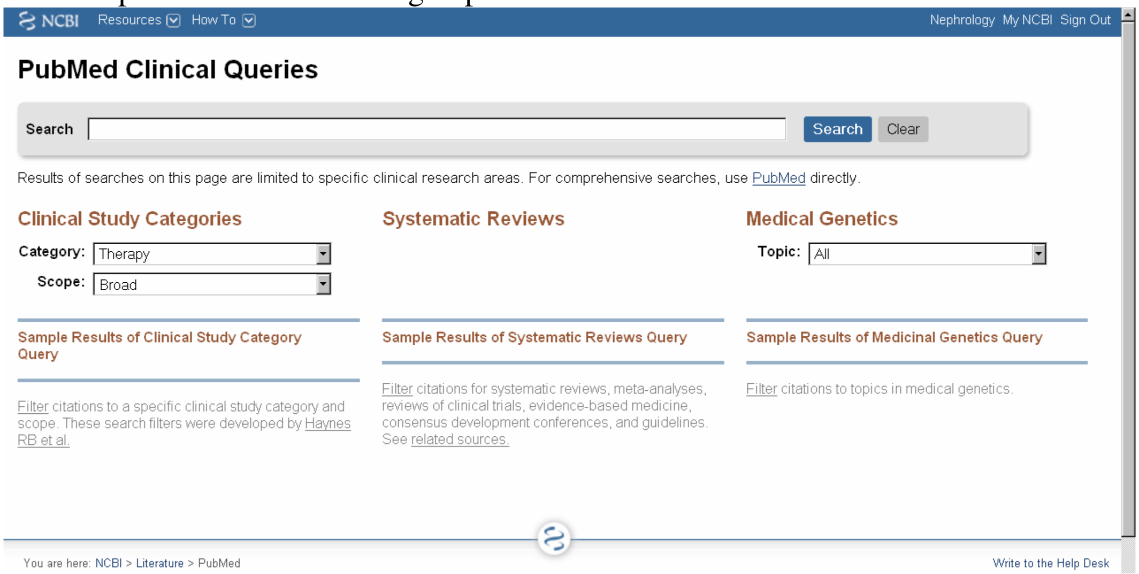
(124) Friedrich JO, Adhikari N, Herridge MS, Beyene J. Meta-analysis: low-dose dopamine increases urine output but does not prevent renal dysfunction or death. *Ann Intern Med* 2005;142:510-524.

(125) Webster A, Woodroffe RC, Taylor RS, Chapman JR, Craig JC. Tacrolimus versus cyclosporin as primary immunosuppression for kidney transplant recipients. *Cochrane Database Syst Rev* 2005;CD003961.

CHAPTER 8: APPENDICES

Appendix 1: Snapshot of PubMed Clinical Queries

High performance methodological search filters developed by the Health Information Research Unit at McMaster University are publicly available on PubMed.
URL: http://www.ncbi.nlm.nih.gov/pubmed/clinical



Appendix 2: Procedure for determining whether an article contains renal information

- Assessments should be made using the full-text of each article.
- An article contains renal information if any content area described in Part A is the main purpose / focus / patient population of the article.
- If the only mention of a content area in Part A is in the patient exclusion criteria, then the article does not contain renal information.
- Content areas described in Part B, by themselves, are NOT considered to be renal information unless they are the main focus / purpose / patient population of the article and a content term from Part A is also mentioned.

Part A

<table>
<tr><th colspan="3">Content Areas</th></tr>
<tr><td rowspan="2">1. Kidney Failure:
• End-stage renal disease (ESRD)
• End-stage renal failure
• End-stage kidney failure
• Nephrotoxicity
• Renal toxicity
• Renal failure
• Uremia/ Uremic</td><td>a) Acute renal failure</td><td>• Acute kidney failure
• Acute kidney injury
• Acute renal injury
• Acute kidney insufficiency
• Acute renal insufficiency
• Acute tubular necrosis
• Acute on chronic renal failure
• Prerenalazotemia
• Prerenal disease</td></tr>
<tr><td>b) Acute or chronic dialysis</td><td>• Artificial kidney
• Continuous renal replacement therapy
• Dialy* (dialysis, dialysate, dialyser, dialyzer, dialysis adequacy, dialysis solutions)
• Extracorporeal dialysis
• Hemodialysis (any type – conventional hemodialysis, nocturnal hemodialysis, daily hemodialysis, frequent hemodialysis, quotidian dialysis, venovenous hemodialysis, also spelled haemo-)
• Hemodiafiltration
• Hemofiltration
• Home dialysis
• Peritoneal dialysis (any type – continuous ambulatory peritoneal dialysis, automated peritoneal dialysis i.e. cycler, cycling, tidal peritoneal dialysis)
• Peritoneal equilibration test
• Renal dialysis
• Slow continuous ultrafiltration</td></tr>
</table>

	c) Kidney transplant recipients, kidney transplantation	• Allograft (*must be kidney)* (i.e. kidney allograft, renal allograft). • Donor or donation *(must be kidney)* (i.e. cadaveric kidney donor, deceased kidney donor, donor nephrectomy, extended criteria donor, expanded criteria donor, kidney donor, living kidney donor, live kidney donor, living kidney donation, non-heart beating donor, renal donor) • Graft or grafting (*must be kidney graft*) (i.e. delayed graft function, graft failure, kidney graft, renal graft etc.) • Kidney rejection • Recipient (*must be kidney)* (i.e. renal recipient, kidney recipient) • Solid organ transplantation *(must include kidney transplant patients, most solid organ transplants are kidneys, although they can be heart, liver, lung etc.)* • Transplant or transplantation(*must be kidney)* (i.e. kidney transplant, kidney transplantation, renal transplant, renal transplantation) • Xenotransplantation (*must be kidney, not another organ*)
	d)Metabolic, inflammatory conditions which are associated with renal disease	The study must describe these conditions in the context of renal patients having/ developing them. For example: • Acidosis • Anemia • Calciphylaxis • Hyperparathyroidism • Osteodystrophy
2. Chronic Kidney Disease:	a) Reduced kidney function .	• *[Kidney function is a common test performed in many types of studies. Measuring kidney function by itself does not make a study eligible unless some of the patients assessed have reduced kidney function or kidney function was a main study outcome]* • Chronic kidney disease • Chronic renal failure • Chronic renal insufficiency • Elevated creatinine (> 135 umol/L [1.5 mg/dL] in men, > 105 umol/L [1.2 mg/dL] in women)

		• Glomerular filtration rate < 60 mL/min per 1.73 m^2, • Pre-dialysis • Reduced (kidney or renal) creatinine clearance • Reduced glomerular filtration rate • Reduced kidney function
	b) Proteinuria	• *[Microalbuminuria by itself is not eligible]* • Overt proteinuria • Random urine albumin to creatinine ratio > 33 mg/mmol • 24 hour urine protein > 300 mg / day • 24 hour urine albumin > 300 mg/day
	c) For outcome studies	May be monitoring changes in kidney function / proteinuria, or development of new onset reduced kidney function or proteinuria
4) Glomerular Diseases: • Glomerulo-nephropathy • Glomerulopathy • Glomerulonephritis • Glomerulosclerosis • Glomerular diseases • Glomerulo-nephritides • Nephropathy	a) Nephrotic/Nephritic syndrome	
	b) Biopsy classification of glomerulo-nephritis	• Anti-GBM antibody disease • Balkan nephropathy • Bright's disease • Benign familial hematuria • Collapsing glomerulosclerosis • Focal (segmental) glomerulosclerosis • Glomerularhematuria • Hereditary nephritis • IgA Nephropathy (aka Berger's disease) • Membranoproliferativeglomerulonephritis • Membranous nephropathy • Mesangiocapillaryglomerulonephritis • Minimal change disease • Post-infectious glomerulonephritis • Post-streptococcal glomerulonephritis • Rapidly progressive glomerulonephritis • Thin basement membrane disease
5) Other Renal Pathology	• Diabetic nephropathy – Diabetic glomerulosclerosis / nephrosclerosis – Diabetic glomerulopathy – Kimmelstiel-Wilson disease / syndrome / nephropathy – Nodular/ Intercapillaryglomerulosclerosis • Cystinosis • Nephrosclerosis - Hypertensive /vascular nephrosclerosis - Kidney / renal nephrosclerosis / arteriosclerosis • Hemolytic uremic syndrome	

	• Interstitial nephritis • Lupus nephritis – Lupus / lupoid nephritis / glomerulonephritis – Lupus nephropathy • Medullary cystic kidney disease • Myeloma kidney / cast nephropathy • Obstructive uropathy / obstructive nephropathy • Polycystic kidney disease • Renal agenesis • Renal atheroemboli • Renal papillary necrosis • Renal tubulointerstitial disease • Renal tubular acidosis • Renovascular / renal hypertension – Renal artery disease – Renal artery stenosis – Scleroderma renal crisis • Thrombotic thrombocytopenia purpura	
6) Vesicoureteral reflux and reflux nephropathy (VUR)		
7) Metabolic acid/base, water disturbance	• Diabetes insipidius (nephrogenic, *NOT central only*) • Renal tubular acidosis • Metabolic acidosis • Metabolic alkalosis	
8) Procedures/ care for kidney patients other than dialysis:	a) Personnel	• Nephrology • Nephrologists • Renal care
	b) Procedure related to dialysis or kidney biopsy	• Hemodialysis catheter insertion (*<u>NOT</u> catheters used for reasons other then dialysis*). • Hemodialysis vascular access • Hemodiaylsis fistula • Hemodialysis (access) graft

Part B

Content Areas	
Systemic illness or genetic conditions which are almost always associated with renal disease:	• Alport's Syndrome • Anti-GBM antibody disease • Goodpasture's Syndrome
Systemic illness or agentswhich have historically been associated with renal diseases *(Note: this list is not exhaustive)*:	• Amyloidosis • Anti-neutrophilcytoplasmic antibody / ANCA • Churg Strauss Syndrome • Cryoglobulinemia • Cystinosis • Endocarditis • Fabry's disease • Henoch-Schönleinpurpura • Hyperoxaluria • HIV • Lithium • Microscopic polyangitis • Multiple myeloma • Non-steroidal anti-inflammatories • Paraproteinemia • Polyarteritisnodosa • Rheumatoid arthritis • Sepsis • Sickle cell disease • Systemic lupus erythematosus • Vasculitis • von Hippel-Lindau disease
Plasma exchange / plasmapheresis / Apheresis	*(relevant if used to treat renal conditions)*
Blood electrolyte disturbances	• Hyper/Hypocalcemia • Hyper/Hypokalemia • Hyper/Hypomagnesiumemia • Hyper/Hyponatremia • Hyper/Hypophosphatemia • Hyperuricemia • Hyperhomocysteinemia • Metabolic acidosis • Metabolic alkalosis • Respiratory acidosis • Respiratory alkalosis
Hematuria (by itself without any other relevant renal terms) Polyuria (by itself without any other relevant renal terms)	
Kidney neoplasms / cancers /tumors /	• Renal cell carcinoma / adenocarcinoma

carcinoma	• Nephroid carcinoma • Wilms tumor • Mesoblasticnephroma
Kidney stones Procedure related to renal stone treatment	• Calcium stones • Calculi (renal) • Cystine stones • Kidney calculi / calculus • Medullary sponge kidney • Nephrolithiasis • Lithotripsy (*NOT gall stones*) • Phosphate calculi • Renal calculi • Staghorn calculi • Struvite stones • Uric acid stones
Ureter or prostate disease	
Urinary tract infection or pyelonephritis	
Pregnancy	• Eclampsia • Preeclampsia

Appendix 3: Details of included systematic reviews

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Interferon monotherapy for dialysis patients with chronic hepatitis C: an analysis of the literature on efficacy and safety.	Russo,M.W.; Goldsweig,C.D.; Jacobson,I.M.; Brown,R.S.,Jr.	2003	12873587	11	11	What is the efficacy and safety of interferon monotherapy in dialysis patients with chronic hepatitis C virus (HCV)?
Laparoscopic donor nephrectomy.	Handschin,A.E.; Weber,M.; Demartines,N.; Clavien,P.A.	2003	14598409	32	32	How does laparoscopic nephrectomy compare to open nephrectomy in terms of donor safety and efficacy?
Interleukin-2 receptor monoclonal antibodies in renal transplantation: meta-analysis of randomised trials.	Adu,D.; Cockwell,P.; Ives,N.J.; Shaw,J.; Wheatley,K.	2003	12689974	8	8	What is the effect of interleukin-2 receptor monoclonal antibodies on renal graft survival, post-transplant malignancy and infectious complications?
Acetylcysteine for prevention of contrast nephropathy: meta-analysis.	Birck,R.; Krzossok,S.; Markowetz,F.; Schnulle,P.; van der Woude,F.J.; Braun,C.	2003	12944058	7	7	Does prophylactic use of acetylcysteine reduce the incidence of contrast nephropathy in patients with renal insufficiency?
Meta-analysis of randomized clinical trials on the usefulness of acetylcysteine for prevention of contrast nephropathy.	Isenbarger,D.W.; Kent,S.M.; O'Malley,P.G.	2003	14675586	7	7	What is the treatment effect of N-acetylcysteine (NAC) for contrast nephropathy (CN) prevention?
N-acetylcysteine for the prevention of contrast-induced nephropathy. A systematic review and meta-analysis.	Liu,R.; Nair,D.; Ix,J.; Moore,D.H.; Bent,S.	2005	15836554	9	8	Does administering N-acetylcysteine around the time of contrast administration reduce the risk of contrast induced nephropathy.?
Theophylline for prevention of contrast-induced nephropathy: a systematic review and meta-analysis.	Bagshaw,S.M.; Ghali,W.A.	2005	15911721	9	9	What is the effect of theophylline on contrast induced nephropathy (CIN)?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Is acetylcysteine effective in preventing contrast-related nephropathy? A meta-analysis.	Nallamothu,B.K.; Shojania,K.G.; Saint,S.; Hofer,T.P.; Humes,H.D.; Moscucci,M.; Bates,E.R.	2004	15629733	20	15	Is acetylcysteine effective in preventing contrast-related nephropathy?
Antiviral medications to prevent cytomegalovirus disease and early death in recipients of solid-organ transplants: a systematic review of randomised controlled trials.	Hodson,E.M.; Jones,C.A.; Webster,A.C.; Strippoli,G.F.; Barclay,P.G.; Kable,K.; Vimalachandra,D.; Craig,J.C.	2005	15964447	16	14	Does antiviral prophylaxis reduce the clinical syndrome associated with cytomegalovirus infection?
Evidence-based systematic literature review of hemoglobin/hematocrit and all-cause mortality in dialysis patients.	Volkova,N.; Arab,L.	2006	16377382	18	17	What is the relationship between hemoglobin and/or hematocrit values and all-cause mortality in dialysis patients?
The role of osmolality in the incidence of contrast-induced nephropathy: a systematic review of angiographic contrast media in high risk patients.	Solomon,R.	2005	16221227	17	17	Are iso-osmolality contrast media (CM) associated with less nephrotoxicity compared to all low-osmolality CM?
Meta-analysis: low-dose dopamine increases urine output but does not prevent renal dysfunction or death.	Friedrich,J.O.; Adhikari,N.; Herridge,M.S.; Beyene,J.	2005	15809463	60	51	What is the efficacy of low-dose dopamine (<5 mcg/kg of body weight per minute) compared with no therapy in patients with or at risk for acute renal failure?
Statins for improving renal outcomes: a meta-analysis.	Sandhu,S.; Wiebe,N.; Fried,L.F.; Tonelli,M.	2006	16762986	27	25	What are the effects of statins on change in kidney function and urinary protein excretion?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Meta-analysis: the efficacy of strategies to prevent organ disease by cytomegalovirus in solid organ transplant recipients.	Kalil,A.C.; Levitsky,J.; Lyden,E.; Stoner,J.; Freifeld,A.G.	2005	16365468	10	10	What is the efficacy of universal prophylaxis and preemptive approaches in preventing cytomegalovirus (CMV) organ disease and other complications in solid organ transplant recipients?
Meta-analysis: the adjuvant role of thymopentin on immunological response to hepatitis B virus vaccine in end-stage renal disease.	Fabrizi,F.; Dixit,V.; Martin,P.	2006	16696803	6	6	What is the efficacy and safety of thymopentin-adjuvanted hepatitis B (HB) vaccine in chronic dialysis patients?
Meta-analysis: intradermal vs. intramuscular vaccination against hepatitis B virus in patients with chronic kidney disease.	Fabrizi,F.; Dixit,V.; Magnini,M.; Elli,A.; Martin,P.	2006	16886915	12	11	How does intradermal vs. intramuscular hepatitis B vaccine compare regarding response rate among chronic kidney disease patients?
Meta-analysis: terlipressin therapy for the hepatorenal syndrome.	Fabrizi,F.; Dixit,V.; Martin,P.	2006	16948805	10	10	What is the efficacy and safety of terlipressin in the treatment of hepatorenal syndrome?
Beneficial impact of fenoldopam in critically ill patients with or at risk for acute renal failure: a meta-analysis of randomized clinical trials.	Landoni,G.; Biondi-Zoccai,G.G.; Tumlin,J.A.; Bove,T.; De,Luca M.; Calabro,M.G.; Ranucci,M.; Zangrillo,A.	2007	17185146	16	12	What is the impact of fenoldopam on acute kidney injury, patient mortality, and length of hospital stay in critically ill patients?
Meta-analysis: the effect of statins on albuminuria.	Douglas,K.; O'Malley,P.G.; Jackson,J.L.	2006	16847294	15	15	Do statins affect albuminuria?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Combination therapy with an angiotensin receptor blocker and an ACE inhibitor in proteinuric renal disease: a systematic review of the efficacy and safety data.	MacKinnon,M.; Shurraw,S.; Akbari,A.; Knoll,G.A.; Jaffey,J.; Clark,H.D.	2006	16797382	21	21	What is the safety and efficacy of combination therapy with an ACE inhibitor and an ARB in patients with chronic proteinuric renal disease?
Extracorporeal blood purification therapies for prevention of radiocontrast-induced nephropathy: a systematic review.	Cruz,D.N.; Perazella,M.A.; Bellomo,R.; Corradi,V.; de,Cal M.; Kuang,D.; Ocampo,C.; Nalesso,F.; Ronco,C.	2006	16931209	8	8	Does periprocedural extracorporeal blood purification prevent radiocontrast-induced nephropathy?
Systematic review of the effectiveness of preventing and treating Staphylococcus aureus carriage in reducing peritoneal catheter-related infections.	McCormack,K.; Rabindranath,K.; Kilonzo,M.; Vale,L.; Fraser,C.; McIntyre,L.; Thomas,S.; Rothnie,H.; Fluck,N.; Gould,I.M.; Waugh,N.	2007	17580002	22	17	What is the clinical effectiveness of alternative strategies for the prevention and eradication of Staphylococcus aureus carriage in patients on peritoneal dialysis (PD)?
Meta-analysis of frusemide to prevent or treat acute renal failure.	Ho,K.M.; Sheridan,D.J.	2006	16861256	9	9	What are the benefits and harms of frusemide in acute renal failure and do these effects differ when used to prevent or to treat acute renal failure?
The effectiveness and cost-effectiveness of cinacalcet for secondary hyperparathyroidism in end-stage renal disease patients on dialysis: a systematic review and economic evaluation.	Garside,R.; Pitt,M.; Anderson,R.; Mealing,S.; Roome,C.; Snaith,A.; D'Souza,R.; Welch,K.; Stein,K.	2007	17462168	7	7	What is the efficacy of cinacalcet for the treatment of secondary hyperparathyroidism in people receiving chronic dialysis?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Combination therapy with an ACE inhibitor and an angiotensin receptor blocker for diabetic nephropathy: a meta-analysis.	Jennings,D.L.; Kalus,J.S.; Coleman,C.I.; Manierski,C.; Yee,J.	2007	17367311	10	10	Does combination rennin angiotensin aldosterone system (RAAS)-inhibiting therapy provide greater benefit in diabetic nephropathy (DN) than angiotensin-converting enzyme inhibitor(s) (ACEIs) and angiotensin receptor blocker(s) (ARBs) therapy alone?
Intermittent versus continuous renal replacement therapy for acute renal failure in adults	Rabindranath,K.; Adams,J.; Macleod,A.M.; Muirhead,N.	2007	17636735	15	12	Is intermittent hemodialysis or continuous renal replacement therapy superior in the treatment of acute renal failure (ARF)?
Mortality and target haemoglobin concentrations in anaemic patients with chronic kidney disease treated with erythropoietin: a meta-analysis.	Phrommintikul,A.; Haas,S.J.; Elsik,M.; Krum,H.	2007	17276778	9	9	In the treatment of anaemic chronic kidney disease patients with recombinant human erythropoietin, do different hemoglobin targets alter all-cause mortality or cardiovascular events?
Progression of chronic kidney disease: the role of blood pressure control, proteinuria, and angiotensin-converting enzyme inhibition: a patient-level meta-analysis.	Jafar,T.H.; Stark,P.C.; Schmid,C.H.; Landa,M.; Maschio,G.; de Jong,P.E.; de,Zeeuw D.; Shahinfar,S.; Toto,R.; Levey,A.S.	2003	12965979	12	12	What level of blood pressure and urine protein excretion is associated with the lowest risk for progression of chronic kidney disease in patients with nondiabetic kidney disease during antihypertensive therapy with and without ACE inhibitors?
Meta-analysis of biochemical and patient-level effects of calcimimetic therapy.	Strippoli,G.F.; Palmer,S.; Tong,A.; Elder,G.; Messa,P.; Craig,J.C.	2006	16632010	8	7	What is the added benefit of calcimimetics on secondary hyperparathyroidism (SHPT) in patients with chronic kidney disease on standard therapy with vitamin D and/or phosphate binders?"

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Effects of angiotensin converting enzyme inhibitors and angiotensin II receptor antagonists on mortality and renal outcomes in diabetic nephropathy: systematic review.	Strippoli,G.F.; Craig,M.; Deeks,J.J.; Schena,F.P.; Craig,J.C.	2004	15459003	47	44	What are the effects of angiotensin converting enzyme (ACE) inhibitors and angiotensin II receptor antagonists (AIIRAs) on renal outcomes and all cause mortality in patients with diabetic nephropathy?
Effectiveness of management strategies for renal artery stenosis: a systematic review.	Balk,E.; Raman,G.; Chung,M.; Ip,S.; Tatsioni,A.; Alonso,A.; Chew,P.; Gilbert,S.J.; Lau,J.	2006	17062633	17	17	Is medical therapy as effective as revascularization for atherosclerotic renal artery stenosis?
Angiotensin converting enzyme inhibitors and angiotensin II receptor antagonists for preventing the progression of diabetic kidney disease	Strippoli,G.F.; Bonifati,C.; Craig,M.; Navaneethan,S.D.; Craig,J.C.	2006	17054288	71	68	What are the benefits and harms of angiotensin converting enzyme inhibitors (ACEi) and angiotensin II receptor antagonists (AIIRA) in patients with diabetic kidney disease (DKD), with major focus on renal and cardiovascular outcomes?
Antifungal agents for preventing fungal infections in solid organ transplant recipients	Playford,E.G.; Webster,A.C.; Sorell,T.C.; Craig,J.C.	2004	15266524	4	4	What are the benefits and harms of prophylactic antifungal agents for the prevention of fungal infections in solid organ transplant recipients?
Antimicrobial agents for preventing peritonitis in peritoneal dialysis patients	Strippoli,G.F.; Tong,A.; Johnson,D.; Schena,F.P.; Craig,J.C.	2004	15495124	20	17	What are the benefits and harms of antimicrobial strategies to prevent peritonitis in peritoneal dialysis (PD)?
Antiviral medications for preventing cytomegalovirus disease in solid organ transplant recipients	Hodson,E.M.; Craig,J.C.; Strippoli,G.F.; Webster,A.C.	2008	18425894	16	15	What are the benefits and harms of antiviral medications for preventing symptomatic cytomegalovirus (CMV) disease in solid organ transplant recipients?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Biocompatible hemodialysis membranes for acute renal failure	Alonso,A.; Lau,J.; Jaber,B.L.	2005	15846749	24	14	Do biocompatible membranes (BCM) confer an advantage in either survival or recovery of renal function over the use of bioincompatible membranes (BCIM) in adult patients with acute renal failure requiring intermittent hemodialysis?
Calcimimetics for secondary hyperparathyroidism in chronic kidney disease patients	Strippoli,G.F.; Tong,A.; Palmer,S.C.; Elder,G.; Craig,J.C.	2006	17054287	7	7	What are the benefits and harms of calcimimetics for the prevention of secondary hyperparathyroid bone disease (including osteitis fibrosa cystica and adynamic bone disease) in dialysis patients with chronic kidney disease?
Calcium channel blockers for preventing acute tubular necrosis in kidney transplant recipients	Shilliday,I.R.; Sherif,M.	2005	15846665	9	9	What are the benefits and harms of using calcium channel blockers in the peri-transplant period in patients at risk of acute tubular necrosis (ATN) following kidney transplantation?
Catheter type, placement and insertion techniques for preventing peritonitis in peritoneal dialysis patients	Strippoli,G.F.; Tong,A.; Johnson,D.; Schena,F.P.; Craig,J.C.	2004	15495125	16	15	Which catheter types, placement and insertion techniques, break in periods and immobilisation devices should be used to prevent of peritonitis in peritoneal dialysis (PD) patients?
Cellulose, modified cellulose and synthetic membranes in the haemodialysis of patients with end-stage renal disease	Macleod,A.M.; Campbell,M.; Cody,J.D.; Daly,C.; Grant,A.; Khan,I.; Rabindranath,K.S.; Vale,L.; Wallace,S.	2005	16034894	36	35	Do synthetic membranes offer clinically important advantages compared with standard or modified cellulose membranes in the haemodialysis of patients with end-stage renal disease (ESRD)?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Continuous ambulatory peritoneal dialysis versus automated peritoneal dialysis for end-stage renal disease?	Rabindranath,K.S.; Adams,J.; Ali,T.Z.; Macleod,A.M.; Vale,L.; Cody,J.; Wallace,S.A.; Daly,C.	2007	17443624	8	4	What are the benefits and harms of continuous ambulatory peritoneal dialysis (CAPD) versus automated peritoneal dialysis (APD) for end-stage renal disease?
Double bag or Y-set versus standard transfer systems for continuous ambulatory peritoneal dialysis in end-stage renal disease	Daly,C.; Campbell,M.; Cody,J.; Grant,A.; Vale,L.; Lawrence,P.; Macleod,A.; Wallace,S.; Khan,I.	2001	11406068	18	13	Is there evidence that supports the use of the Y-set (and modifications) or double bag systems versus standard transfer systems for the prevention of peritonitis in peritoneal dialysis (PD) patients?
Effects of nonsteroidal anti-inflammatory drugs on postoperative renal function in adults with normal renal function?	Lee,A.; Cooper,M.G.; Craig,J.C.; Knight,J.F.; Keneally,J.P.	2007	17443518	22	21	What are the effects of nonsteroidal anti-inflammatory drugs (NSAIDs) on postoperative renal function in adults with normal preoperative renal function?
Fish oil for kidney transplant recipients	Lim,A.K.; Manley,K.J.; Roberts,M.A.; Fraenkel,M.B.	2007	17443580	20	14	What are the benefits and harms of fish oil treatment in ameliorating the kidney and cardiovascular adverse effects of CNI-based immunosuppressive therapy in kidney transplant recipients?
HMG CoA reductase inhibitors (statins) for dialysis patients	Navaneethan,S.D.; Nigwekar, S.U., Perkovic, V., Johnson, D.W., Craig, J.C., Strippoli, G.F.M.	2004	15495097	7	5	What are the benefits and harms of statins in dialysis patients?
Haemodiafiltration, haemofiltration and haemodialysis for end-stage kidney disease	Rabindranath,K.S.; Strippoli,G.F.; Daly,C.; Roderick,P.J.; Wallace,S.; Macleod,A.M.	2006	17054289	23	22	What is the comparative efficacy of predominantly convective modes of extracorporeal renal replacement therapy (RRT) with hemodialysis (HD) in patients with end stage kidney disease (ESKD)?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Immunoglobulins, vaccines or interferon for preventing cytomegalovirus disease in solid organ transplant recipients?	Hodson,E.M.; Jones,C.A.; Strippoli,G.F.; Webster,A.C.; Craig,J.C.	2007	17443573	28	25	What are the benefits and harms of immunoglobulins, vaccines or interferon for preventing cytomegalovirus disease in solid organ transplant recipients?
Immunosuppressive agents for treating IgA nephropathy	Baskarat, R., Molony, D.A., Samuels, J.A.	2003	14584001	21	19	What are the benefits and harms of immunosuppression for the treatment of IgA nephropathy?
Immunosuppressive treatment for idiopathic membranous nephropathy in adults with nephrotic syndrome	Schieppati,A.; Perna,A.; Zamora,J.; Giuliano,G.A.; Braun,N.; Remuzzi,G.	2004	15495098	17	15	Is immunosuppressive treatment effective and safe in the treatment of idiopathic membranous nephropathy (IMN) in adults with nephrotic syndrome?
Interleukin 2 receptor antagonists for kidney transplant recipients	Webster,A.C.; Playford,E.G.; Higgins,G.; Chapman,J.R.; Craig,J.	2004	14974043	105	48	What are the benefits and harms of Interleukin 2 receptor antagonists versus standard immunosuppression for kidney transplant recipients when they are added to a standard dual or triple therapy regimen, or used in place of another agent?
Interventions for preventing bone disease in kidney transplant recipients	Palmer,S.C.; McGregor,D.O.; Strippoli,G.F.	2007	17636784	27	18	Do pharmacotherapeutic agents used to treat bone disease following kidney transplantation change the incidence of complications of bone disease, particularly the incidence of fractures?
Low protein diets for chronic kidney disease in non diabetic adults	Fouque,D.; Laville,M.; Boissel,J.P.	2006	16625550	19	9	What is the efficacy of low protein diets in preventing the natural progression of chronic kidney disease towards end-stage renal disease and retard the need for starting maintenance dialysis?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Pre-emptive treatment for cytomegalovirus viraemia to prevent cytomegalovirus disease in solid organ transplant recipients	Strippoli,G.F.; Hodson,E.M.; Jones,C.J.; Craig,J.C.	2006	16437521	6	5	What are the benefits and harms of pre-emptive treatment for cytomegalovirus viraemia to prevent cytomegalovirus disease in solid organ transplant recipients?
Recombinant human erythropoietin for chronic renal failure anaemia in pre-dialysis patients	Cody,J.; Daly,C.; Campbell,M.; Khan I.; Rabindranath,K.; Vale,L.; Wallace,S.; Macleod,A.	2005	16034896	15	9	What are the potential benefits (prevention of kidney failure, improvement of Hb, improvement of QOL) and harms of rHu EPO in pre-dialysis CKD patients?
Routine intraoperative ureteric stenting for kidney transplant recipients	Wilson,C.H.; Bhatti,A.A.; Rix,D.A.; Manas,D.M.	2005	16235385	7	6	What are the benefits and harms of routine ureteric stenting to prevent urological complications in kidney transplants recipients?
Tacrolimus versus cyclosporin as primary immunosuppression for kidney transplant recipients	Webster,A.; Taylor,R.S.; Chapman,J.R.; Craig,J.C.	2005	16235347	30	23	What is the effect on transplant outcomes, toxicity and adverse effects of tacrolimus as compared directly to cyclosporin, in the treatment of kidney transplant recipients?
Target of rapamycin inhibitors (TOR-I; sirolimus and everolimus) for primary immunosuppression in kidney transplant recipients	Webster,A.C.; Lee,V.W.; Chapman,J.R.; Craig,J.C.	2006	16625599	33	14	What are the short and long-term benefits and harms of sirolimus and everolimus when used in primary immunosuppressive regimens for kidney transplant recipients?
Meta-analysis: antibiotics for prophylaxis against hemodialysis catheter-related infections	James MT; Conley J; Tonelli M	2008	18413621	14	11	Do topical or intraluminal antibiotics reduce catheter-related bloodstream infection compared with no antibiotic therapy in adults undergoing hemodialysis?
Continuous versus intermittent renal replacement therapy for critically ill patients with acute kidney injury: a meta-analysis	Bagshaw SM; Berthiaume LR; Delaney A	2008	18216610	8	5	Does continuous, compared with intermittent, renal replacement therapy (RRT) portend any meaningful difference on mortality, renal recovery, or treatment-related complications?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Effects of corticosteroid on Henoch-Schonlein purpura: a systematic review	Weiss PF; Feinstein JA; Luan X	2007	17974746	15	15	Does corticosteroid therapy ameliorate the acute manifestations of Henoch-Schánlein purpura or mitigate renal injury?
Biocompatible hemodialysis membranes for acute renal failure	Alonso,A.; Lau,J.; Jaber,B.L.	2008	18254074	16	10	Does the use of biocompatible membranes (BCM) confer an advantage in either survival or recovery of renal function over the use of bioincompatible membranes (BICM) in adult patients with acute renal failure (ARF) requiring intermittent hemodialysis?
Ultrasound monitoring to detect access stenosis in hemodialysis patients: a systematic review	Tonelli M; James M; Wiebe N	2008	18371539	11	10	Does vascular access screening for fistulas and grafts improve clinically relevant outcomes in hemodialysis patients?
Meta-analysis: vitamin D compounds in chronic kidney disease	Palmer SC; McGregor DO; Macaskill P	2007	18087055	76	62	Does vitamin D therapy improve biochemical markers of mineral metabolism and cardiovascular and mortality outcomes in chronic kidney disease?
Evidence-based emergency medicine review. Prevention of contrast-induced nephropathy in the emergency department	Sinert R; Doty CI	2007	17512638	7	7	How do different prophylactic therapies compare in how they prevent contrast-induced nephropathy?
Effects of statins in patients with chronic kidney disease: meta-analysis and meta-regression of randomised controlled trials	Strippoli GF, Navaneethan SD, Johnson DW	2008	18299289	50	49	How effective and safe are statins for renal and cardiovascular outcomes in each stage of chronic kidney disease (pre-dialysis, dialysis, and transplantation)?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
A meta-analysis of hemodialysis catheter locking solutions in the prevention of catheter-related infection	Jaffer Y.; Selby, N.; Taal, M.W.; Fluck, R.J.; McIntyre, C.W.	2008	18215701	7	7	How effective are antimicrobial lock solutions (ALSs) at decreasing catheter-related infection (CRI), catheter thrombosis, mortality, and other side-effect rates?
Interventions for minimal change disease in adults with nephrotic syndrome	Palmer SC; Nand K; Strippoli GF	2008	18253993	3	3	What are the benefits and harms of interventions for the nephrotic syndrome in adults caused by minimal change disease?
Meta-analysis: effect of monotherapy and combination therapy with inhibitors of the renin angiotensin system on proteinuria in renal disease	Kunz R; Friedrich C; Wolbers M	2008	17984482	49	47	What are the effects of angiotensin-receptor blockers (ARBs) on urinary protein excretion in patients with nephropathy compared with placebo and other antihypertensive agents and their combinations?
Outcomes in patients with chronic kidney disease referred late to nephrologists: a meta-analysis	Chan MR; Dall AT; Fletcher KE	2007	18060927	22	21	What differences exist in mortality and the duration of hospitalization in patients with chronic kidney disease who are referred early versus late to nephrologists?
Renal replacement therapy in patients with acute renal failure: a systematic review	Pannu N; Klarenbach S; Wiebe N	2008	18285591	30	30	What evidence is available to guide the provision of dialysis to patients with acute renal failure (ARF)?
Warfarin anticoagulation in hemodialysis patients: a systematic review of bleeding rates	Elliott MJ; Zimmerman D; Holden RM	2007	17720522	5	4	What is known about the rates of bleeding episodes per patient-year in HD patients treated with warfarin compared with no warfarin or subcutaneous heparin?
Change in proteinuria after adding aldosterone blockers to ACE inhibitors or angiotensin receptor blockers in CKD: a systematic review	Bomback AS; Kshirsagar AV; Amamoo MA	2008	18215698	15	12	What is the effect of adding mineralocorticoid receptor blockers (MRBs) to angiotensin-converting enzyme (ACE)-inhibitor and/or angiotensin receptor blocker (ARB) therapy on proteinuria, glomerular filtration rate (GFR), blood pressure, and risk of hyperkalemia?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Interferon treatment in hemodialysis patients with chronic hepatitis C virus infection: a systematic review of the literature and meta-analysis of treatment efficacy and harms	Gordon CE; Uhlig K; Lau J	2008	18215704	25	24	What are the effects and harms of interferon (IFN) and pegylated IFN (PEG-IFN) treatment of hemodialysis patients with chronic HCV infection?
The efficacy of loop diuretics in acute renal failure: assessment using Bayesian evidence synthesis techniques	Sampath S; Moran JL; Graham PL	2007	18084840	13	12	What is the efficacy of loop diuretics in acute renal failure?
Mycophenolate mofetil for induction therapy of lupus nephritis: a systematic review and meta-analysis	Walsh M; James M; Jayne D	2007	17702723	6	4	What is the risk for failure to induce remission of lupus nephritis in patients who were treated with mycophenolate mofetil compared with cyclophosphamide?
Mycophenolate mofetil decreases acute rejection and may improve graft survival in renal transplant recipients when compared with azathioprine: a systematic review	Knight, S. R.; Russell, N. K.; Barcena, L.; Morris, P. J. ;;	2009	19300178	27	23	Does mycophenolate mofetil (MMF) improve outcomes compared with azathioprine (AZA) in renal transplant recipients in incidence of acute rejection, patient and graft survival, and toxicity?
Calcineurin inhibitor sparing with mycophenolate in kidney transplantation: a systematic review and meta-analysis	Moore, J.; Middleton, L.; Cockwell, P.; et al. ;;	2009	19307799	19	19	How does calcineurin inhibitor (CNI) sparing with mycophenolate mofetil (MMF) as sole adjunctive immunosuppression effect transplant outcome?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Antiproteinuric response to dual blockade of the renin-angiotensin system in primary glomerulonephritis: meta-analysis and metaregression	Catapano, F.; Chiodini, P.; De Nicola, L.; et al. ;;	2008	18468748	13	13	How does the antiproteinuric efficacy and safety of combination therapy compare to monotherapy with either an angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blockers (ARBs) in patients with primary glomerulonephritis (GN)?
Nephrotoxicity of iso-osmolar iodixanol compared with nonionic low-osmolar contrast media: meta-analysis of randomized controlled trials	Heinrich, M. C.; Haberle, L.; Muller, V.; Bautz, W.; Uder, M. ;	2009	19092091	27	21	How does the nephrotoxicity of iodixanol compare with that of nonionic low-osmolar contrast media (LOCM)?
Educational interventions in kidney disease care: a systematic review of randomized trials	Mason, J.; Khunti, K.; Stone, M.; Farooqi, A.; Carr, S. ;	2008	18440681	21	21	How effective are educational interventions in people with kidney disease?
Management of chronic allograft nephropathy: a systematic review	Birnbaum, L. M.; Lipman, M.; Paraskevas, S.; et al. ;;	2009	19339427	12	12	How effective are various immunosuppressive management strategies of chronic allograft nephropathy (CAN) and of chronic allograft dysfunction (CAD)?
Sodium bicarbonate-based hydration prevents contrast-induced nephropathy: a meta-analysis	Meier, P.; Ko, D. T.; Tamura, A.; Tamhane, U.; Gurm, H. S. ;	2009	19439062	17	8	How effective is normal saline versus sodium bicarbonate for prevention of contrast-induced nephropathy?
Steroid avoidance or withdrawal for kidney transplant recipients	Pascual, J.; Zamora, J.; Galeano, C.; Royuela, A.; Quereda, C. ;	2009	19160257	29	24	How safe and effective is steroid withdrawal or avoidance in patients receiving a kidney transplant?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Aldosterone antagonists for preventing the progression of chronic kidney disease: a systematic review and meta-analysis	Navaneethan, S. D.; Nigwekar, S. U.; Sehgal, A. R.; Strippoli, G. F. ;;	2009	19261819	11	10	What are the benefits and harms of adding selective and nonselective aldosterone antagonists (AA) in chronic kidney disease (CKD) patients already on renin-angiotensin system blockers (RAS)?
Interventions for renal vasculitis in adults	Walters, G.; Willis, N.S.; Craig, J.C.	2009	18646089	18	13	What are the benefits and harms of different interventions for the treatment of renal vasculitis in adults?
HMG CoA reductase inhibitors (statins) for kidney transplant recipients	Navaneethan, S. D.; Perkovic, V.; Johnson, D. W.; Nigwekar, S. U.; Craig, J. C.; Strippoli, G. F.	2009	19370615	62	28	What are the benefits and harms of statin therapy on mortality and renal outcomes in kidney transplant recipients?
HMG CoA reductase inhibitors (statins) for people with chronic kidney disease not requiring dialysis	Navaneethan, S. D.; Pansini, F.; Perkovic, V.; et al. ;;	2009	19370693	52	42	What are the benefits and harms of statins in chronic kidney disease (CKD) patients not receiving renal replacement therapy?
HMG CoA reductase inhibitors (statins) for dialysis patients	Navaneethan, S. D.; Nigwekar, S. U.; Perkovic, V.; Johnson, D. W.; Craig, J. C.; Strippoli, G. F.	2009	19370598	32	18	What are the benefits and harms of statins in dialysis patients?
Atrial natriuretic peptide for management of acute kidney injury: a systematic review and meta-analysis	Nigwekar, S. U.; Navaneethan, S. D.; Parikh, C. R.; Hix, J. K. ;;	2009	19073785	10	10	What are the benefits of atrial natriuretic peptide (ANP) in the prevention and treatment of acute kidney injury (AKI)?
Sodium bicarbonate therapy for prevention of contrast-induced nephropathy: a systematic review and meta-analysis	Navaneethan, S. D.; Singh, S.; Appasamy, S.; Wing, R. E.; Sehgal, A. R. ;	2009	19027212	12	6	What are the benefits of hydration with sodium bicarbonate compared with normal saline in the prevention of contrast-induced nephropathy?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
Medical adjuvant treatment to increase patency of arteriovenous fistulae and grafts	Osborn, G.; Escofet, X.; Da Silva, A. ;;;	2008	18843633	10	9	What are the effects of adjuvant drug treatment on the patency of fistulae and grafts in patients with end-stage renal disease (ESRD) who are undergoing haemodialysis?
Immunosuppressive treatment for focal segmental glomerulosclerosis in adults	Braun, N.; Schmutzler, F.; Lange, C.; et al. ;;	2008	18646090	4	3	What are the effects of different immunomodulatory and immunosuppressive regimens in adults with focal and segmental glomerulosclerosis (FSGS)?
Meta-analysis of N-acetylcysteine to prevent acute renal failure after major surgery	Ho, K. M.; Morgan, D. J.	2009	18649982	10	10	What are the effects of N-Acetylcysteine (NAC) on mortality, acute renal failure requiring dialysis, allogeneic blood transfusion, surgical reexploration for bleeding, and length of intensive care unit (ICU) stay?
Does perioperative hemodynamic optimization protect renal function in surgical patients? A meta-analytic study	Brienza, N.; Giglio, M. T.; Marucci, M.; Fiore, T.	2009	19384211	20	20	What are the effects of perioperative hemodynamic optimization on postoperative acute renal dysfunction?
Intravenous versus oral iron supplementation for the treatment of anemia in CKD: systematic review and meta-analysis	Rozen-Zvi, B.; Gafter-Gvili, A.; Paul, M.; Leibovici, L.; Shpilberg, O.; Gafter, U.	2008	18845368	13	11	What is the best method of iron administration for the treatment of patients with anemia of chronic kidney disease (CKD)?
Low-protein diet for diabetic nephropathy: a meta-analysis of randomized controlled trials	Pan, Y.; Guo, L. L.; Jin, H. M.	2008	18779281	8	8	What is the effect of a low-protein diet (LPD) on renal function in patients with type 1 or 2 diabetic renal diseases?
Timing of renal replacement therapy initiation in acute renal failure: a meta-analysis	Seabra, V. F.; Balk, E. M.; Liangos, O.; Sosa, M. A.; Cendoroglo, M.; Jaber, B. L.	2008	18562058	20	17	What is the effect of early renal replacement therapy (RRT) on mortality in patients with acute renal failure (ARF)?

Systematic review title	Authors	Year	PMID	Total no. of included studies	No. included studies in PubMed	Clinical question
N-acetylcysteine to reduce renal failure after cardiac surgery: a systematic review and meta-analysis	Naughton, F.; Wijeysundera, D.; Karkouti, K.; Tait, G.; Beattie, S.	2008	19050086	7	7	What is the effect of N-acetylcysteine (NAC) on acute renal failure and important clinical outcomes after cardiac surgery?
Effect of lowering blood pressure on cardiovascular events and mortality in patients on dialysis: a systematic review and meta-analysis of randomised controlled trials	Heerspink, H. J.; Ninomiya, T.; Zoungas, S.; et al. ;;	2009	19249092	7	6	What is the effect of treatments that reduce blood pressure in patients receiving maintenance dialysis?
Effects of L-carnitine on dialysis-related hypotension and muscle cramps: a meta-analysis	Lynch, K. E.; Feldman, H. I.; Berlin, J. A.; Flory, J.; Rowan, C. G.; Brunelli, S. M.	2008	18706751	7	6	What is the efficacy of L-carnitine supplementation for treatment of patients with intradialytic hypotension and cramping?

Appendix 4: Survey of nephrologists

Survey: Exploring how Nephrologists Search for Medical Information

London Health Sciences Centre, Room ELL-101, London, Ontario, N5A 4G5

LETTER OF INFORMATION AND CONSENT

Investigators
Ms. Salimah Shariff, PhD Student, Department of Epidemiology & Biostatistics, University of Western Ontario
Dr. Amit Garg, Nephrologist, London Health Sciences Centre, Victoria Campus Director, London Kidney Clinical Research Unit, 519-685-8502.

Participation
We invite you to participate in this research survey; participation is voluntary and takes 5-10 minutes to complete. To be eligible you must be a licensed, practicing Nephrologist in Canada.

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time. You indicate your consent to participate by completing and returning the attached questionnaire.

Purpose of the survey
The research survey explores how nephrologists search for medical information. The survey consists of a variety of questions regarding the information sources and procedures nephrologists use to search for medical information.

The purpose of the research is to try and develop better ways to help nephrologists find relevant medical information.

Possible risks and benefits to you for participating in the survey
Possible risks: There are no risks to you for participating in this survey.

Possible benefits: If we develop better ways to find medical information in nephrology this could be a benefit to all physicians.

Confidentiality
Your information will be kept confidential and reported anonymously as grouped data. All electronic data will be stored in a password protected, secure, database and print material will be kept under lock at the London Health Sciences Centre. Only the research team will have access to any collected data. All data will be destroyed and permanently deleted after 25 years. When the results of the study are published, your name will not be used. Representatives of the University of Western Ontario Health Sciences Research Ethics Board may require access to study-related records for the purpose of monitoring the research.

Contact persons
If you have any questions about the content of this study, please contact the study coordinator, <*insert name*>, at <*insert email*> or call at: <*insert phone number*>. If you have any questions about the conduct of this study or your rights as a research participant you may contact the Director of the Office of Research Ethics at 519-661-3030 or at ethics@uwo.ca.

Funding Sources
The study has been funded by the Canadian Institutes of Health Research.

Consent
I have read the Letter of Information and agree to participate. All questions have been answered to my satisfaction. Completion of this survey indicates your consent to participate.

Thank you for participating in our survey!

Survey: Exploring how Nephrologists Search for Medical Information

1. **Are you a practicing nephrologist in Canada?**

   ☐ Yes ☐ No

2. **In the past year, have you used the following online sources to find information to guide the treatment of a patient?**

Cochrane Collaboration	☐ Yes	☐ No
Elsevier's Scirus	☐ Yes	☐ No
EMBASE	☐ Yes	☐ No
Google	☐ Yes	☐ No
Google Scholar	☐ Yes	☐ No
PubMed (MEDLINE)	☐ Yes	☐ No
Ovid (MEDLINE)	☐ Yes	☐ No
PubMed using *Clinical Queries* Feature	☐ Yes	☐ No
UpToDate	☐ Yes	☐ No
Yahoo!	☐ Yes	☐ No

3. **Have you previously received training in literature searching?**
   Examples of training include Searching Skills Workshops, Library Training Sessions, PubMed Tutorials

   ☐ Yes ☐ No

4. **On average, how many times per month do you search for information to guide the treatment of your patients?**
   Searching for information includes reading textbooks, searching the internet, using online bibliographic databases like PubMed, or using software on your palm pilot / blackberry.

   I search for information to guide the treatment of my patients an average of _____ times per month.
   If you never search for information, indicate 0.

5. **On average, how many times per month do you search a bibliographic database for medical literature?**
   Examples of online bibliographic databases are PubMed (MEDLINE), EMBASE, or Google Scholar

I search a bibliographic database an average of _____ times per month.
If you never search bibliographic databases for medical literature, indicate 0.

6. When you search a bibliographic database, do you scan results from the top of the page and make your way down?

☐ Yes ☐ No

7. When you search a bibliographic database, how many results do you generally scan per search? *(e.g. 10 results)*
(For your reference: by default, PubMed displays 20 results per page and Google Scholar displays 10 results per page)

_____ Number of results

8. For the following statements, please indicate your level of agreement:

a. I am comfortable searching for general information on the Internet.
*(e.g. using **Yahoo! or Google** to search for things to do in New York City)*

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
☐	☐	☐	☐	☐

b. I am comfortable searching for medical information on the Internet.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
☐	☐	☐	☐	☐

c. I am comfortable searching for medical literature in PubMed.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
☐	☐	☐	☐	☐

9. **When searching online for an answer to guide the treatment of a patient, which of the following information sources would you go to first?** (please select only one response)

- ☐ Cochrane Collaboration
- ☐ Elsevier's Scirus
- ☐ EMBASE
- ☐ Google
- ☐ Google Scholar
- ☐ PubMed (MEDLINE)
- ☐ Ovid (MEDLINE)
- ☐ PubMed using *Clinical Queries* Feature
- ☐ UpToDate
- ☐ Yahoo!
- ☐ Other (Please specify:____________________________)

10. **Consider the following scenario:**

Suppose you performed a search in PubMed and 10 of the first 20 results were relevant to your search. You try the same search using a new online bibliographic database. Of the first 20 results, how many would have to be relevant to compel you to use the new database instead of PubMed next time?

Using PubMed, 10 out of 20 results were relevant.
I would use the new database next time instead of PubMed if _____ out of 20 results were relevant.

This next question (question 11) is the most important question of the survey.

11. Suppose you wanted to search for an answer to the following clinical question:

<insert clinical question>

Please enter the terms or phrase you would type into a search box of an **online bibliographic search database** to obtain an answer to the above question. *(e.g. using PubMed, Google Scholar, EMBASE)*

Enter search phrase in the following box:

If you have any additional information related to your search strategy, please enter it here:

12. When searching an online bibliographic database, do you use any of the following search options?

Boolean searching (using AND, OR & NOT to connect search terms)	☐ Yes	☐ No
Limits (e.g. limiting the scope of your search by language, publication type, date, author, age of participants, type of article)	☐ Yes	☐ No
Controlled vocabularies (e.g. searching using Medical Subject Headings MeSH terms in PubMed/Medline)	☐ Yes	☐ No
Truncation or wildcards (e.g. * or $)	☐ Yes	☐ No

13. Do you practice nephrology at a centre with a nephrology fellowship training program?

☐ Yes ☐ No

14. How many years have you practiced nephrology since completing your residency training? *(e.g.. 3 years)*

Number of years: _____

15. What is your gender?

☐ Male
☐ Female

16. What is your age?

_____ years old.

Thank you for completing the survey.

We appreciate you taking the time to participate.

Appendix 5: Research ethics approval

Office of Research Ethics

The University of Western Ontario
Room 4180 Support Services Building, London, ON, Canada N6A 5C1
Telephone: (519) 661-3036 Fax: (519) 850-2466 Email: ethics@uwo.ca
Website: www.uwo.ca/research/ethics

Western

Use of Human Subjects - Ethics Approval Notice

Principal Investigator: Dr. A.X. Garg
Review Number: 13668E **Revision Number:** 2
Review Date: June 19, 2009 **Review Level:** Expedited
Protocol Title: Survey of how nephrologists search for medication information
Department and Institution: Nephrology, London Health Sciences Centre
Sponsor: CIHR-CANADIAN INSTITUTE OF HEALTH RESEARCH
Ethics Approval Date: July 08, 2009 **Expiry Date:** December 31, 2010
Documents Reviewed and Approved: Revised study end date, administrative changes, study instruments and study methods. First Email to Nephrologists, Second Email to Nephrologists, Letter of Information and Consent.
Documents Received for Information:

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices: Consolidated Guidelines; and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced revision(s) or amendment(s) on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the UWO Updated Approval Request Form.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the HSREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of monitor, telephone number). Expedited review of minor change(s) in ongoing studies will be considered. Subjects must receive a copy of the signed information/consent documentation.

Investigators must promptly also report to the HSREB:

a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to this office for approval.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

Chair of HSREB: Dr. Joseph Gilbert

Ethics Officer to Contact for Further Information			
☐ Janice Sutherland (jsutherl@uwo.ca)	☐ Elizabeth Wambolt (ewambolt@uwo.ca)	☑ Grace Kelly (grace.kelly@uwo.ca)	☐ Denise Grafton (dgrafton@uwo.ca)

This is an official document. Please retain the original in your files.

cc: ORE File
LHRI

UWO HSREB Ethics Approval - Revision
V.2008-07-01 (rptApprovalNoticeHSREB_REV) 13668E Page 1 of 1

Appendix 6: Sample-size calculation for Objective 1 - Determinants of search success

Sample size estimates for each study hypothesis were calculated using the SAS power procedure 'twosamplemeans'. The formula for calculation is outlined below. For all calculations, power was specified as 80% and the significance level was specified as 0.05. Ratios of unexposed to exposed for each predictor is summarized in Table 1 were determined from the 20 responses received from the pilot phase of Objective 2 (see Appendix 15).

$$n_1 = \frac{(Z_{\alpha/2} + Z_{\beta})^2 \sigma^2 (r+1)}{(\Delta)^2 r}$$

where:

n_1 is the number of 'exposed' cases studied.

Δ is the minimum difference in means, between exposure groups, that one wishes to detect.

r is the ratio of the number of unexposed cases to the number of exposed cases.

σ is the standard deviation in the population for a continuously distributed (outcome) variable.

Table 1: Definitions of exposures and ratios

Variable	**Definition / Groups**	**Estimated Ratio Referent (unexposed): Comparator (exposed)**
Use of multifaceted search queries	Use of no features (referent) Use of at least 1 feature	1:1
Number of search concepts	Less than 4 terms (referent) 4 or more terms	2:1
Frequency of searching	Less than once a week (referent) Once a week or more	1:2
Previous training in literature searching	no (referent) yes	1:1

Primary Outcome: Sensitivity

Sample size estimates for a range of ratios for the proportion of unexposed to exposed responses are summarized in Table 2 (SAS output on the next page). The **standard deviation** for sensitivity was estimated at 0.19 using the pilot data collected for Objective 2. The calculations incorporate a minimum detectable difference of 15% in sensitivity.

Table 2: Sample size estimates for different ratios of unexposed to exposed responses[£]

Unexposed : Exposed **r** (% Exposed)	Sample Size of Exposed (n_1)	Total Sample Size n_1*(r + 1)
1:1 (50 %)	27	54
1:2 (66.7%)	40	60
2:1 (33.3%)	20	60

[£]Where $\sigma = 0.19$ and $\Delta = 15\%$ difference in Sensitivity;

Secondary Outcome: Precision

Given a sample size of 60 and using the standard deviation of 0.05 for the outcome measure precision, this study will provide 80% power to detect a minimum difference of 4% in precision, should this difference in truth exist.

SAS Output

Outcome: Sensitivity, Ratio 1:1

```
The POWER Procedure
Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution                Normal
Method                       Exact
Mean Difference               0.15
Standard Deviation            0.19
Group 1 Weight                   1
Group 2 Weight                   1
Nominal Power                  0.8
Number of Sides                  2
Null Difference                  0
Alpha                         0.05

Computed N Total

Actual        N
 Power    Total

 0.812       54
```

Outcome: Sensitivity, Ratio 1:2 [or 2:1]

```
The POWER Procedure
Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution                Normal
Method                       Exact
Mean Difference               0.15
Standard Deviation            0.19
Group 1 Weight                   1
Group 2 Weight                   2
Nominal Power                  0.8
Number of Sides                  2
Null Difference                  0
Alpha                         0.05

Computed N Total

Actual        N
 Power    Total

 0.809       60
```

Outcome: Precision, Fixed Sample Size: 60

The POWER Procedure
Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution	Normal
Method	Exact
Standard Deviation	0.05
Group 1 Weight	1
Group 2 Weight	2
Total Sample Size	60
Power	0.8
Number of Sides	2
Null Difference	0
Alpha	0.05

Computed Mean Diff

Mean Diff
0.039

Appendix 7: MODEL #1; Predictor: Search Query Characteristics; Outcome: Sensitivity; Question 1

A7.1 Multivariable linear regression

<u>The assumptions of linear regression include:</u>

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

<u>Examination of outcome (sensitivity)</u>

The examination of the outcome revealed that sensitivity is positively skewed (Figure 1). A solution to remedying a positive skew is to take the log or the square root of the outcome measure. As the outcome consisted of a significant number of zeros, a log could not be taken. Instead, I transformed the data with a square root. Results from the transformed outcome are presented in Figure 2. The transformation reduced the skewness and thus I continued the analysis with this transformed outcome.

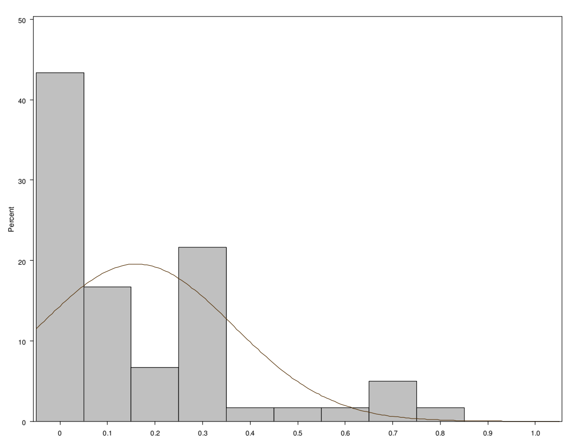


Figure 1a: Histogram of sensitivity

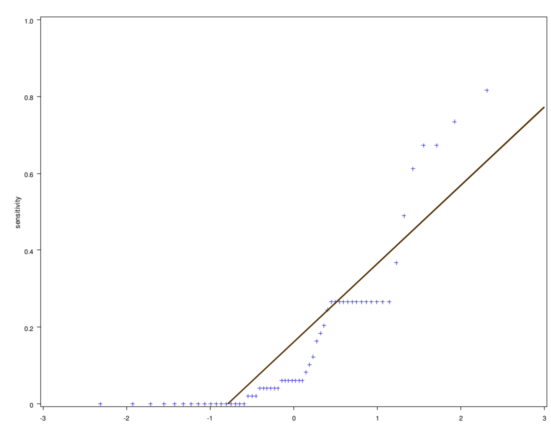


Figure 1b: Normal probability plot of sensitivity

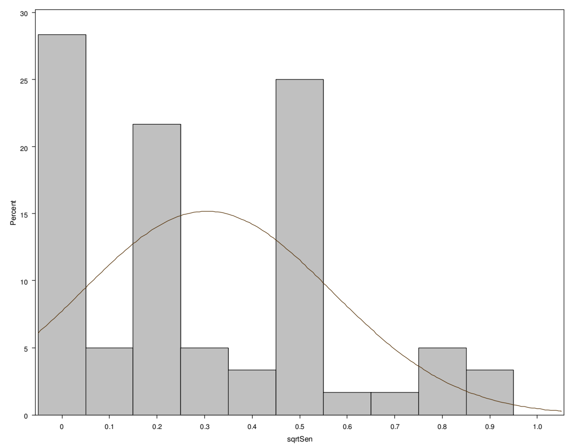


Figure 2a: Histogram of square root of sensitivity

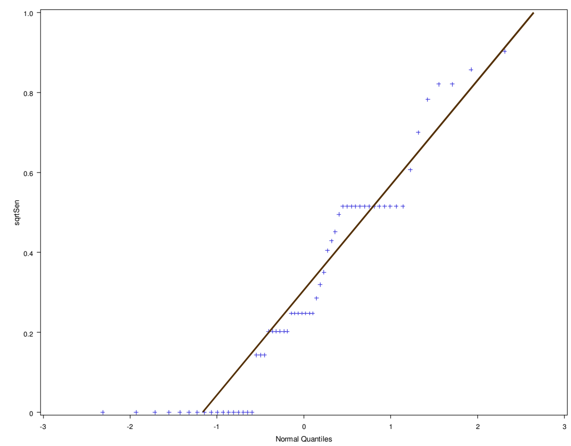


Figure 2b: Normal probability plot of square root of sensitivity

Univariate analysis

An examination of the descriptive statistics revealed that almost all search queries specified the patient and intervention terms, but no control term and thus these variables could not be considered for the regression models to follow. In addition, the use of quotations in the query was also not further considered as only one search used quotes. The relationship between the primary predictors (number of concept terms and use of multifaceted search features) and square root of sensitivity are presented in Figure 3. There appears to be a linear relationship between the predictors and the outcome, which is required for a linear regression (assumption #1).

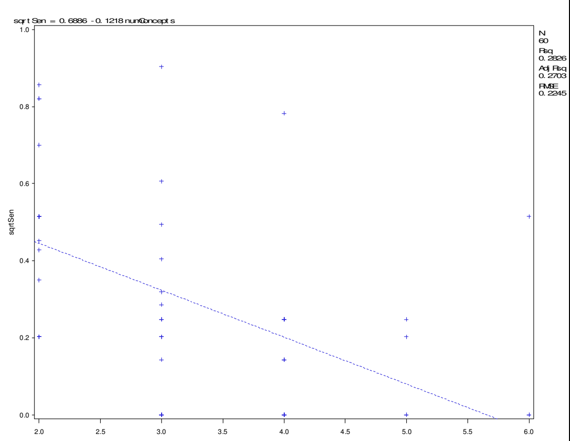


Figure 3a: Relationship between number of concept terms and the square root of sensitivity

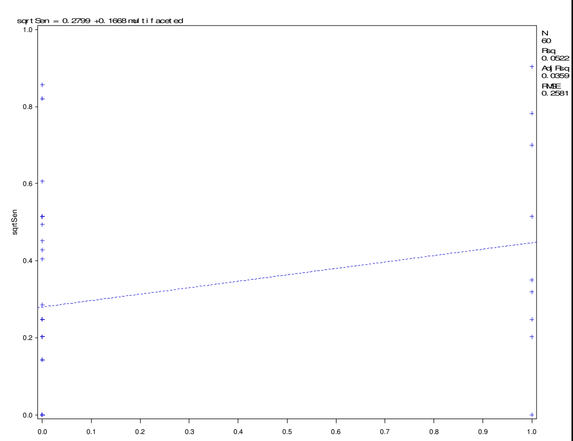


Figure 3b: Relationship between the use of multifaceted search features and the square root of sensitivity

Build base model

The base model consists of including the two primary predictors without any other covariates/confounders. The measure of number of concepts was modeled as a quantitative covariate. An alternative to this method would be to use dummy variables. However, as there are 6 categories for the number of concepts, this would require using 5 degrees of freedom and with only 60 observations the model specification should not exceed 6 degrees of freedom as this makes it more difficult to assure the normality of the residuals. Multifaceted searching was included into the model as a binary measure. Details of the analysis are presented in Table 1. The adjusted R-squared for this model was 47%, suggesting that 47% of the variance in the outcome is explained by the inclusion of the variables.

Table 1: Results of the analyses of the association the primary predictors and square root of sensitivity

Variable	**Estimate for change in the square root of sensitivity (95% Confidence intervals)**	**p-value**
Number of concepts	-0.13 (-0.18 to -0.08)	<0.001
Use of multifaceted search features (referent group: No)	0.27 (0.12 to 0.42)	<0.001

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by a minimum of 10%. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 2. As both potential confounders changed the estimate of the regression coefficients by more than 10%, they were both included into the fitted model.

Table 2: Percent change in the estimates with the addition of potential confounders

Confounding variable	Percent change in square root of sensitivity estimate for number of concepts	Percent change in square root of sensitivity estimate for use of multifaceted search features
Outcome term used in search (referent group: No)	48%	40%
Acronym term used in search (referent group: No)	11%	18%

Fitted model

Results of the full model are presented in Table 3. The adjusted R-squared for this model increased to 64%, suggesting that 64% of the variance in the outcome is explained by the inclusion of the variables. This provided evidence supporting a good model specification.

Table 3: Regression estimates of final model

Variable	Estimate for change in the square root of sensitivity (95% Confidence intervals)	p-value
Number of concept terms	-0.05 (-0.09 to 0.00)	0.041
Use of multifaceted search features (referent group: No)	0.14 (-0.01 to 0.26)	0.029
Outcome term used in search (referent group: No)	-0.25 (-0.36 to -0.15)	<0.001
Acronym term used in search (referent group: No)	-0.19 (-0.29 to -0.10)	<0.001

Diagnostics

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The residuals appear to deviate slightly from equal variance and more than 5% (8%) of residuals lie outside the 95% confidence interval. However, the normal probability plot appears quite fitted, except for one potential outlier which was examined.

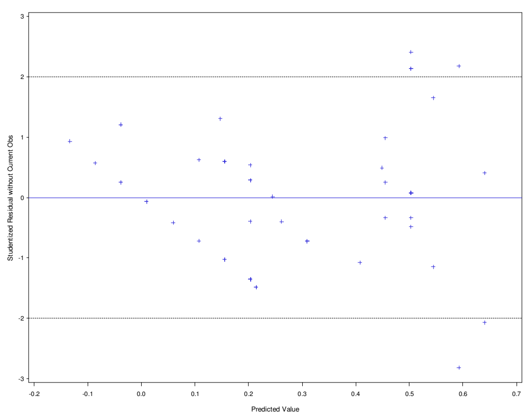


Figure 4a: Scatter plot of residuals

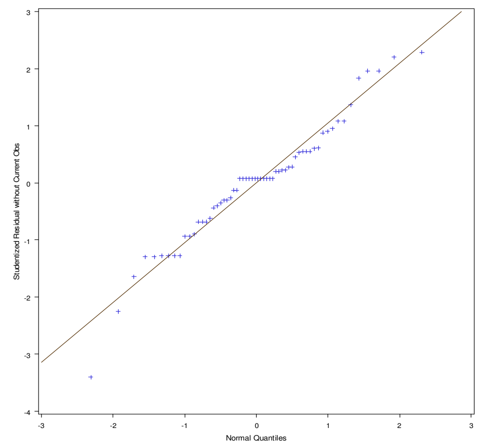


Figure 4b: Normal probability plot of residuals

A7.2 Poisson and negative binomial regression

Since there was some concern about the residuals in the linear regression model not exhibiting equal variance and because other models in this thesis used Poisson or negative binomial regression modeling, I repeated the model building for this outcome (see Appendix 8 for explanation of Poisson and negative binomial modeling). If the Poisson or negative binomial proved to also be an accurate modeling type, this would make comparisons across the models simpler.

<u>The assumptions of Poisson regression include:</u>

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

<u>Examination of outcome (sensitivity)</u>

To test the first assumption, I plotted the log of sensitivity against the primary predictors (number of concept terms and multifaceted searching). Since some searches did not find any relevant articles, the outcome included several zero values; instead I added a slight offset of 1 in order to be able to calculate the log. This is presented in Figure 5. The relationships appear linear as is required (assumption #1).

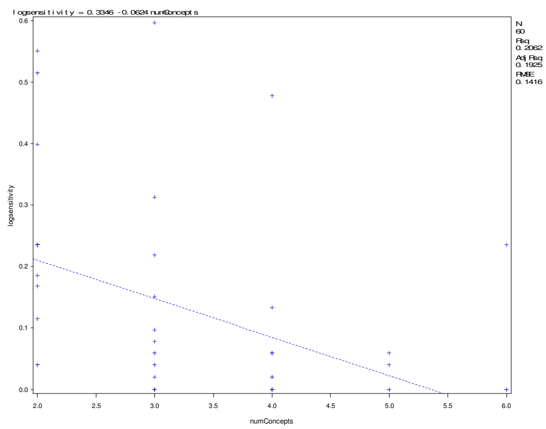


Figure 5a: Relationship between number of concept terms and the log of sensitivity

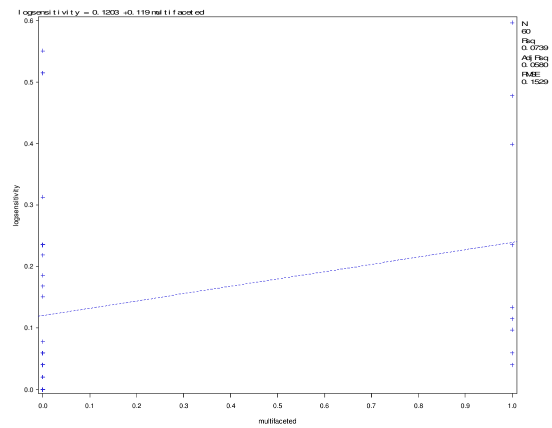


Figure 5b: Relationship between the use of multifaceted search features and the log of sensitivity

Build base model

The base model consists of including the two primary predictors without any other covariates/confounders. The Deviance for the Poisson model l is 3.8 suggesting that the Poisson regression is not appropriate. Instead, I applied the negative binomial regression and compared the Log Likelihood ratio which revealed a p-value of <0.001 suggesting that the negative binomial regression is a better fit. Estimates of the regression coefficients are presented in Table 4.

Table 4: Results of the analyses of the association between the primary predictors and sensitivity

Variable	**Estimate of the rate ratio (95% Confidence Interval)**	**p-value**
Number of concept terms	0.41 (0.31 to 0.56)	<0.001
Use of multifaceted search features (referent group: No)	7.73 (2.89 to 20.71)	<0.001

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by a minimum of 10%. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 5. As the inclusion of both potential confounders changed the estimates of the rate ratio by more than 10%, both were added to the fitted model.

Table 5: Percent change in the estimates with the addition of potential confounders

Confounding variable	**Percent change in rate ratio estimate for number of concepts**	**Percent change in rate ratio estimate for use of multifaceted search features**
Outcome term used in search (referent group: No)	61%	63%
Acronym term used in search (referent group: No)	12%	31%

Fitted model

Results of the full model are presented in Table 6.

Table 6: Regression estimates of final model

Variable	**Estimate for rate ratio (95% Confidence intervals)**	**p-value**
Number of concept terms	0.69 (0.53 to 0.89)	0.005
Use of multifaceted search features (referent group: No)	2.64 (1.39 to 5.00)	0.003
Outcome term used in search (referent group: No)	0.21 (0.12 to 0.39)	<0.001
Acronym term used in search (referent group: No)	0.19 (0.09 to 0.36)	<0.001

Diagnostics

The assessment of residuals is presented in Figure 6. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model appears to fit well as only three residual values are outside the 95% confidence interval (as would be expected) and the normal probability plot appears to be well fitted.

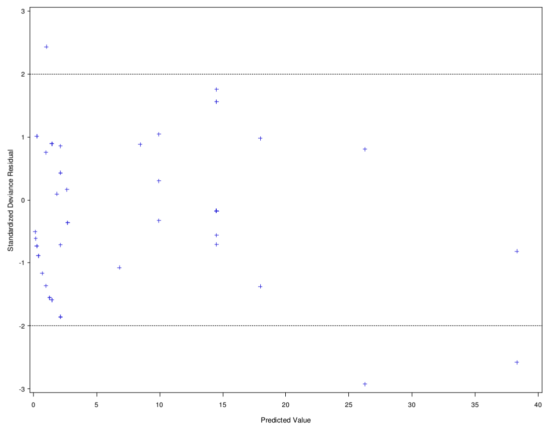


Figure 6a: Scatter plot of residuals

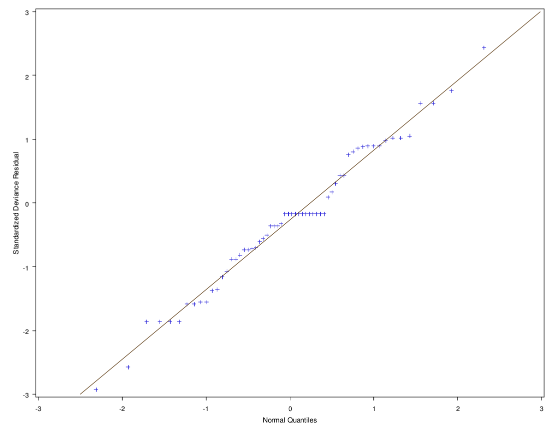


Figure 6b: Normal probability plot of residuals

A7.3 Comparing results from linear regression model and negative binomial model

An analysis of the residual suggests that the negative binomial model fits the data better than the linear regression model. However, the estimates received from the linear regression support the same directions of effect seen in the negative binomial regression, providing evidence of the robustness of the data and its effects.

Appendix 8: MODEL #2; Predictor: Search Query Characteristics; Outcome: Precision; Question 1

A8.1 Multivariable linear regression

Examination of outcome (precision)

An examination of the outcome revealed that in six cases the physician-provided search query did not retrieve any articles, rendering the precision undefined. These six cases accordingly could not be used in the regression. However, I also explored the inclusion of these values (as zero precision) in additional analyses. The histogram of the outcome (Figure 1) revealed that the outcome is positively skewed. A solution to remedying a positive skew is to take the log or the square root of the outcome. As the outcome consists of a number of zeros (no relevant articles were found), a log cannot be taken. Instead, I transformed the data with a square root. The transformed outcome is presented in Figure 2. The transformation reduced the skewness slightly and so I continued the analysis with this transformed outcome.

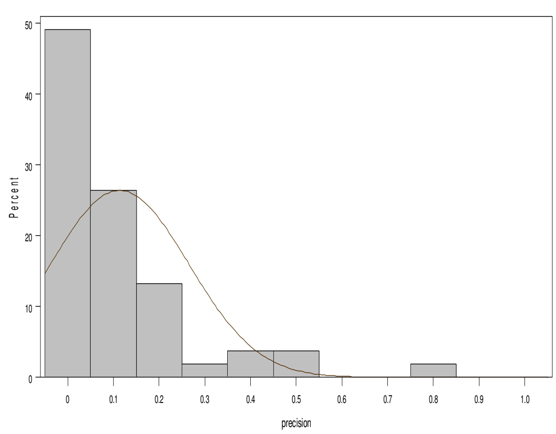


Figure 1a: Histogram of precision

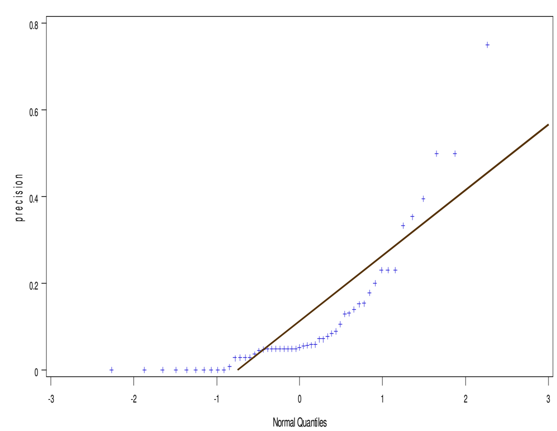


Figure 1b: Normal probability plot of precision

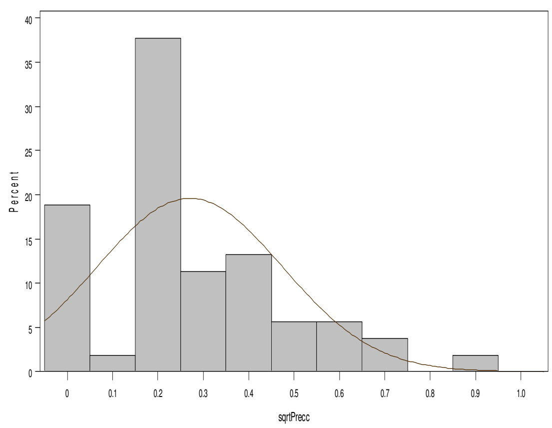


Figure 2a: Histogram of square root of precision

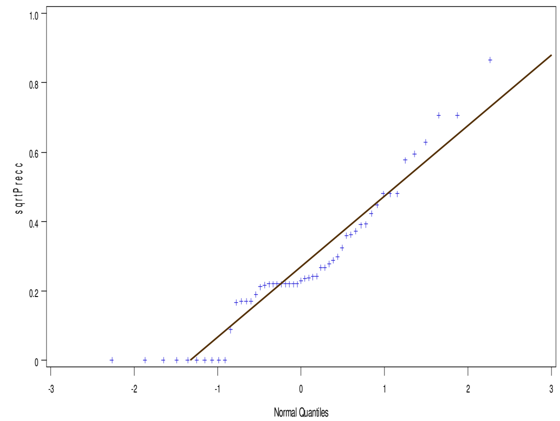


Figure 2b: Normal probability plot of square root of precision

Univariate analysis

An examination of the descriptive statistics revealed that almost all search queries specified the patient and intervention terms, but no control term and thus these variables were not further considered as they would not provide any information to the regression model. In addition, the use of quotation was also not further considered as only one search query included quotes. The relationship between the primary predictors (number of concept terms and use of multifaceted search features) and square root of precision are presented in Figure 3. There appears to be a linear relationship between the predictors and the outcome, which is required for a linear regression (assumption #1).

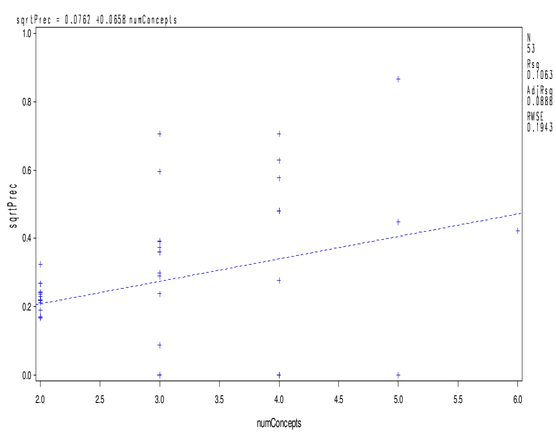


Figure 3a: Relationship between number of concept terms and the square root of precision

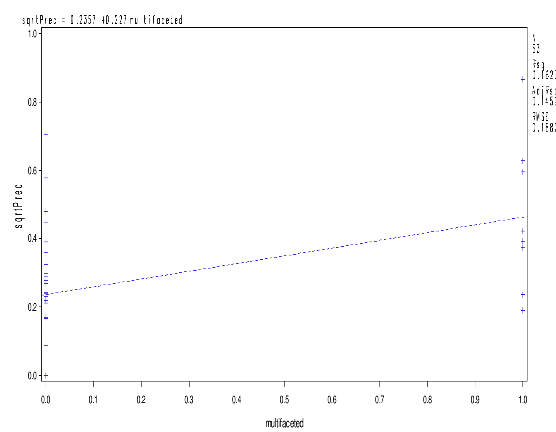


Figure 3b: Relationship between the use of multifaceted search features and the square root of precision

Build base model

The base model consists of including the two primary predictors without any potential confounders. Details of the analysis are presented in Table 1. The adjusted R-squared for this model was 19%, suggesting that 19% of the variance in the outcome is explained by the inclusion of the variables.

Table 1: Results of the analyses of the association between the primary predictors and square root of precision

Variable	Estimate for square root of precision (95% Confidence intervals)	p-value
Number of concept terms	0.05 (-0.003 to 0.10)	0.063
Use of multifaceted search features (referent group: No)	0.19 (0.04 to 0.32)	0.011

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 2. As both potential confounders changed the effect measures for the number of concepts by more than 10%, they were both included into the fitted model.

Table 2: Percent change in the estimates with the addition of potential confounders

Confounding variable	**Percent change in square root of precision estimate for number of concepts**	**Percent change in square root of precision estimate for use of multifaceted search features**
Outcome term used in search (referent group: No)	17%	5%
Acronym term used in search (referent group: No)	17%	6%

<u>Fitted model</u>

Results of the full model are presented in Table 3. The adjusted R-squared for this model was decreased by 3% to 16%, when compared to the base model. This is surprising as it would be expected that the R-squared would increase with the addition of a confounder.

Table 3: Regression estimates of final model

Variable	**Estimate for change in the square root of precision (95% Confidence intervals)**	**p-value**
Number of concept terms	0.06 (-0.01 to 0.12)	0.096
Use of multifaceted search features (referent group: No)	0.18 (0.02 to 0.33)	0.024
Outcome term used in search (referent group: No)	-0.02 (-0.16 to 0.11)	0.711
Acronym term used in search (referent group: No)	0.03 (-0.11 to 0.16)	0.698

<u>Diagnostics</u>

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model specification appears to be appropriate as the residuals appear to exhibit equal variance, albeit with a little clustering. However, 5/53 (9%) points appear outside the 95% confidence lines, and only 2-3 would be expected; three of the points are very close to the boundaries. The normal probability appears to be well fitted with light tails. Due to the concern about equal variance I explored another model type that might be more appropriate for the data.

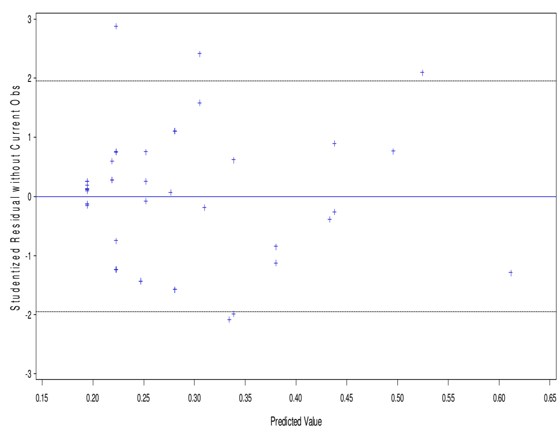


Figure 4a: Scatter plot of residuals

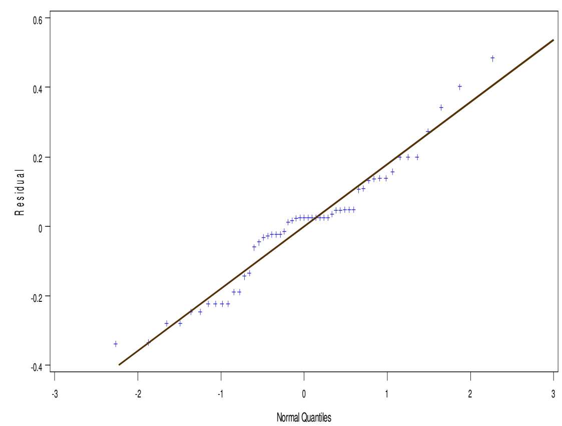


Figure 4b: Normal probability plot of residuals

A8.2 Selecting an appropriate regression model[112-114]

When a linear regression model is not appropriate for the data, several discrete response regression models can be selected instead.

The first is a binary response model (also known as logistic regression). For this model, the outcome variable must be binary. In the case of precision, this would require me to categorize the outcome. Unfortunately, the literature did not elicit any meaningful cut-point by which to categorize the outcome, and thus this model is not appropriate for analyzing precision.

A second potential model is ordinal logistic regression. However, this model also requires categorization of the outcome, and so is not appropriate for analyzing precision.

The third set of models are Poisson or negative binomial regressions. These regressions are used to model the number of occurrences of an event of interest or the rate of occurrence of an event. In the case of rates the numerator is used as the response variable (outcome) and the log of the denominator is included within the model as an offset term. For precision, the numerator represents the number of relevant citations found, while the denominator represents the total number of citations retrieved by the search.

A8.3 Poisson or negative binomial regression

The assumptions of Poisson regression include:

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

Examination of outcome (precision)

To test the first assumption, I plotted the log precision against the primary predictors (number of concept terms and multifaceted searching). Since some searches did not find any relevant articles, the outcome included several zero values; instead I added a slight offset of 1 in order to be able to calculate the log. This is presented in Figure 5. The relationships appeared linear as is required.

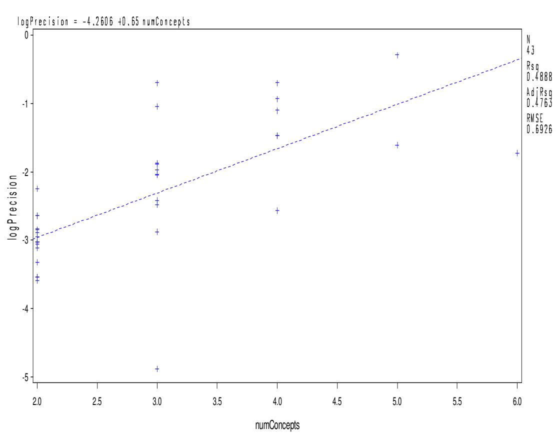


Figure 5a: Plot comparing the log of precision against the number of concepts

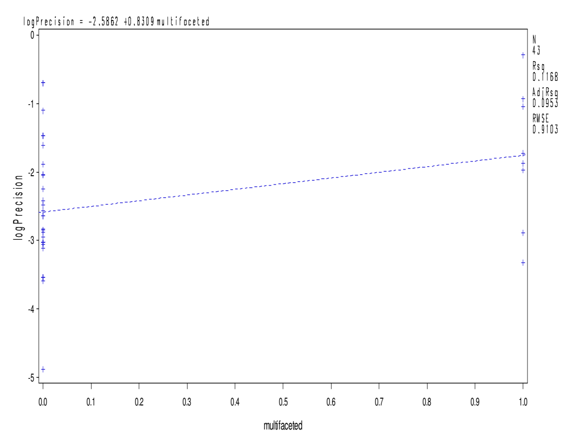


Figure 5b: Plot comparing the log of precision against the use of multifaceted search features

Build base model

To test the second assumption, I ran the base model using Poisson regression and assessed the deviance. The criteria for assessing goodness of fit are presented here:

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	50	166.2982	3.3260
Scaled Deviance	50	166.2982	3.3260
Pearson Chi-Square	50	179.3931	3.5879
Scaled Pearson X2	50	179.3931	3.5879
Log Likelihood		759.9814	
Full Log Likelihood		-164.9284	
AIC (smaller is better)		335.8569	
AICC (smaller is better)		336.3467	
BIC (smaller is better)		341.7678	

Deviance and Pearson Chi-Square divided by the degrees of freedom are used to detect overdispersion or underdispersion. For the Poisson regression the mean and the variance should be equal (equidispersion), which implies that the deviance and the Pearson statistic divided by the degrees of freedom should be approximately one. Values greater than 1 indicate overdispersion, and values smaller than 1 indicate underdispersion. Evidence of underdispersion or overdispersion indicates inadequate fit of the Poisson model. For the current case, there is indication of overdispersion as the Value/DF = 3.3 which is far from one. In the case of overdispersion, running the negative binomial regression instead may be more appropriate, which I chose to do.

The criteria for assessing goodness of fit for the negative binomial regression are presented here:

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	50	51.8960	1.0379
Scaled Deviance	50	51.8960	1.0379
Pearson Chi-Square	50	42.8698	0.8574
Scaled Pearson X2	50	42.8698	0.8574
Log Likelihood		795.1989	
Full Log Likelihood		-129.7110	
AIC (smaller is better)		267.4220	
AICC (smaller is better)		268.2553	
BIC (smaller is better)		275.3032	

In order to determine whether the negative binomial is a better model, a Likelihood Ratio test must be performed. The following are the steps for the test:

1. Record the Log Likeihood (LL) for both the Poisson and negative binomial regressions
2. Compute the likelihood ratio(LR) statistic: -2(LL (Poisson) – LL (negative binomial)).
3. The asymptotic distribution of the LR statistic follows a Chi-squared distribution with 1 degree of freedom. Conduct a Chi-squared test at a significance level of 0.05.

The LR statistic is equal to -2(759.9814-795.1989)=70.4 which is <0.0001 confirming that the negative binomial is a more appropriate model.

Build base model

Regression coefficient estimates of the base model created using negative binomial are presented in Table 4. Regression coefficients equate to the log of the rate ratio for every one unit increase in the predictor variable. Accordingly, Table 4 also presents the exponentiated regression coefficients (the rate ratio).

Table 4: Results of the analyses of the association between the primary predictors and precision

Variable	Estimate log of the rate ratio (95% Confidence Interval)	Estimate rate ratio (95% Confidence Interval)	p-value
Number of concept terms	0.49 (0.26 to 0.71)	1.63 (1.29 to 2.04)	<0.001
Use of multifaceted search features (referent group: No)	0.70 (0.22 to 1.18)	2.01 (1.25 to 3.26)	0.004

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 5. As the addition of each of the potential confounder did not change the estimates of the rate ratio by more than 10%, none of them were included in the final model.

Table 5: Percent change in estimates with the addition of potential confounders

Confounding variable	Percent change in rate ratio of precision for number of concepts	Percent change in rate ratio for use of multifaceted search features
Outcome term used in search (referent group: No)	4%	3%
Acronym term used in search (referent group: No)	0.3%	0.4%

Fitted model

The full model is the same as the base model presented in Table 5.

Diagnostics

The assessment of residuals is presented in Figure 6. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model specification appears to be appropriate as few points lie outside the 95% confidence lines. Three points are outside the limits as would be expected; two may be outliers and were examined. The normal probability appears to be well fitted with light tails.

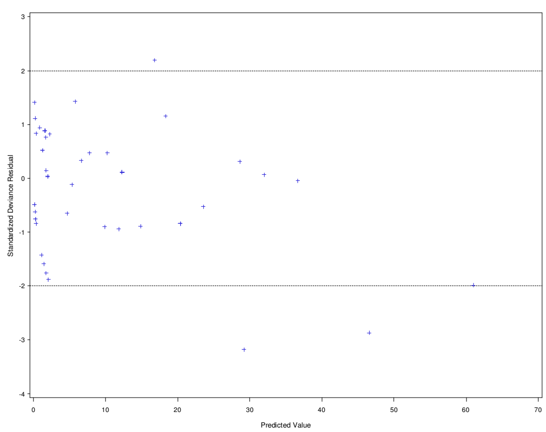


Figure 6a: Scatter plot of residuals

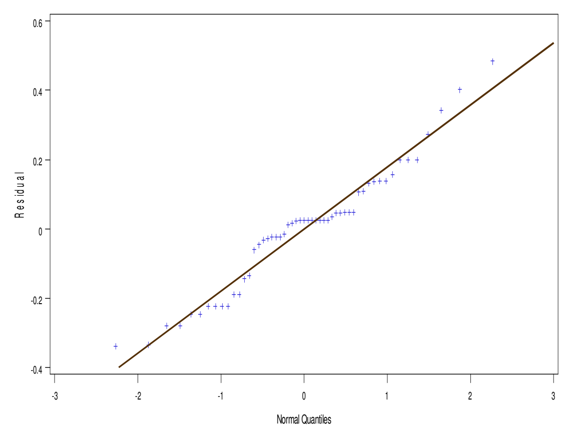


Figure 6b: Normal probability plot of residuals

A8.4 Comparing results from linear regression and negative binomial

An analysis of the residual suggests that the negative binomial model fits the data better than the linear regression model. However, the estimates received from the linear regression support the same directions of effects seen in the negative binomial regression, providing evidence of the robustness of the data and its' effects.

Appendix 9: MODEL #3; Predictor: Nephrologist Characteristics; Outcome: Sensitivity; Question 1

A9.1 Multivariable linear regression

The assumptions of linear regression include:

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

Examination of outcome (sensitivity)

The examination of the outcome revealed that sensitivity is positively skewed (Figure 1). A solution to remedying a positive skew is to take the log or the square root of the outcome measure. As the outcome consists of a significant number of zeros, a log cannot be taken. Instead, I transformed the data with a square root. Results from the transformed outcome are presented in Figure 2. The transformation reduced the skewness and thus I continued the analysis with this transformed outcome.

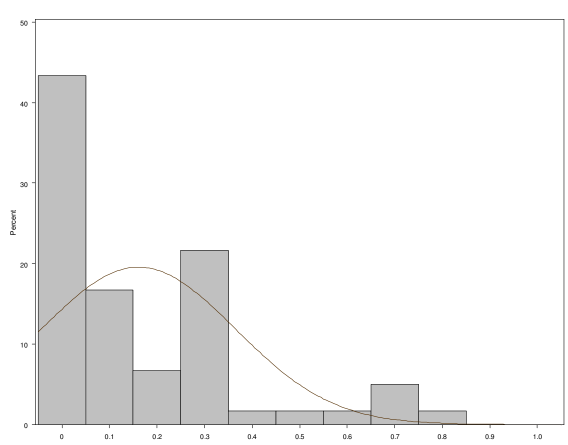


Figure 1a: Histogram of sensitivity

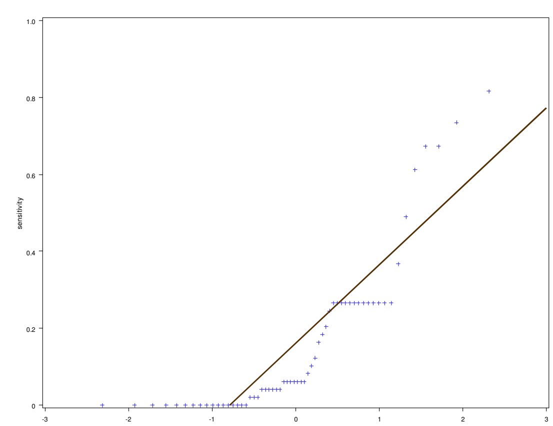


Figure 1b: Normal probability plot of sensitivity

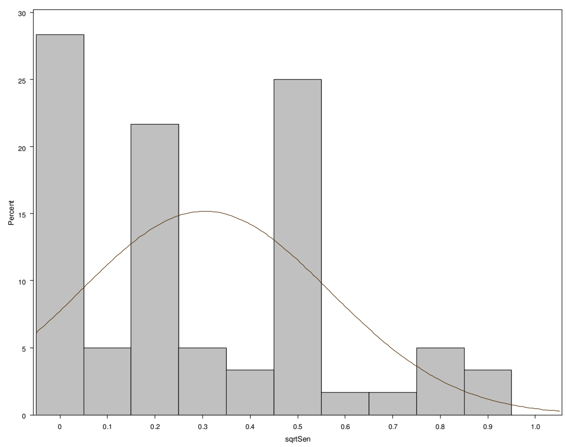


Figure 2a: Histogram of square root of sensitivity

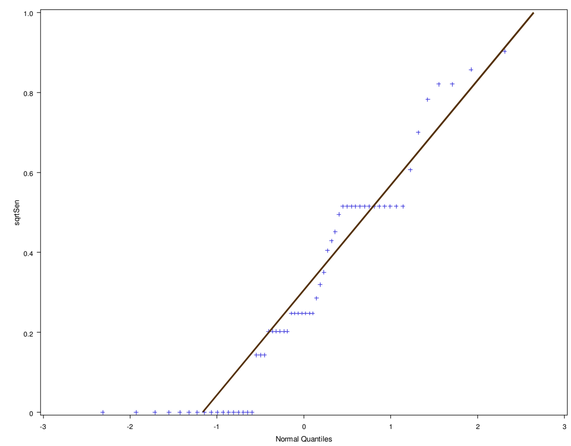


Figure 2b: Normal probability plot of square root of sensitivity

<u>Univariate analysis</u>

An examination of the descriptive statistics revealed that almost all search queries specified the patient and intervention terms, but no control term and thus these variables could not be considered for the regression models to follow. In addition, the use of quotations in the query was also not further considered as only one search used quotes. The relationship between the primary predictors (frequency of searching and previous training in literature searching) and square root of sensitivity are presented in Figure 3. There appeared to be a linear relationship between the predictors and the outcome, which is required for a linear regression (assumption #1).

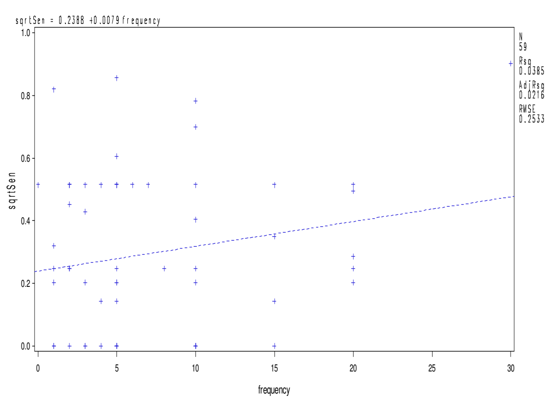


Figure 3a: Relationship between frequency of searching and the square root of sensitivity

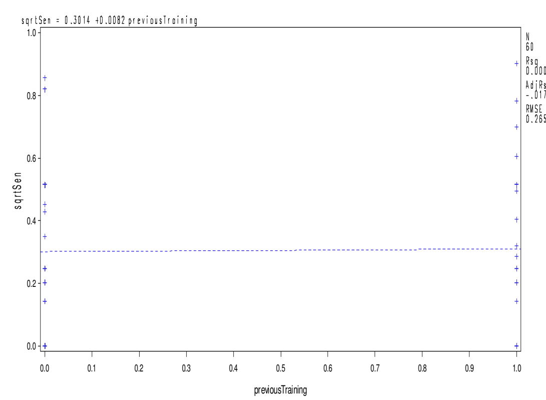


Figure 3b: Relationship between the use of previous training in literature searching the square root of sensitivity

<u>Build base model</u>

The base model consists of including the two primary predictors without any other covariates/confounders. Frequency of searching was included into the model as a continuous variable while previous training in literature searching was included into the model as a binary measure. Details of the analysis are presented in Table 1. The adjusted R-squared for this model was 0.06%, suggesting only 0.06% of the variance in the outcome is explained by the inclusion of the variables.

Table 1: Results of the analyses of the association between the primary predictors and square root of sensitivity

Variable	Estimate for change in the square root of sensitivity (95% Confidence intervals)	p-value
Frequency of searching	0.01 (-0.003 to 0.02)	0.198
Previously training in literature searching (referent group: No)	-0.004 (-0.1 to 0.14)	0.838

<u>Assess potential confounders</u>

To assess confounders, I calculated the percent change in the regression coefficients with and without the inclusion of a potential confounders (one confounder at a time); these results are presented in Table 2. Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. While all three potential confounders drastically changed the estimates of previous literature searching training, the inclusion of the confounders caused the R-squared to change to a negative measure. A negative R-squared may suggest that the model is not well specified with the inclusion of the confounders. I continued with the linear regression using all potential confounders, but also checked whether a Poisson or negative binomial regression was more appropriate.

Table 2: Percent change in the estimates with the addition of potential confounders

Confounding variable	Percent change in square root of sensitivity estimate for frequency of searching	Percent change in square root of sensitivity estimate for previous training in literature searching
Years practicing nephrology	11%	1085%
Sex (referent group: Females)	9%	56%
Age	8%	1169%
Practice in academic setting (referent group: No)	21%	63%

<u>Fitted model</u>

Results of the full model are presented in Table 3.

Table 3: Regression estimates of final model

Variable	Estimate for change in the square root of sensitivity (95% Confidence intervals)	p-value
Frequency of searching	0.01 (-0.006 to 0.02)	0.358
Previously received training in literature searching (referent group: No)	-0.01 (-0.16 to 0.18)	0.901
Years practicing nephrology	-0.01 (-0.04 to 0.01)	0.383
Age	0.01 (-0.02 to 0.04)	0.477
Sex (referent group: Females)	-0.04 (-0.25 to 0.17)	0.693
Practice in academic setting (referent group: No)	0.03 (-0.14 to 0.20)	0.731

<u>Diagnostics</u>

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model residuals appear to exhibit equal variance, and 5% of the residuals are outside the 95% confidence interval as would be expected. The normal probability plot also appears to be fitted with light tails and potential some outliers.

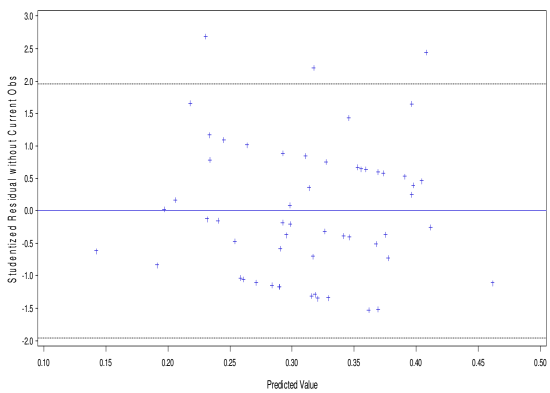


Figure 4a: Scatter plot of residuals

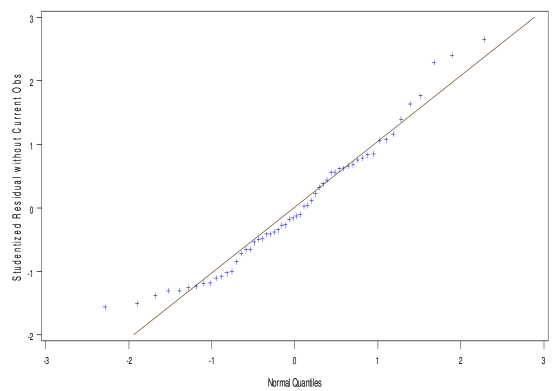


Figure 4b: Normal probability plot of residuals

A9.2 Poisson and negative binomial regression

Since there was some concern that the linear regression model was not specified well as indicated by the negative R-squared values, I repeated the model building for this outcome. If the Poisson or negative binomial proved to also be an accurate modeling type, this would also make comparisons across the models simpler.

<u>The assumptions of Poisson regression include:</u>

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

<u>Examination of outcome (sensitivity)</u>

To test the first assumption, I plotted the log sensitivity against the primary predictors (number of concepts and multifaceted searching). Since some searches did not find any relevant articles, the outcome includes several zero values; I added a slight offset of 1 in order to be able to calculate the log. This is presented in Figure 5. The relationships appeared linear as is required.

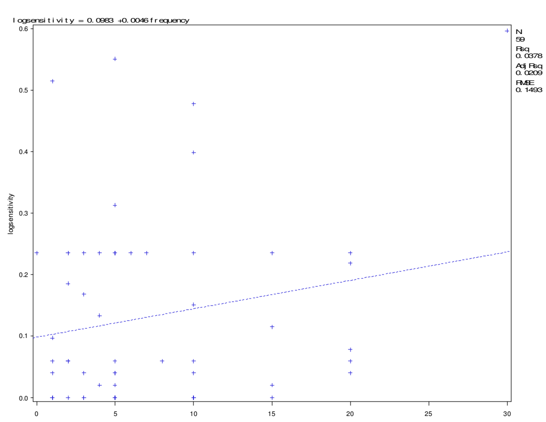


Figure 5a: Relationship between frequency of searching the log of sensitivity

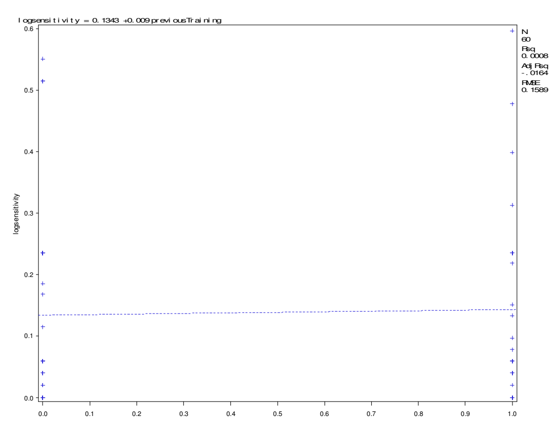


Figure 5b: Relationship between previous training and the log of sensitivity

Build base model

The base model consists of including the two primary predictors without any other covariates/confounders. The Deviance for the model is 10.6 suggesting that the Poisson regression is not appropriate. Instead, I applied the negative binomial regression and compared the Log Likelihood ratio which revealed a p-value of <0.001 suggesting that the negative binomial regression is a better fit. Estimates of the regression coefficients are presented in Table 4.

Table 4: Results of the analyses of the association between the primary predictors and sensitivity

Variable	Estimate of the rate ratio (95% Confidence Interval)	p-value
Frequency of searching	1.03 (0.97 to 1.09)	0.334
Previous training in literature searching (referent group: No)	0.99 (0.44 to 2.21)	0.983

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 5. As no potential confounders changed the regression coefficients for the main predictors by more than 10% none of the variables were included as confounders.

Table 5: Percent change in estimates with the addition of potential confounders

Confounding variable	Percent change in rate ratio estimate for frequency of searching	Percent change in rate ratio estimate for previous training in literature searching
Years practicing nephrology	0.1%	6%
Sex (referent group: Females)	0.4%	0.6%
Age	0.3%	6%
Practice in academic setting (referent group: No)	0.4%	0%

<u>Fitted model</u>

Results of the full model are same as the base model presented in table 5.

<u>Diagnostics</u>

The assessment of residuals is presented in Figure 6. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model appears adequately fitted as no residual values lie outside the 95% confidence interval (3 would be expected); however and the normal probability plot seems to have one heavy tail. Outliers were assessed.

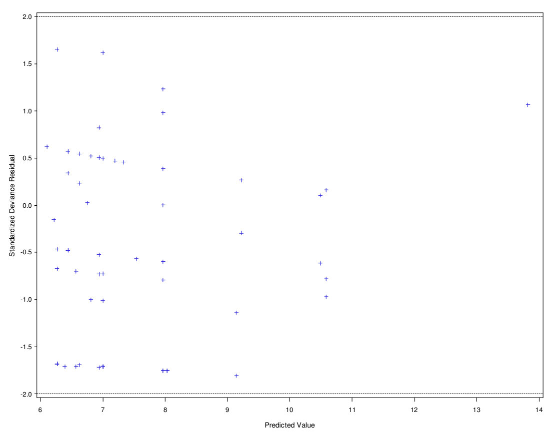


Figure 6a: Scatter plot of residuals

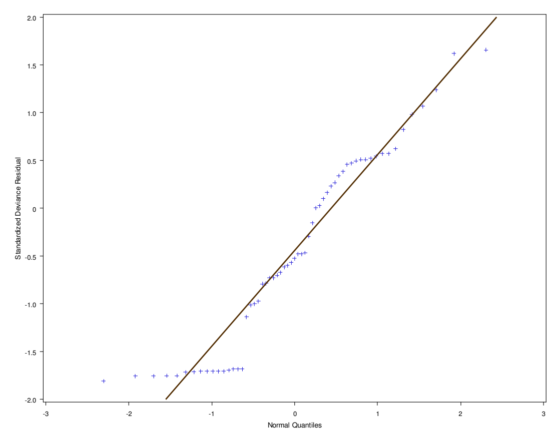


Figure 6b: Normal probability plot of residuals

A9.3 Comparing results from linear regression and negative binomial

Both models appear to equally fit the data from an analysis of the residuals; however, no association between physician characteristics and sensitivity was evident in either model. The effect measures for both models were close to unity with large p-values. In addition, the linear regression model resulted in a negative R-squared value.

Appendix 10: MODEL #4; Predictor: Nephrologist Characteristics; Outcome: Precision; Question 1

A10.1 Multivariable linear regression

The assumptions of linear regression include:

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

An examination of this model using linear regression identified similar discrepancies as with Model #3 with an R-squared reduction from 5% to 1% after the inclusion of candidate confounders. Results of the final fitted model and residuals are included here. Thus, I also conducted a Poisson/negative binomial regression for this model.

Fitted model

Results of the full model are presented in Table 1.

Table 1: Regression estimates of final model

Variable	Estimate for square root of precision (95% Confidence Interval)	p-value
Frequency of searching	0.004 (-0.004 to 0.1)	0.308
Previously received training in literature searching (referent group: No)	0.06 (-0.07 to 0.20)	0.346
Years practicing nephrology	-0.01 (-0.03 to 0.01)	0.365
Sex (referent group: Females)	-0.02 (0.01 to 0.03)	0.802
Age	0.008 (-0.01 to 0.03)	0.421

<u>Diagnostics</u>

The assessment of residuals is presented in Figure 1. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model specification appears to be appropriate as only two points lie outside the 95% confidence lines. The normal probability also appears to be well fitted with light tails and potentially a couple outliers.

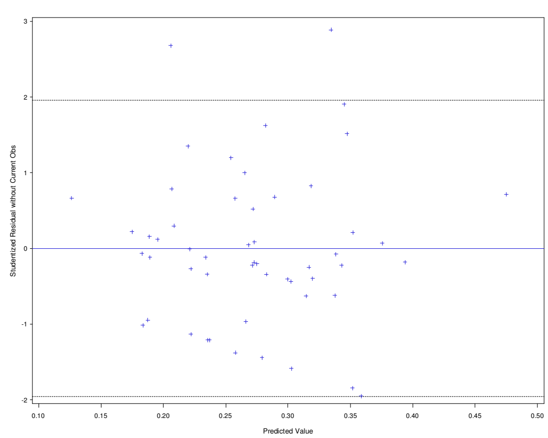


Figure 1a: Scatter plot of residuals

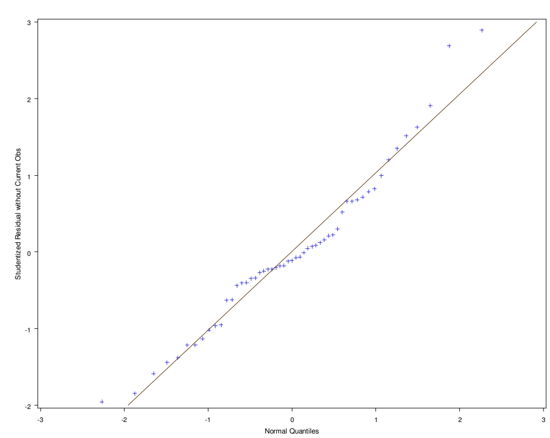


Figure 1b: Normal probability plot of residuals

A10.2 Poisson and negative binomial regression

<u>The assumptions of Poisson regression include:</u>

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

<u>Univariate analysis</u>

To test the first assumption, I plotted the log precision against the primary predictors (frequency of searching and previous training in literature searching). Since some searches did not find any relevant articles, the outcome includes several zero values; I added a slight offset of 1 in order to be able to calculate the log. This is presented in Figure 2. The relationships appear linear as is required.

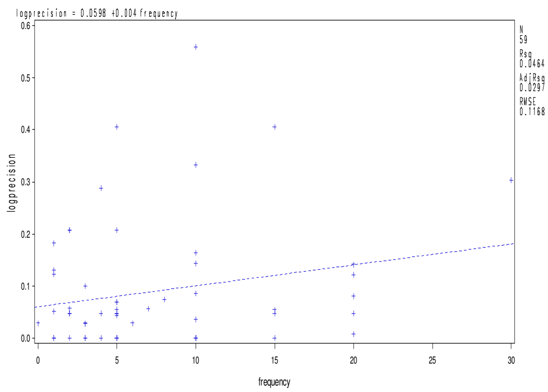


Figure 2a: Plot comparing the log of precision against the frequency of searching

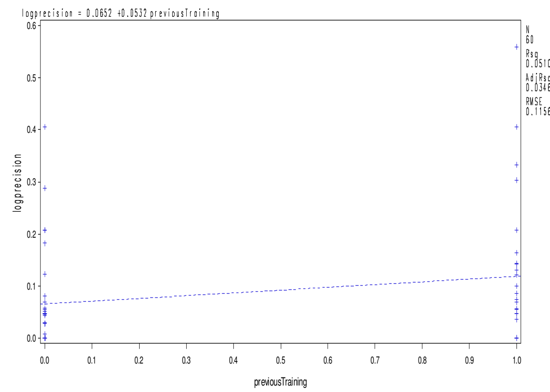


Figure 2b: Plot comparing the log of precision against previous training in literature searching

<u>Build base model</u>

The base model consists of including the two primary predictors without any other covariates/confounders. The Deviance for the model is 4.1 suggesting that the Poisson regression is not appropriate. Instead, I applied the negative binomial regression and compared the Log Likelihood ratio which revealed a p-value of <0.001 suggesting that the negative binomial regression is a better fit. Estimates of the regression coefficients are presented in Table 2. Regression coefficients equate to the log of the rate ratio for every one unit increase in the predictor variable. Accordingly, Table 2 also presents the exponentiated regression coefficients (the rate ratio).

Table 2: Results of the analyses of the association between the primary predictors and precision

Variable	Estimate log rate ratio of precision (95% CI)	Estimate for the rate ratio of precision (95% CI)	p-value
Frequency of searching	0.02 (-0.01 to 005)	1.01 (0.99 to 1.05)	0.187
Previously received training in literature searching (referent group: No)	0.88 (0.43 to 1.33)	2.41 (1.53 to 3.79)	<0.01

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 3. As only the number of years practicing nephrology and sex changed effect measures by more than 10%, they were included in the final model.

Table 3: Percent change in the estimates with the addition of potential confounders

Confounding variable	Percent change in rate ratio of precision one unit change in frequency of searching	Percent change in rate ratio for precision when comparing previous training in literature searching to no training
Years practicing nephrology	0.3%	12%
Sex (referent group: Females)	0%	19%
Age	0.3%	3%
Practice in academic setting	0.1%	0 %

Fitted model

Results of the full model are presented in Table 4.

Table 4: Regression estimates of final model

Variable	Estimate of rate ratio for precision (95% Confidence Interval)	p-value
Frequency of searching	1.02 (0.99 to 1.05)	0.118
Previously received training in literature searching (referent group: No)	2.27 (1.43 to 3.62)	<0.001
Years practicing nephrology	0.99 (0.97 to 1.01)	0.562
Sex (referent group: Females)	0.72 (0.42 to 1.22)	0.219

Diagnostics

The assessment of residuals is presented in Figure 3. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model specification appears to be appropriate as few points lie outside the 95% confidence lines. Four points are outside the limits; however two of them are very close to the boundaries, the other two may be outliers and were examined. The normal probability appears to be well fitted with light tails.

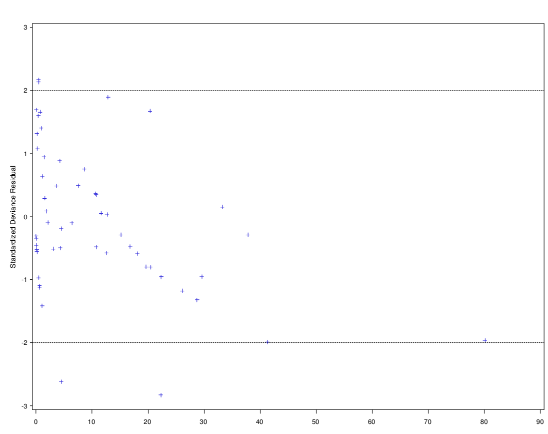


Figure 3a: Scatter plot of residuals

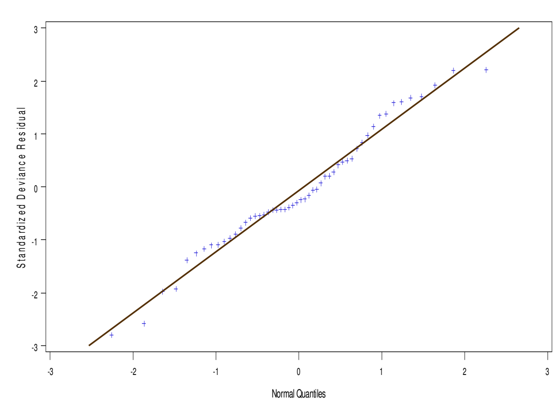


Figure 2b: Normal probability plot of residuals

A10.3 Comparing results from linear regression and negative binomial regression

Both models appear to equally fit the data from an analysis of the residuals; however the linear regression model resulted in a lower R-squared value for the final model, compared to the base model. The estimates received from the linear regression support the same directions of effects seen in the negative binomial regression, providing evidence of the robustness of the data and its' effects.

Appendix 11: MODEL #5; Predictor: Search Query Characteristics; Outcome: Sensitivity; Question 2

A11.1 Multivariable linear regression

<u>The assumptions of linear regression include:</u>

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

<u>Examination of outcome (sensitivity)</u>

The examination of the outcome revealed that sensitivity operates like a discrete measure. Sensitivity takes the values of 0, 0.25, 0.5, 0.75 or 1, corresponding to the maximum of 4 relevant articles for this clinical question (Figure 1). In this situation linear regression was not a good option as the corresponding residuals did not appear normally distributed with unequal variance (Figure 2; Table 1). However, the R-squared was 33% for the model, which was a slight improvement from the base model.

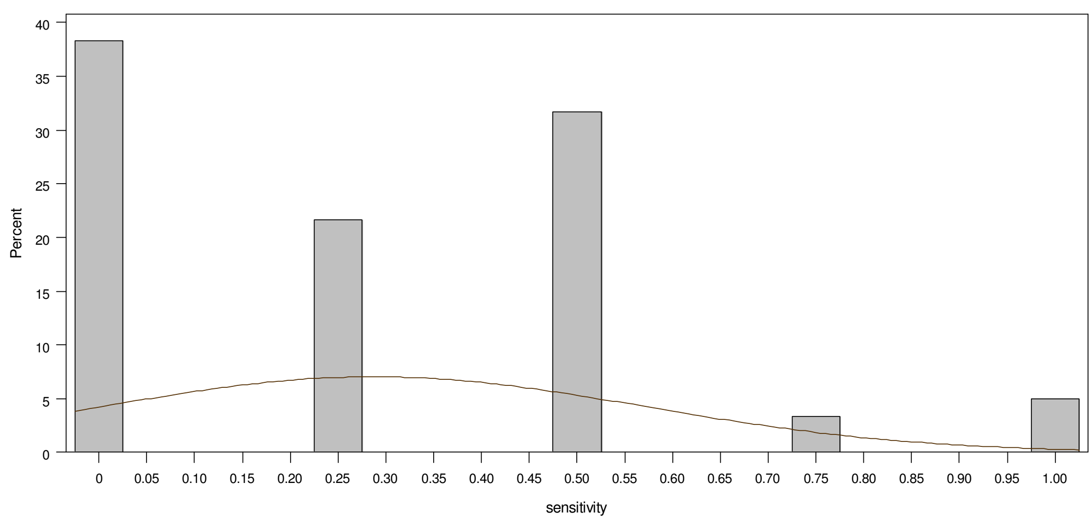


Figure 1: Histogram of sensitivity

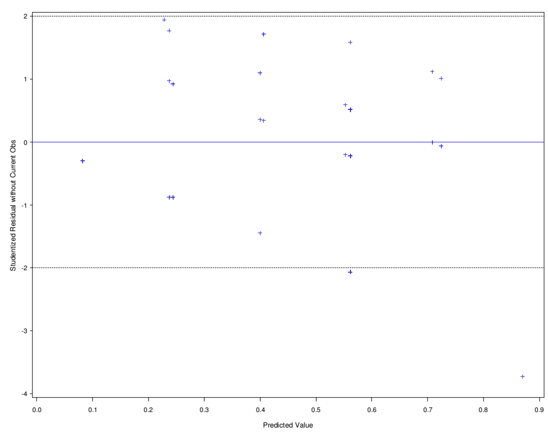


Figure 2a: Scatter plot of residuals

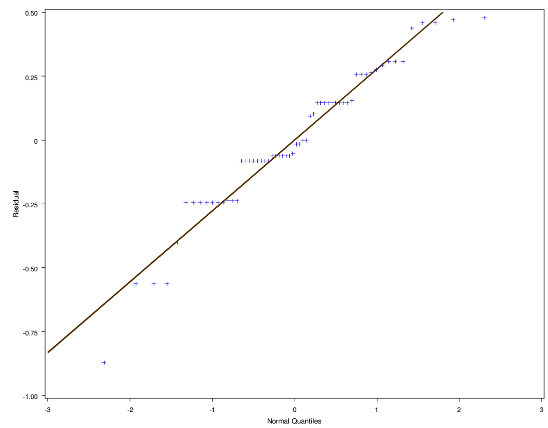


Figure 2b: Normal probability plot of residuals

Table 1: Regression estimates of final linear regression model

Variable	**Estimate for square root of sensitivity (95% Confidence intervals)**	**p-value**
Number of concepts	-0.16 (-0.26 to -0.07)	0.001
Use of multifaceted search features (referent group: No)	0.31 (0.08 to 0.54)	0.010
Outcome term used in search (referent group: No)	-0.16 (-0.33 to 0.02)	0.074

A11.2 Selecting appropriate regression model

An alternative option would be to use ordinal logistic regression, categorizing the data based on the number of relevant articles found. To assess whether ordinal logistic regression could be used, I created frequency tables of the outcome versus the primary predictors (Tables 2 & 3). These tables reveal several cells with very few data points, and thus ordinal logistic regression was not an option as the model would not run adequately. Instead, I chose to use Poisson regression.

Table 2: Frequency table comparing the number of relevant articles found to the number of concepts.

	Number of Concepts						
	Frequency	**1**	**2**	**3**	**4**	**5**	***Total***
Number of relevant articles found	**0**	0	4	8	11	0	*23*
	1	0	7	5	1	0	*13*
	2	2	9	6	1	1	*19*
	3	0	2	0	0	0	*2*
	4	1	1	1	0	0	*3*
	Total	*3*	*23*	*20*	*13*	*1*	*60*

Table 3: Frequency table comparing the number of relevant articles found to the use of multifaceted search features

	Use of multifaceted search features			
	Frequency	**0**	**1**	***Total***
Number of relevant articles found	**0**	22	1	*23*
	1	12	1	*13*
	2	15	4	*19*
	3	2	0	*2*
	4	2	1	*3*
	Total	*53*	*7*	*60*

A11.3 Poisson and negative binomial regression

<u>The assumptions of Poisson regression include:</u>

4. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
5. Outcome has variance equal to the mean (equidispersion).
6. The standardized deviance residuals are approximately normally distributed with equal variance.

<u>Examination of outcome (sensitivity)</u>

The relationship between the primary predictors (number of concept terms and use of multifaceted search features) and the log of sensitivity is presented in Figure 3. As some searches retrieved no results, this resulted in sensitivity values of zero. To calculate the log, I added a nominal value of 1 to the sensitivity. There appears to be a linear relationship between the predictors and the outcome, which is required for a Poisson regression (assumption #1).

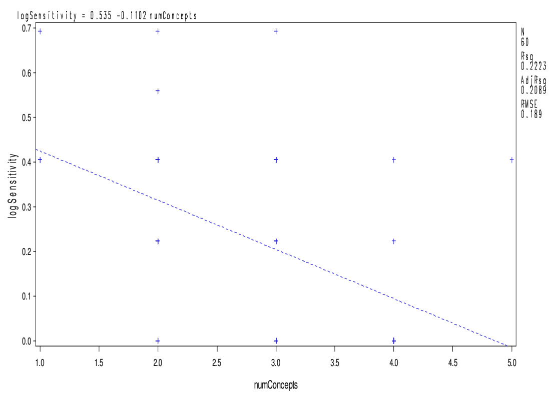


Figure 3a: Relationship between number of concept terms and the log of sensitivity

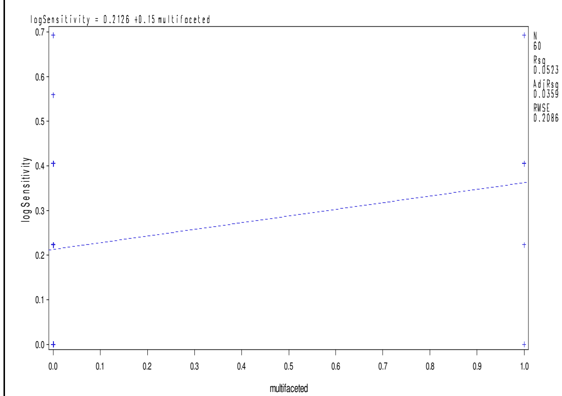


Figure 3b: Relationship between the use of multifaceted search features and the log of sensitivity

Build base model

The base model consists of including the two primary predictors without any other covariates/confounders. The Deviance for the model is 1.04 suggesting that the Poisson regression can be used. The measure of number of concept terms was modeled as a quantitative covariate. Multifaceted searching was included into the model as a binary measure. Estimates of the regression coefficients are presented in Table 4.

Table 4: Results of the analyses of the association between the primary predictors and sensitivity

Variable	Estimate log rate ratio (95% Confidence Interval)	Estimate for rate ratio (95% Confidence Interval)	p-value
Number of concept terms	-0.62 (-0.92 to -0.32)	0.54 (0.39 to 0.73)	<0.001
Use of multifaceted search features (referent group: No)	0.87 (0.24 to 1.49)	2.34 (1.27 to 1.49)	0.007

<u>Assess potential confounders</u>

Confounders were added to the model if they changed the rate ratio estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 5. As only the inclusion of the outcome term variable changed the rate ratio by more than 10%, it was the only confounder added into the fitted model.

Table 5: Percent change in the estimates with the addition of potential confounders

Confounding variable	**Percent change in rate ratio estimate for number of concepts**	**Percent change in rate ratio estimate for use of multifaceted search features**
Outcome term used in search (referent group: No)	13%	5%
Control term used in search (referent group No)	3%	6%
Acronym term used in search (referent group: No)	0.2%	3%

<u>Fitted model</u>

Results of the full model are presented in Table 6.

Table 6: Regression estimates of final model

Variable	**Estimate for rate ratio of sensitivity (95% Confidence intervals)**	**p-value**
Number of concepts	0.61 (0.43 to 0.85)	0.004
Use of multifaceted search features (referent group: No)	2.27 (1.21 to 4.25)	0.011
Outcome term used in search (referent group: No)	0.64 (0.34 to 1.22)	0.176

Diagnostics

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model appears to fit adequately. Only one residual value is outside the 95% confidence interval. The normal probability plot, on the other hand appears granular, while the points are clustered around the normal line. This suggests some deviation from the normality assumption. In addition, the two plots appear very similar to those seen in the linear regression model.

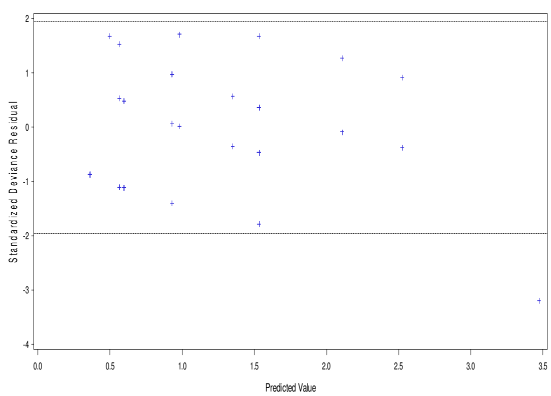


Figure 4a: Scatter plot of residuals

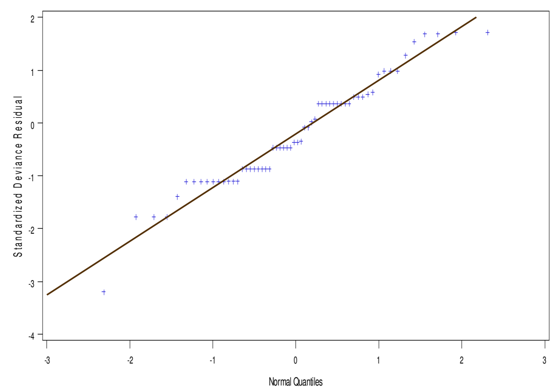


Figure 4b: Normal probability plot of residuals

A11.4 Comparing results from linear regression and Poisson regression

An analysis of the residuals suggests that the Poisson regression fits the data better than the linear regression model (as there is no assumption on the equal variance of the residuals for Poisson regression). However, the estimates received from the linear regression support those seen in the Poisson regression, providing evidence of the robustness of the data and its associations.

Appendix 12: MODEL #6; Predictor: Search Query Characteristics; Outcome: Precision; Question 2

A12.1 Multivariable linear regression

<u>The assumptions of linear regression include:</u>

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

<u>Examination of outcome (precision)</u>

The histogram of the outcome (Figure 1) revealed that the outcome is positively skewed. A solution to remedying a positive skew is to take the log or the square root of the outcome. As the precision consists of a significant number of zeros (no relevant articles were found), a log cannot be taken. Instead, I transformed the data with a square root. The transformed outcome is presented in Figure 2. The transformation made a slight difference in the skewness and so I continued the analysis with this transformed outcome.

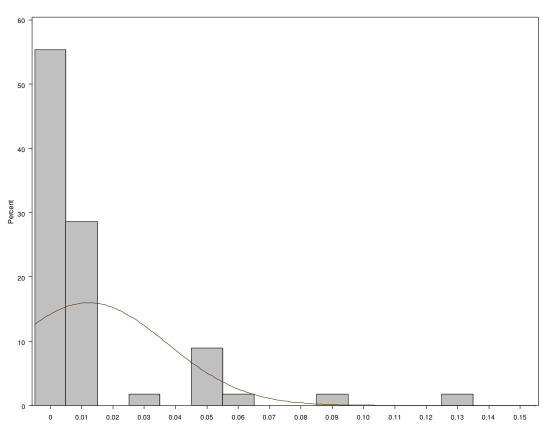


Figure 1a: Histogram of precision

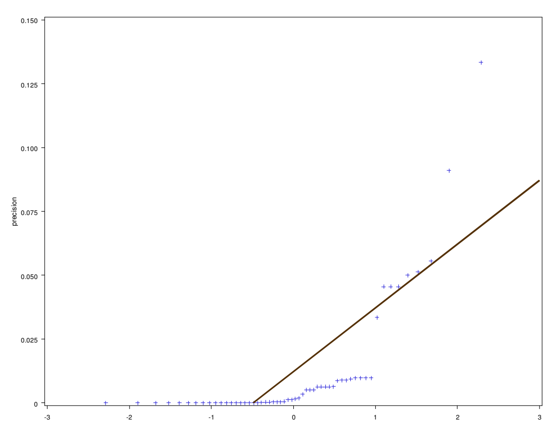


Figure 1b: Normal probability plot of precision

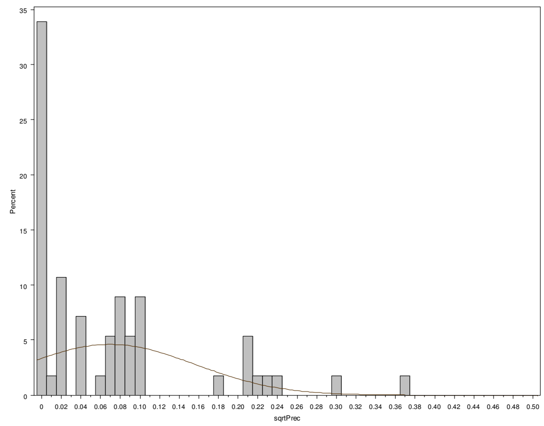


Figure 2a: Histogram of square root of precision

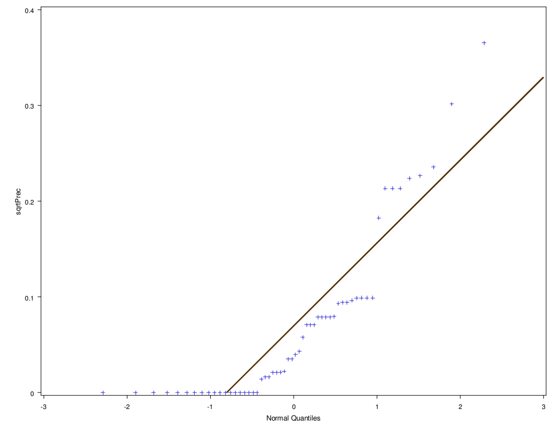


Figure 2b: Normal probability plot of square root of precision

Univariate analysis

The relationship between the primary predictors (number of concept terms and use of multifaceted search features) are presented in Figure 3. There appears to be a linear relationship between the predictors and the outcome, which is required for a linear regression (assumption #1).

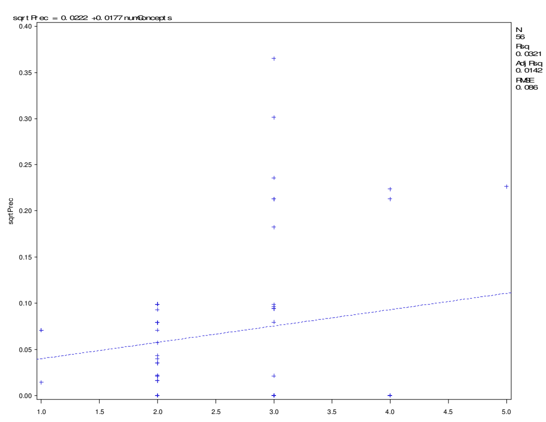


Figure 3a: Relationship between number of concept terms and the square root of precision

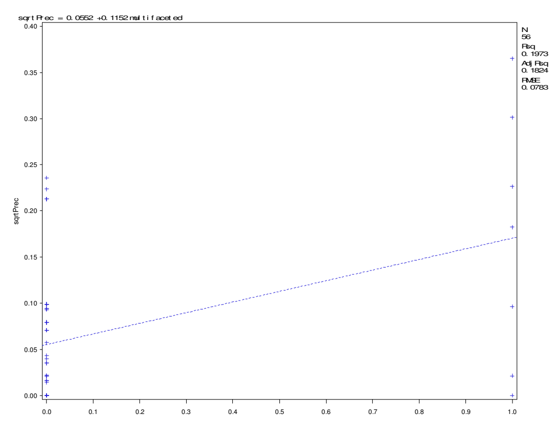


Figure 3b: Relationship between the use of multifaceted search features and square root of precision

<u>Base model</u>

The base model consists of including the two primary predictors without any potential confounders. The measure of number of concepts was modeled as a quantitative covariate, while multifaceted searching was included as binary measure. Details of the analysis are presented in Table 1. The adjusted R-squared for this model was 18%, suggesting that 18% of the variance in the outcome is explained by the inclusion of the variables.

Table 1: Results of the analyses of the association between the primary predictors and the square root of precision.

Variable	Estimate for square root of precision (95% Confidence intervals)	p-value
Number of concept terms	0.01 (-0.02 to 0.04)	0.460
Use of multifaceted search features (referent group: No)	0.11 (0.05 to 0.18)	0.001

<u>Assess potential confounders</u>

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 2. As only the outcome variable changed the regression estimate for the number of concept terms by more than 10%, it was the only variable included into the fitted model.

Table 2: Percent change in estimates with the addition of potential confounders

Confounding variable	Percent change in square root of sensitivity estimate for number of concepts	Percent change in square root of sensitivity estimate for use of multifaceted search features
Outcome term used in search (referent group: No)	5%	0%
Acronym term used in search (referent group: No)	41%	7%
Control term used in search (referent group: No)	6%	1%

Fitted model

Results of the full model are presented in Table 3. The adjusted R-squared for this model remained at 18%, when compared to the base model.

Table 3: Regression estimates of final model

Variable	Estimate for change in the square root of sensitivity (95% Confidence intervals)	p-value
Number of concept terms	0.07 (-0.02 to 0.03)	0.674
Use of multifaceted search features (referent group: No)	0.12 (0.05 to 0.18)	<0.001
Acronym term used in search (referent group: No)	0.02 (-0.02 to 0.07)	0.283

Diagnostics

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model specification appears to be poor. The residuals appear to exhibit equal variance, albeit with some clustering, however 4/56 (7%) points appear outside the 95% confidence lines, and only 2-3 would be expected. The normal probability appears to be poorly fitted with heavy tails. Instead, I chose to explore the Poisson/negative binomial regression models for this data.

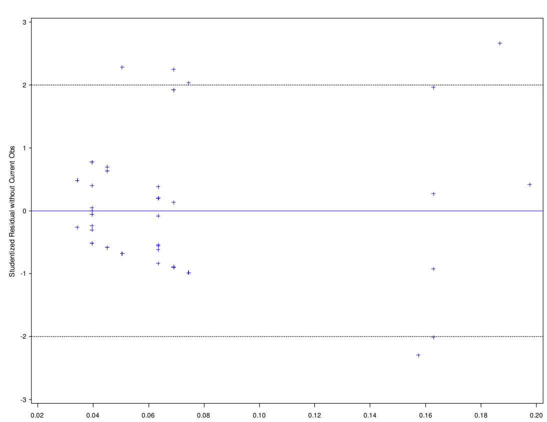


Figure 4a: Scatter plot of residuals

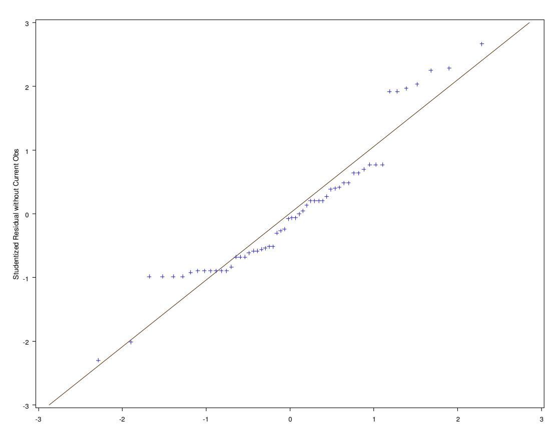


Figure 4b: Normal probability plot of residuals

A12.2 Poisson or negative binomial regression

The assumptions of Poisson regression include:

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

Univariate analysis

To test the first assumption, I plotted the log precision against the primary predictors (number of concepts and multifaceted searching). Since some searches did not find any relevant articles, the outcome includes several zero values; I added a slight offset of 1 in order to be able to calculate the log. This is presented in Figure 5. The relationships appear linear as is required.

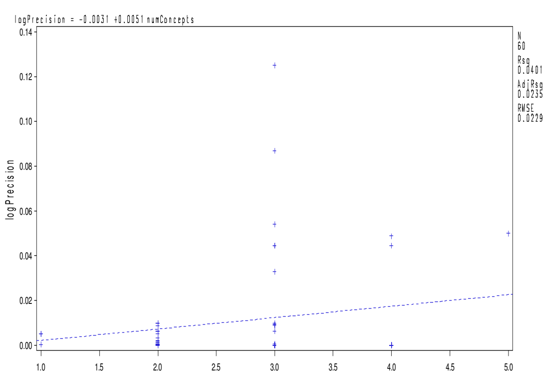


Figure 5a: Relationship between number of concept terms and the log of precision

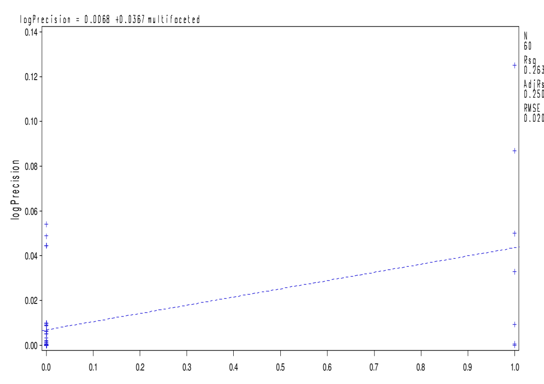


Figure 5b: Relationship between the use of multifaceted search features and the log of precision

Build base model

The base model consists of including the two primary predictors without any other covariates/confounders. The measure of number of concept terms was modeled as a quantitative covariate. Multifaceted searching was included into the model as a binary measure. The Deviance for the Poisson model is 2.6 suggesting that the Poisson regression is not an appropriate model. Instead, I applied the negative binomial regression and compared the Log Likelihood ratio which revealed a p-value of <0.001 suggesting that the negative binomial regression is a better fit. Estimates of the rate ratio are presented in Table 4.

Table 4: Results of the analyses of the association between the primary predictors and precision

Variable	**Estimate for rate ratio of precision (95% Confidence Interval)**	**p-value**
Number of concept terms	2.28 (1.30 to 3.98)	0.004
Use of multifaceted search features (referent group: No)	3.00 (0.77 to 11.66)	0.118

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 5. As only the inclusion of the outcome variable changed the rate ratio by more than 10%, it was the only confounder added into the fitted model.

Table 5: Percent change in the estimates with the addition of potential confounders

Confounding variable	Percent change in rate ratio estimate for number of concepts	Percent change in rate ratio estimate for use of multifaceted search features
Outcome term used in search (referent group: No)	3%	10%
Control term used in search (referent group No)	4%	4%
Acronym term used in search (referent group: No)	2%	4%

<u>Fitted model</u>

Results of the full model are presented in Table 6.

Table 6: Regression estimates of final model

Variable	Estimate for rate ratio of precision (95% Confidence intervals)	p-value
Number of concept terms	2.21 (1.24 to 3.91)	0.007
Use of multifaceted search features (referent group: No)	2.71 (0.67 to 10.90)	0.159
Outcome term used in search (referent group: No)	1.39 (0.44 to 4.31)	0.566

<u>Diagnostics</u>

The assessment of residuals is presented in Figure 6. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model appears to fit well as only one residual value is outside the 95% confidence interval and the normal probability plot appears fitted, except for a couple potential outliers, which were assessed.

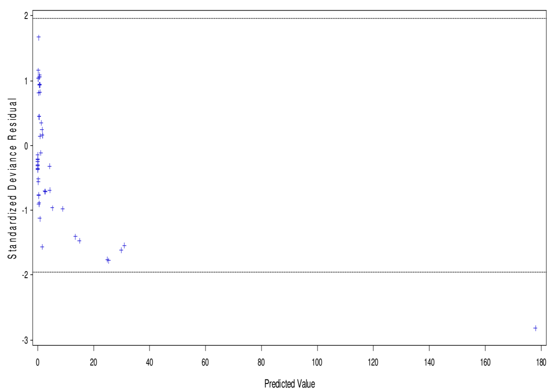
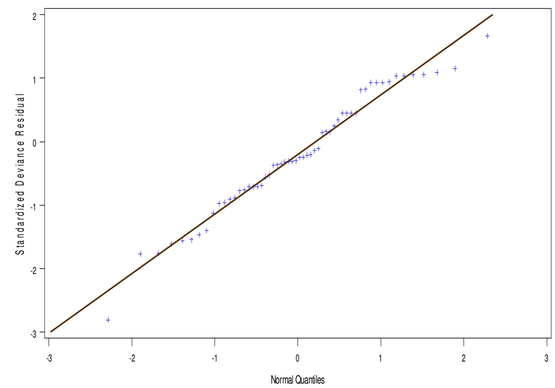
	

Figure 6a: Scatter plot of residuals	**Figure 6b: Normal probability plot of residuals**

A12.3 Comparing results from linear regression and negative binomial

An analysis of the residuals suggests that the negative binomial model fits the data better than the linear regression model. Unfortunately, the significance tests of the linear regression did not support those received from negative binomial regression, albeit the associations followed the same directions of effect. Sensitivity analyses removing potential outliers did not resolve these differences.

Appendix 13: MODEL #7; Predictor: Nephrologist Characteristics; Outcome: Sensitivity; Question 2

A13.1 Multivariable linear regression

<u>The assumptions of linear regression include:</u>

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

<u>Examination of outcome (sensitivity)</u>

The examination of the outcome revealed that sensitivity operates like a discrete measure. Sensitivity takes the values of 0, 0.25, 0.5, 0.75 or 1, corresponding to the maximum of 4 relevant articles for this clinical question (Figure 1). In this situation linear regression was not a good option as the corresponding residuals did not appear normally distributed with unequal variance (Figure 2; Table 1). In addition, the R-squared for the model was -9% suggesting a poor fit.

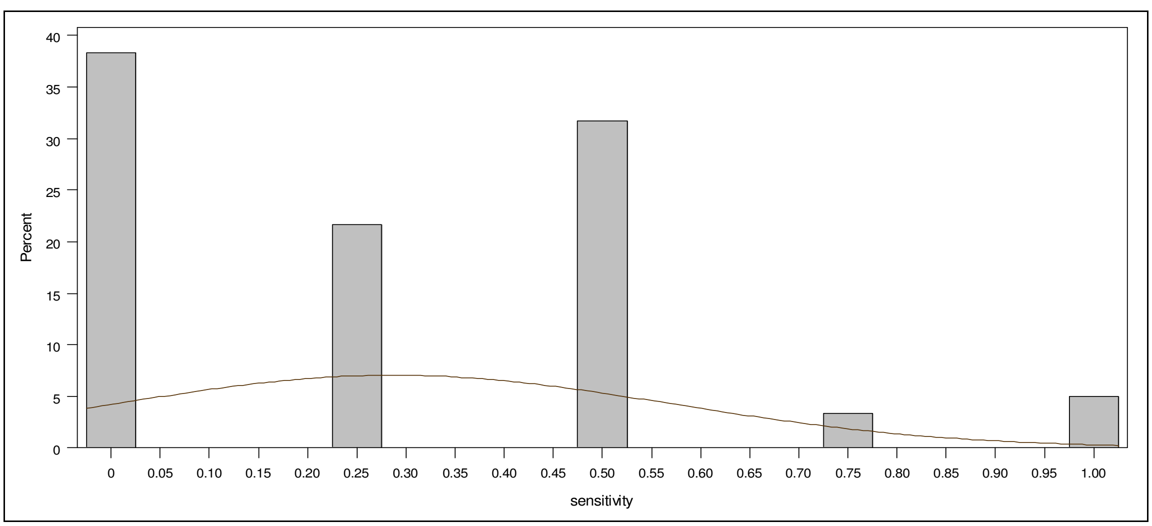


Figure 1: Histogram of sensitivity

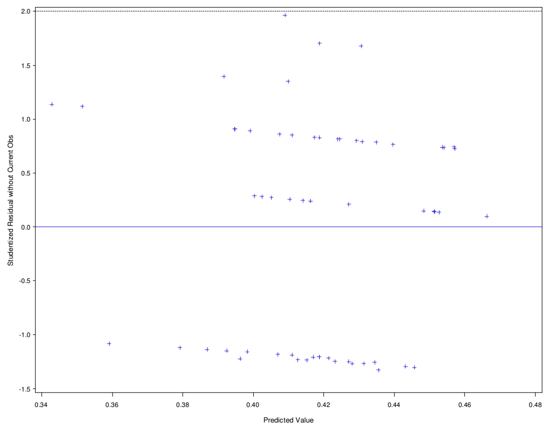


Figure 2a: Scatter plot of residuals

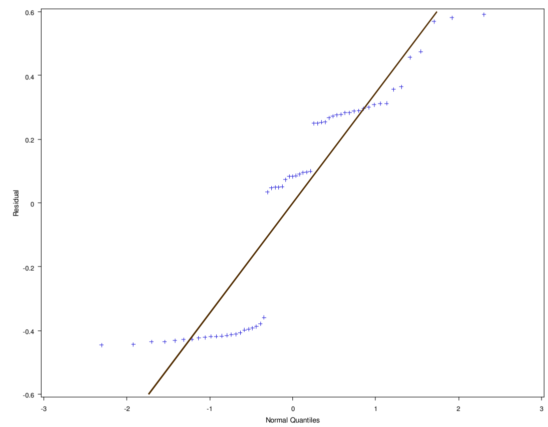


Figure 2b: Normal probability plot of residuals

Table 1: Regression estimates of final linear regression model

Variable	**Estimate for square root of sensitivity (95% Confidence intervals)**	**p-value**
Frequency of searching	-0.03 (-0.02 to 0.01)	0.701
Previous training in literature searching (referent group: No)	-0.004 (-0.24 to 0.23)	0.973
Age	0.001 (-0.04 to 0.04)	0.948
Sex (referent group: Females)	-0.04 (-0.33 to 0.25)	0.778
Years practicing nephrology	-0.003 (-0.04 to 0.03)	0.870
Practice in an academic setting (referent group: No)	0.11 (-0.13 to 0.35)	0.362

A13.2 Selecting appropriate regression model

An alternative option would be to use ordinal logistic regression, categorizing the data based on the number of relevant articles found. To assess whether ordinal logistic regression can be used, I created frequency tables of the outcome versus the primary predictors. These tables revealed several cells with very few data points, and thus ordinal logistic regression was not an option as the model will not run adequately. Instead, I chose to use Poisson regression.

A13.3 Poisson and negative binomial regression

<u>The assumptions of Poisson regression include:</u>

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

<u>Univariate analysis</u>

The relationship between the primary predictors (frequency of searching and previous training in literature searching) and the log of sensitivity is presented in Figure 3. As some searches received no results, this produced values of zero for sensitivity. Thus, in order to calculate the log, I added a nominal value of 1 to the sensitivity values. There appears to be a linear relationship, albeit very little relationship, between the predictors and the outcome, which is required for a Poisson regression (assumption #1).

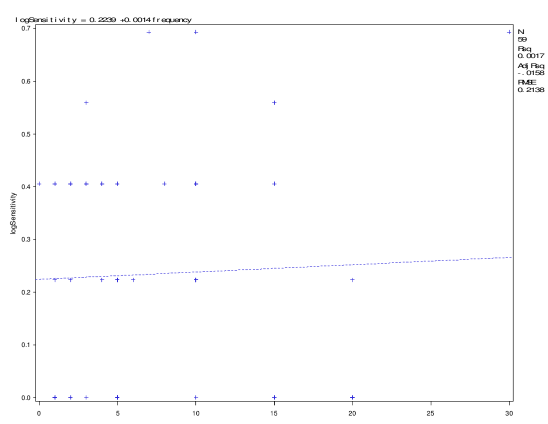


Figure 3a: Relationship frequency of searching and the log of sensitivity

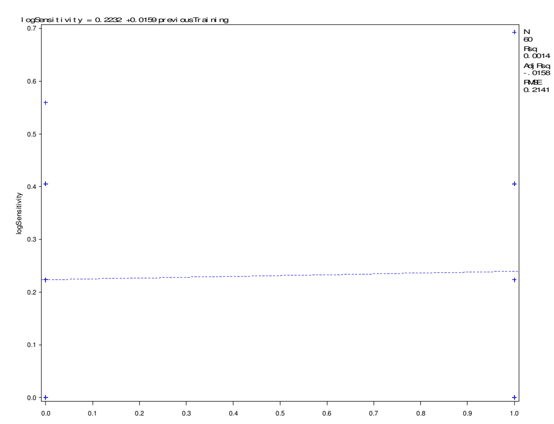


Figure 3b: Relationship between previous training in literature searching and the log of sensitivity

<u>Build base model</u>

The base model consists of including the two primary predictors without any other covariates/confounders. The Deviance for the model is 1.38, a value close to 1, suggesting that the Poisson regression can be used. The measure of frequency of searching was modeled as a quantitative covariate. Previous training in literature searching was included into the model as a binary measure. Estimates of the rate ratio are presented in Table 2.

Table 2: Results of the analyses of the association between the primary predictors and sensitivity

Variable	Estimate for rate ratio (95% Confidence intervals)	p-value
Frequency of searching	1.01 (0.97 to 1.05)	0.604
Previous training in literature searching (referent group: No)	1.07 (0.65 to 1.73)	0.814

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by a minimum of 10%. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 3. As none of the variables changed the effect measures by more than 10%, none of them were included in the fitted model.

Table 3 Percent change in the estimates the addition of potential confounders

Confounding variable	Percent change in rate ratio estimate for frequency of searching	Percent change in rate ratio estimate for previous training in literature searching
Years practicing nephrology	0%	3%
Sex (referent group: Females)	0%	1%
Age	0%	4%
Practice in academic setting (referent group: No)	1%	1%

Fitted model

Results of the full model are the same as the base model presented in Table 2.

Diagnostics

The assessment of residuals is presented in Figure 3. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model appears to be poorly fitted. The residuals appear to exhibit equal variance (assumption #2) and only two residual values lie outside the 95% confidence interval (3 would be expected); however and the normal probability plot is granular with points clustered around the normal probability line. The regression diagnostics for this model suggest an improvement from the linear regression specification discussed earlier, although both are poorly fitted.

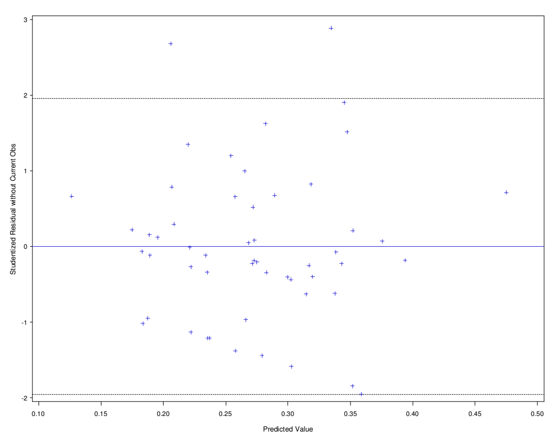


Figure 4a: Scatter plot of residuals

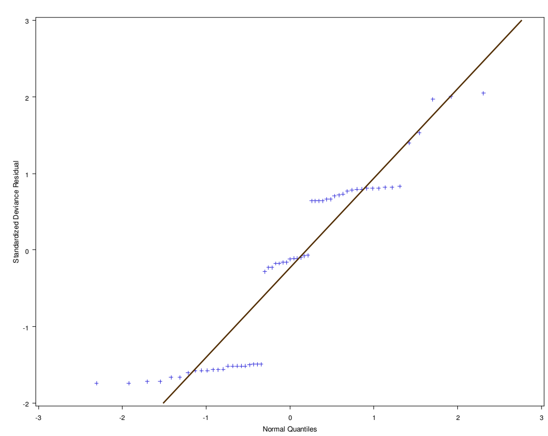


Figure 4b: Normal probability plot of residuals

A13.3 Comparing results from linear regression and Poisson regression

An analysis of the residuals suggests that the Poisson regression model fits the data better than the linear regression model; however, no associations between physician characteristics and sensitivity were evident in either model. The effect measures for both models were close to unity with large p-values. In addition, the linear regression model resulted in a negative R-squared value.

Appendix 14: MODEL #8; Predictor: Nephrologist Characteristics; Outcome: Precision; Question 1

A14.1 Multivariable linear regression

<u>The assumptions of linear regression include:</u>

1. The relationship between the outcomes and the predictors is (approximately) linear.
2. The error term has zero mean.
3. The error term has constant variance.
4. The errors are uncorrelated.
5. The errors are normally distributed or we have an adequate sample size to rely on large sample theory.

<u>Examination of outcome (precision)</u>

The histogram of the outcome (Figure 1) revealed that the precision is positively skewed. A solution to remedying a positive skew is to take the log or the square root of the outcome. As the outcome consists of a significant number of zeros (no relevant articles were found), a log cannot be taken. Instead, I transformed the data with a square root. The transformed outcome is presented in Figure 2. The transformation made a slight difference in the skewness and so I continued the analysis with this transformed outcome.

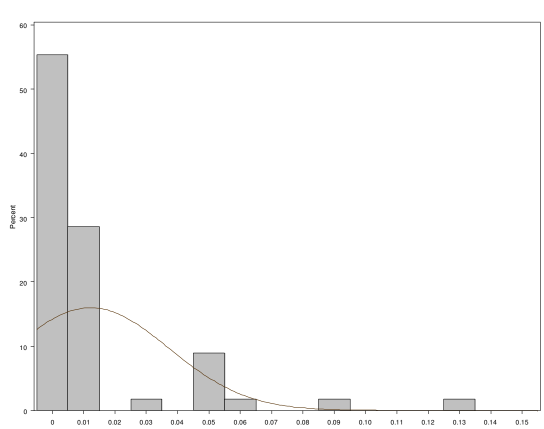


Figure 1a: Histogram of precision

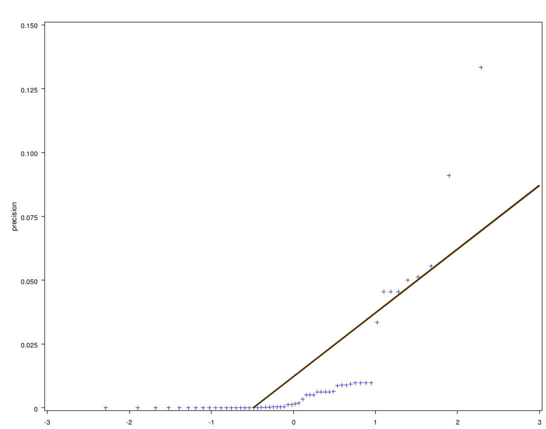


Figure 1b: Normal probability plot of precision

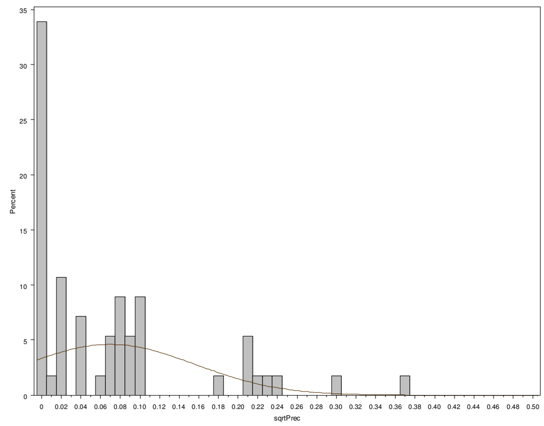


Figure 2a: Histogram of square root of precision

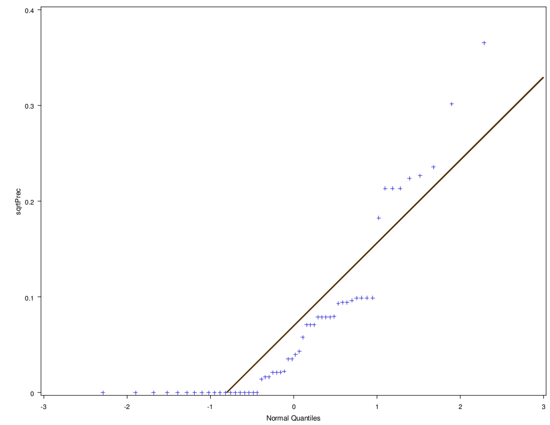


Figure 2b: Normal probability plot of square root of precision

<u>Univariate analysis</u>

The relationship between the primary predictors (frequency of searching and previous training in literature searching) and square root of precision are presented in Figure 3. There appears to be a linear relationship between the predictors and the outcome, which is required for a linear regression (assumption #1).

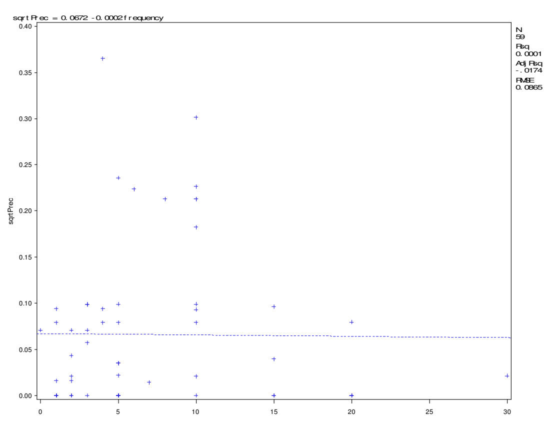


Figure 3a: Relationship frequency of searching and the square root of precision

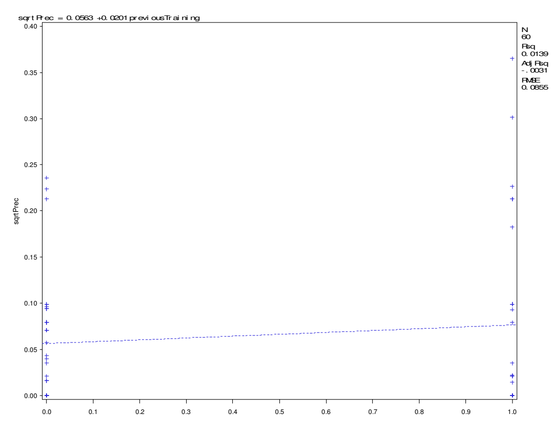


Figure 3b: Relationship between previous training in literature searching and the square root of precision

Build base model

The base model consists of including the two primary predictors without any potential confounders (Table 1). The adjusted R-squared for this model was negative suggesting a poor fit for the data.

Table 1: Results of the analyses of the association between the primary predictors and the square root of precision.

Variable	Estimate for change in the square root of precision (95% Confidence intervals)	p-value
Frequency of searching	0.0 (-.0004 to 0.003)	0.752
Previous training in literature seraching (referent group: No)	0.02 (-0.03 to 0.07)	0.390

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 2. As all variables changed the regression estimates by more than 10%, all were included into the fitted model. However, the R-squared continued to be negative with the addition of the confounders.

Table 2 Percent change in the estimates with the addition of potential confounders

Confounding variable	Percent change in square root of sensitivity estimate for number of concepts	Percent change in square root of sensitivity estimate for use of multifaceted search features
Years practicing nephrology	57%	57%
Sex (referent group: Females)	39%	17%
Age	51%	73%
Practice in academic setting	150%	25%

<u>Fitted model</u>

Results of the full model are presented in Table 3. The adjusted R-squared was also negative.

Table 3: Regression estimates of final model

Variable	**Estimate for change in the square root of precision (95% Confidence intervals)**	**p-value**
Frequency of searching	-0.002 (-0.006 to 0.003)	0.555
Previous training in literature searching (referent group: No)	0.04 (-0.02 to 0.09)	0.217
Age	0.003 (-0.01 to 0.01)	0.608
Sex (referent group: Females)	-0.003 (-0.08 to 0.07)	0.928
Years practicing nephrology	-0.002 (-0.01 to 0.01)	0.745
Practice in academic setting (referent group: No)	0.02 (-0.04 to 0.08)	0.521

<u>Diagnostics</u>

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model specification appears to be poor. The residuals appear to exhibit equal variance, albeit with some clustering. Four (4/56; 7%) points lie outside the 95% confidence lines, and only 2-3 would be expected. However, three of them are very close to the boundaries. The normal probability appears to be poorly fitted with heavy tails. Instead, I chose to explore the Poisson/negative binomial regression models for this data.

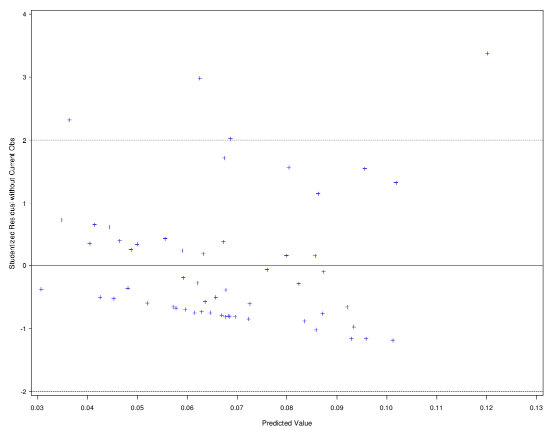


Figure 4a: Scatter plot of residuals

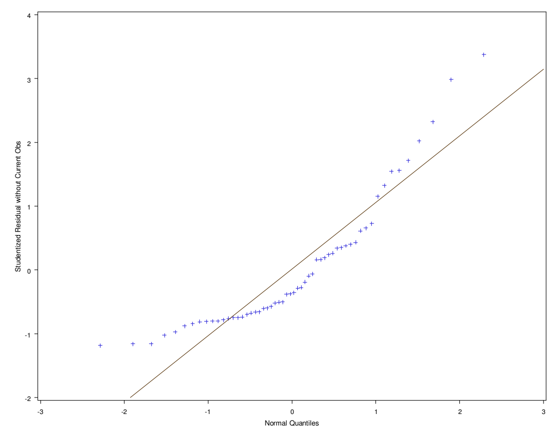


Figure 4b: Normal probability plot of residuals

A14.2 Poisson or negative binomial regression

<u>The assumptions of Poisson regression include:</u>

1. Logarithm of the outcome rate changes linearly with equal increment increases in the exposure variable.
2. Outcome has variance equal to the mean (equidispersion).
3. The standardized deviance residuals are approximately normally distributed with equal variance.

<u>Univariate analysis</u>

The relationship between the primary predictors (frequency of searching and previous training in literature searching) and the log of precision is presented in Figure 5. As some searches received no results, this resulted in precision values of zero. To calculate the log, I added a nominal value of 1 to the precision. There appears to be a linear relationship between the predictors and the outcome, which is required for a Poisson regression (assumption #1).

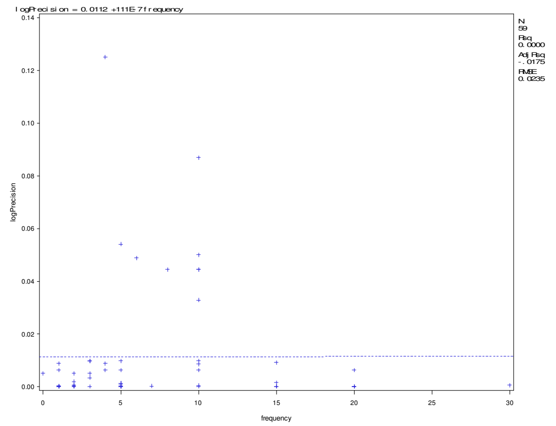


Figure 5a: Relationship frequency of searching and the log of precision

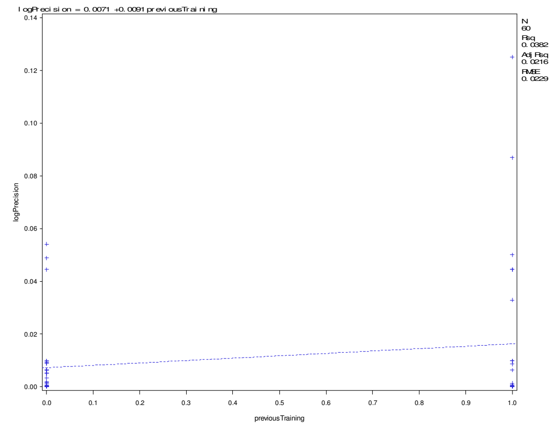


Figure 5b: Relationship between previous training in literature searching and the log of precision

<u>Build base model</u>

The base model consists of including the two primary predictors without any other covariates/confounders. The Deviance for the Poisson model is 3.5 suggesting that the Poisson regression is not an appropriate model. Instead, I applied the negative binomial regression and compared the Log Likelihood ratio which revealed a p-value of <0.001 suggesting that the negative binomial regression is a better fit. Estimates of the regression coefficients are presented in Table 4.

Table 4: Results of the analyses of the association between the primary predictors and precision

Variable	**Estimate for change in the rate ratio for precision (95% Confidence intervals)**	**p-value**
Frequency of searching	0.98 (0.89 to 1.08)	0.661
Previous training in literature searching (referent group: No)	2.11 (0.72 to 6.21)	0.175

Assess potential confounders

Confounders were added to the model if they changed the estimate of either of the primary predictors by 10% or more. Percent changes in the estimate with the addition of the potential confounders (only one confounder at a time) are presented in Table 5. As only the inclusion of age changed the effect measures by more than 10%, it was the only confounder added into the fitted model.

Table 5: Percent change in the estimates with the addition of potential confounders

Confounding variable	**Percent change in rate ratio estimate for frequency of searching**	**Percent change in rate ratio estimate for previous training in literature searching**
Practice in academic setting (referent group: No)	0%	6%
Years practicing nephrology	0%	0%
Age	0%	16%
Sex (referent group: Females)	0%	1%

Fitted model

Results of the full model are presented in Table 6.

Table 6: Regression estimates of final model

Variable	**Estimate for rate ratio of precision (95% Confidence intervals)**	**p-value**
Frequency of searching	0.98 (0.89 to 1.08)	0.683
Previous training in literature searching (referent group: No)	2.46 (0.80 to 7.50)	0.114
Age	1.02 (0.96 to 1.08)	0.456

Diagnostics

The assessment of residuals is presented in Figure 4. The first panel (a) shows a scatter plot of the residuals while (b) presents the normal probability plot of the residuals. The model appears to fit well as only one residual value is outside the 95% confidence interval and the normal probability plot appears fitted.

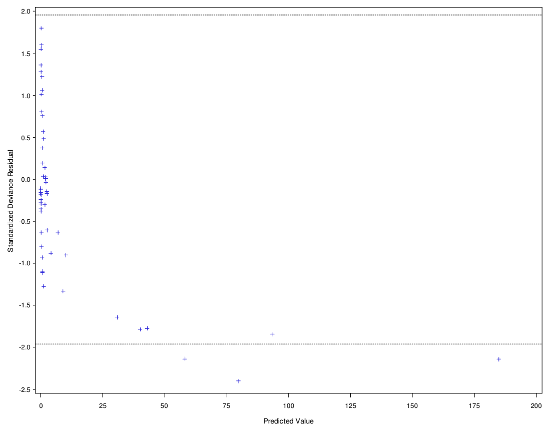


Figure 6a: Scatter plot of residuals

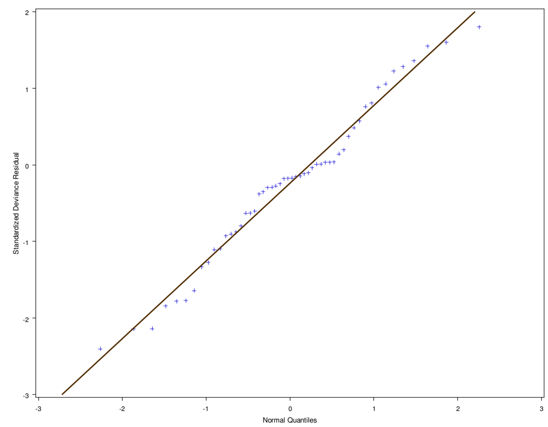


Figure 6b: Normal probability plot of residuals

A14.3 Comparing results from linear regression and negative binomial

An analysis of the residuals suggests that the negative binomial regression model fits the data much better than the linear regression model; however, no associations between physician characteristics and precision were evident in either model. The effect measures for both models were close to unity with p-values greater than 0.1. In addition, the linear regression model resulted in a negative R-squared value.

Appendix 15: Pilot data used to calculate standard deviation values

To test the feasibility of the survey, 20 clinical questions of therapy were randomly selected in July 2008. A sample of 26 physicians was approached to receive results for 20 questions, a response rate of 76% as achieved. Two nephrologists declined to participate and four did not complete the survey within 6 weeks. Respondents completed all questions in the survey.

To demonstrate the feasibility of testing the filters five physician-generated searches were selected from the set of 20 responses. Table 1(a-c) shows preliminary data for the five clinical questions. The first table (a) lists the clinical questions and the search queries provided by the nephrologists. The search query performances (sensitivity and precision) of the unaided searches are provided. The next two tables (b, c) provide the difference in search performances for each of the 17 filter combinations compared to the unaided searches. Table (b) shows changes in search sensitivity; a positive value indicates better performance for the filter. Table (c) shows changes in the precision; a positive value indicates a better performance for the filter. Twelve (12) filters improved sensitivity, 16 filters improved precision and 11 filters improved both sensitivity and precision. Combinations of all three filter-types appeared to maximally improve search performance.

Table 1: Preliminary Data - Testing the Application of Filters to Physician-generated Search Queries
(a) Five Clinical Questions & Physician-generated Search Queries

Clinician Search #	Clinical Question	Physician-provided Search Query	Sensitivity	Precision
1 (95)	What are the effects of statins on change in kidney function and urinary protein excretion?[121]	statins and kidney function	25%	1%
2 (100)	How does intradermal vs. intramuscular hepatitis B vaccine compare regarding response rate among chronic kidney disease patients?[122]	hepatitis b vaccination in chronic kidney disease	55%	5%
3 (115)	What is the impact of fenoldopam on acute kidney injury, patient mortality, and length of hospital stay in critically ill patients?[123]	fenoldopam and acute kidney injury	8%	14%
4 (72)	What is the efficacy of low-dose dopamine (<5 mcg/kg of body weight per minute) compared with no therapy in patients with or at risk for acute renal failure?[124]	low-dose dopamine AND acute renal failure	12%	8%
5 (53)	When tacrolimus is compared directly with cyclosporin, in the treatment of kidney transplant recipients, what is the evidence on transplant outcomes, toxicity and adverse effects?[125]	kidney transplant outcome tacrolimus cyclosporin	14%	5%

(b) Change in sensitivity between physician-generated search and filter aided searches
(Formula: Difference in Sensitivity = Sensitivity of filter – Sensitivity of physician-generated search)

Filter					Clinician Search #					
Journal	Methods		Content		1	2	3	4	5	Mean
	Broad	Narrow	Broad	Narrow						
x					-8%	-9%	-8%	0%	0%	-5%
	x				0%	0%	0%	0%	0%	0%
		x			0%	-9%	-8%	0%	0%	-3%
			x		54%	36%	67%	17%	78%	50%
				x	38%	36%	67%	12%	78%	46%
x	x				-8%	-9%	-8%	0%	0%	-5%
x		x			-8%	-18%	-8%	0%	0%	-7%
x			x		46%	27%	50%	17%	76%	43%
x				x	29%	27%	50%	12%	76%	39%
	x		x		54%	36%	67%	15%	78%	50%
	x			x	38%	36%	67%	12%	78%	46%
		x	x		50%	27%	58%	13%	70%	44%
		x		x	33%	27%	58%	10%	70%	40%
x	x		x		46%	27%	50%	15%	76%	43%
x	x			x	29%	27%	50%	12%	76%	39%
x		x	x		42%	18%	50%	13%	68%	38%
x		x		x	25%	18%	50%	10%	68%	34%

(c) Change in precision between physician-generated search and filter aided searches
(Formula: Difference in Precision = Precision of physician-generated search – precision of filter)

Filter					Clinician Search #					
Journal	Methods		Content		1	2	3	4	5	Mean
	Broad	Narrow	Broad	Narrow						
x					0%	1%	-14%	4%	1%	-2%
	x				1%	2%	6%	1%	1%	2%
		x			8%	31%	-14%	35%	32%	19%
			x		1%	-2%	-12%	-1%	0%	-3%
				x	1%	-1%	-10%	-1%	2%	-2%
x	x				0%	3%	-14%	5%	2%	-1%
x		x			6%	32%	0%	39%	36%	22%
x			x		1%	-1%	-11%	2%	1%	-1%
x				x	1%	0%	-9%	2%	2%	-1%
	x		x		1%	0%	-9%	0%	1%	-1%
	x			x	2%	0%	-6%	1%	3%	0%
		x	x		14%	22%	18%	27%	33%	23%
		x		X	13%	22%	24%	25%	43%	25%
x	x		x		2%	1%	-7%	3%	2%	0%
x	x			X	2%	1%	-4%	4%	3%	1%
x		x	x		13%	27%	23%	31%	35%	26%
x		x		x	12%	27%	29%	29%	45%	29%

Appendix 16: Sample-size calculations for Objective 2 – Impact of search filters on search query performance

Due to the paired nature of the analysis, sample size estimates for the objective were calculated using the SAS procedure ‘oneamplemeans’. The formula for calculation is outlined below. For all calculations, power was specified as 80% and the significance level was specified as 0.0015.

$$n_1 = \frac{\left(Z_{\alpha/2} + Z_{\beta}\right)^2 2\sigma^2}{(\Delta)^2}$$

where:

n is the total number of observations.

Δ is the minimum difference in means, between exposure groups, that one wishes to detect.

σ is the standard deviation in the population for a continuously distributed (outcome) variable.

In this analysis the physician’s unaided search will be considered ‘unexposed’ and the filter-aided search will be considered ‘exposed’. It was believed that a sample of 100 systematic reviews would meet the inclusion criteria. Using the value of 100 observations (n), the detectable mean difference in sensitivity and precision was calculated (See SAS output on the following page). With an estimated standard deviation of 0.28, the study was able to identify a difference in average sensitivity of 11.5%. The study was also able to identify a difference in average precision of 6% with an estimated standard deviation of 0.14.

SAS Output

Outcome: Sensitivity

```
                  The POWER Procedure
              One-sample t Test for Mean

               Fixed Scenario Elements

          Distribution                   Normal
          Method                          Exact
          Alpha                          0.0015
          Standard Deviation               0.28
          Total Sample Size                 100
          Power                             0.8
          Number of Sides                     2
          Null Mean                           0

                     Computed Mean

                         Mean

                        0.115
```

Outcome: Precision

```
                  The POWER Procedure
              One-sample t Test for Mean

               Fixed Scenario Elements

          Distribution                   Normal
          Method                          Exact
          Alpha                          0.0015
          Standard Deviation               0.14
          Total Sample Size                 100
          Power                             0.8
          Number of Sides                     2
          Null Mean                           0

                     Computed Mean

                          Mean

                         0.0577
```

Appendix 17: IndividualPower SAS macro to evaluate power for multiple comparisons using Dunnett's method

```
/* The %IndividualPower Macro */

/* This macro computes power for various multiple comparisons tests */
/* using the ``Individual Power" definition. */


/*---------------------------------------------------------------*/
/* Name:      IndividualPower                                    */
/* Title:     Macro to evaluate individual power of multiple     */
/*            comparisons                                        */
/* Author:    Randy Tobias, sasrdt@sas.com                       */
/* Release:   Version 7.01                                       */
/*---------------------------------------------------------------*/
/* Inputs:                                                       */
/*                                                               */
/*        MCP = RANGE, DUNNETT2, DUNNETT1, or MAXMOD (required) */
/*                                                               */
/*          G = Number of groups (excluding control for          */
/*              DUNNETT2 and DUNNETT1; required)                 */
/*                                                               */
/*          D = Meaningful mean difference (required)            */
/*                                                               */
/*          S = Standard deviation (required)                    */
/*                                                               */
/*        FWE = Desired Familywise Error (0.05 default)          */
/*                                                               */
/*     TARGET = Target power level (0.80 default)                */
/*                                                               */
/*---------------------------------------------------------------*/
/* Output:  This macro plots individual power for a variety of  */
/* Multiple comparisons methods, and plots it as a function of  */
/* n, the within-group sample size                               */
/*---------------------------------------------------------------*/

%macro IndividualPower(mcp=,g=,d=,s=,FWE=0.05,target=0.80);
%let mcp = %upcase(&mcp);
options nonotes;
data power;
   keep C_a N NCP DF Power;
   label N="Group size, N";

   ntarget = 1;
   nactual = .;
   dtarget = 1000;

   do N=2 to 1000 until (Power>.99);
      %if (&mcp = MAXMOD) %then %do; ncp = sqrt(N  )*(&d/&s); %end;
      %else                     %do; ncp = sqrt(N/2)*(&d/&s); %end;

      %if (   (&mcp = DUNNETT1)
           or (&mcp = DUNNETT2)) %then %do; df = (&g+1)*(N-1); %end;
      %else                              %do; df = (&g  )*(N-1); %end;
```

```
      conf = 1-&fwe;

      %if (&mcp = RANGE) %then %do;
         c_a = probmc("&mcp",.,conf,df,&g)/sqrt(2);
         %end;
      %else %do;
         c_a = probmc("&mcp",.,conf,df,&g);
         %end;

      %if (&mcp = DUNNETT1) %then %do;
         Power = 1-probt(c_a     ,df,ncp   );
         %end;
      %else %do;
         Power = 1-probf(c_a**2,1,df,ncp**2);
         %end;

      if (abs(Power - &target) < dtarget) then do;
         ntarget = N;
         nactual = Power;
         dtarget = abs(Power - &target);
         end;
      output;
   end;
   call symput('ntarget',trim(left(ntarget)));
   call symput('nactual',trim(left(nactual)));
run;


data target;
   length xsys ysys position $ 1;
   retain xsys ysys hsys color;
   xsys = '2'; ysys = '2'; color = 'black';
   x = 0       ; y = &nactual; function = 'MOVE ';
output;
   x = &ntarget; y = &nactual; function = 'DRAW '; line=1; size=1;
output;
   x = &ntarget; y = 0;        function = 'DRAW '; line=1; size=1;
output;
   x = &ntarget+2; y = &nactual/2; function = 'LABEL';
   style = 'swissb';
   text  = "Power(N=&ntarget)";
   position = '0';
   output;
   x = &ntarget+2; y = &nactual/2-0.12; function = 'LABEL';
   style = 'swissb';
   text  = "  = "||put(&nactual,pvalue6.);
   position = '0';
   output;

goptions ftext=swissb vsize=6 in hsize=6 in;
axis1 style=1 width=2 minor=none order=0 to 1 by 0.2;
axis2 style=1 width=2 minor=none;
symbol1 i=join;
proc gplot data=power annotate=target;
   title2 "Power for detecting an individual difference of &d";
   title3 "Using the &mcp method with FWE=&FWE";
```

```
   title4 "With &g groups and standard deviation = &s";
   plot power*n=1 / vaxis=axis1 haxis=axis2 frame;
run;
quit;
title2;
title3;
title4;
options notes;
%mend;
```

Appendix 18: Sample-size calculations for Objective 2 using Dunnett's method

Sample size estimates were calculated using the SAS macro '%IndividualPower' with the following parameters: standard deviation=0.28; groups=34; mean detectable difference=0.115. These parameters were identified from the sample size calculations used for a paired t-test with an alpha value of 0.0015, after applying the Bonferroni correction (Appendix 17).

Sample Size using Dunnett's method

Given the above parameters, for the primary outcome of sensitivity, using the Dunnett's method, this study would require search queries for 179 clinical questions and have power of ~80% to detect a difference of 11.5% in the primary outcome of sensitivity.

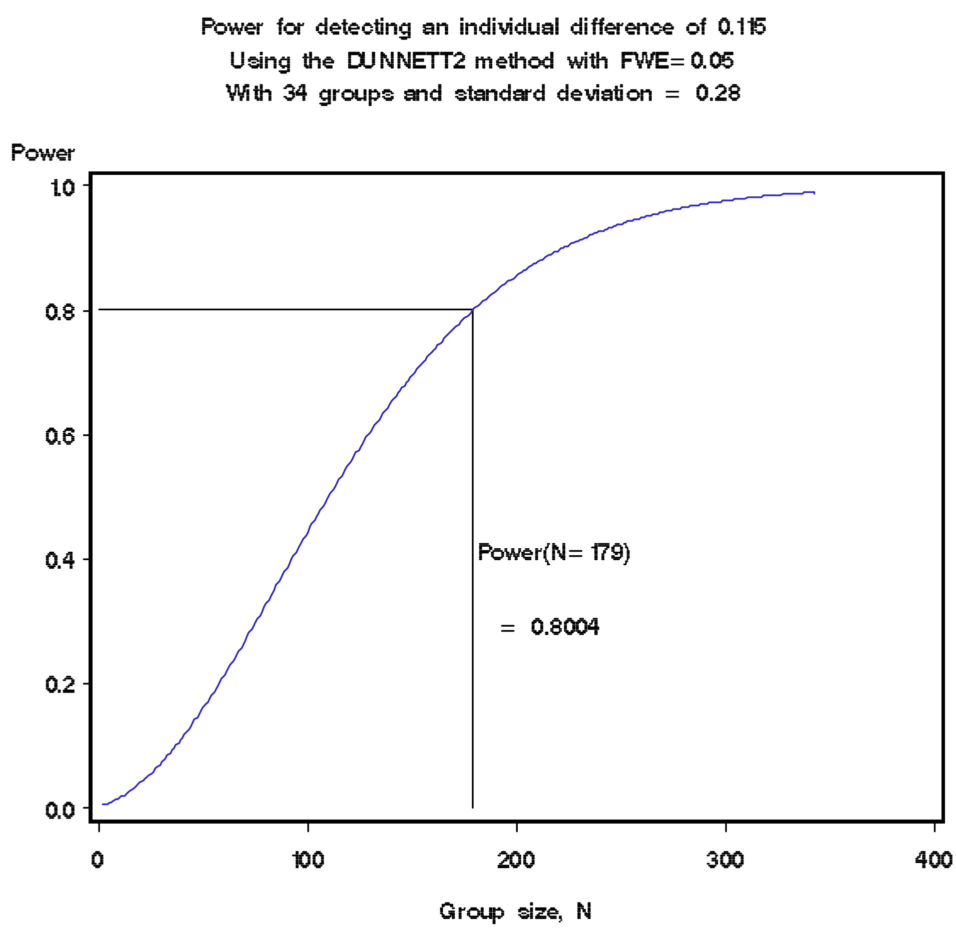
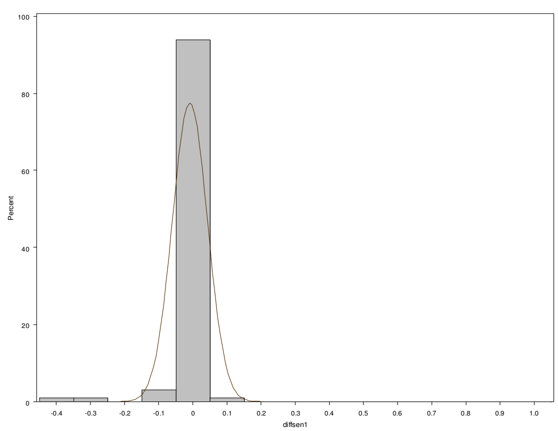


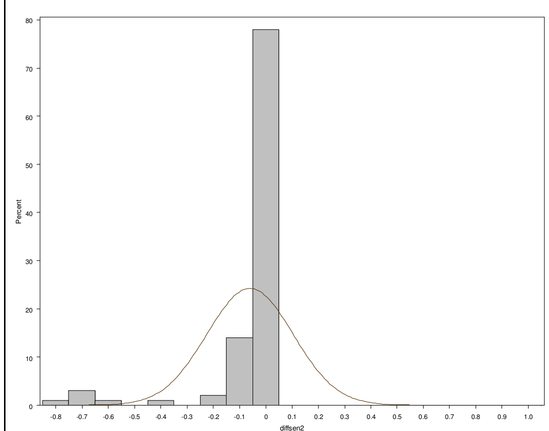
Figure 1: Power calculation in SAS using Dunnett's method

Appendix 19: Objective 2 - Histograms of the differences in search sensitivity and precision between filtered and non-filtered searches

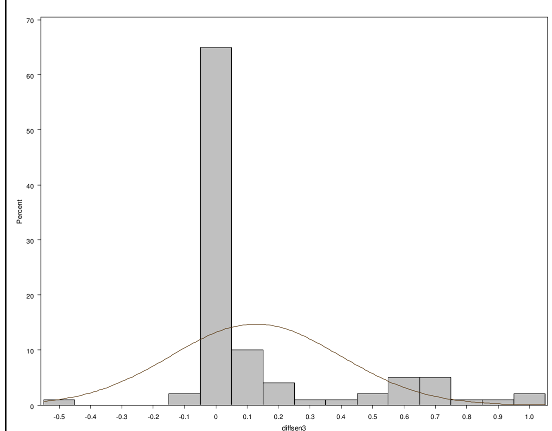
Figure 1 (a-q): Histograms of the differences in sensitivity between filtered and non-filtered searches



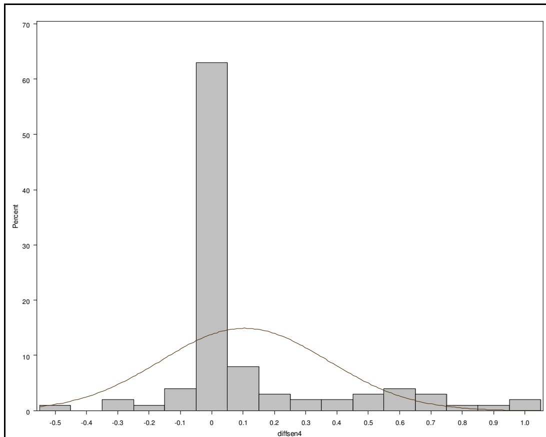
Panel a: Difference between non-filtered searches and the searches aided with the methods broad filter



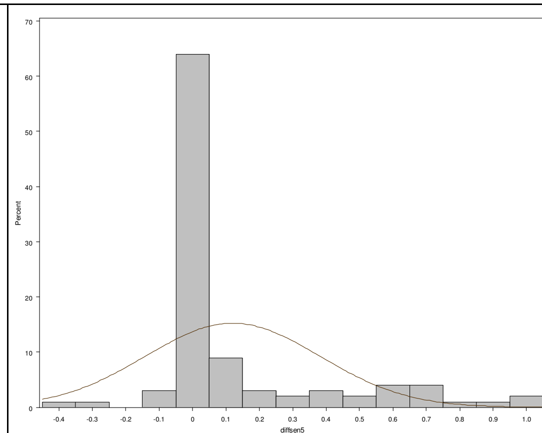
Panel b: Difference between non-filtered searches and the searches aided with the methods narrow filter



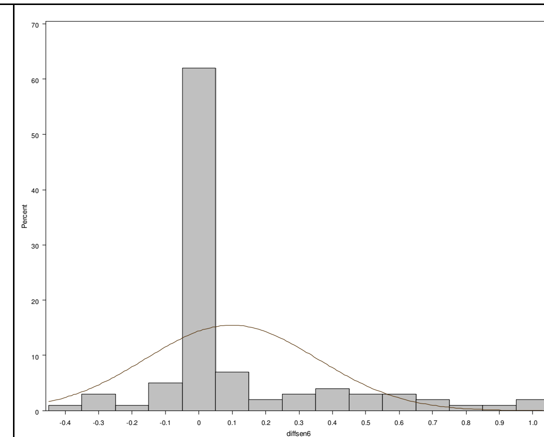
Panel c: Difference between non-filtered searches and the searches aided with the content broad filter



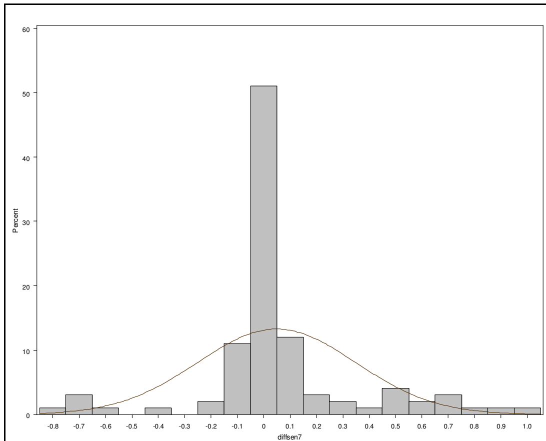
Panel d: Difference between non-filtered searches and the searches aided with the content narrow filter



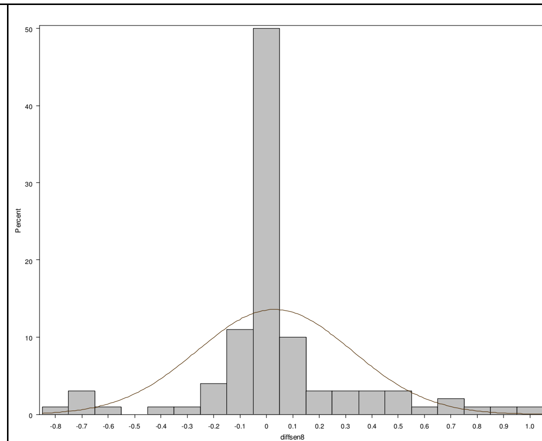
Panel e: Difference between non-filtered searches and the searches aided with the methods broad filter and contents broad filter



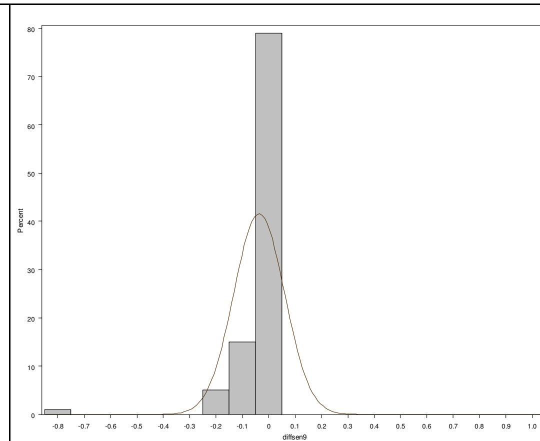
Panel f: Difference between non-filtered searches and the searches aided with the methods broad filter and contents narrow filter



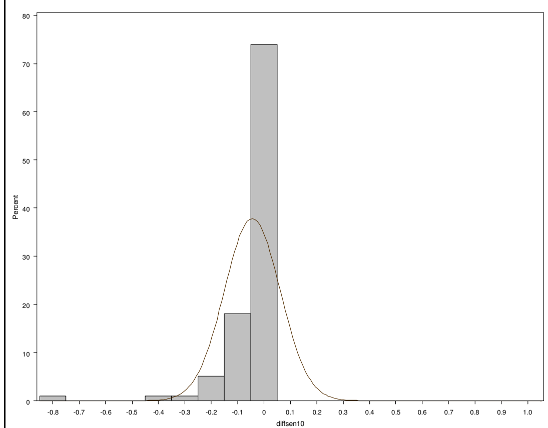
Panel g: Difference between non-filtered searches and the searches aided with the methods narrow filter and contents broad filter



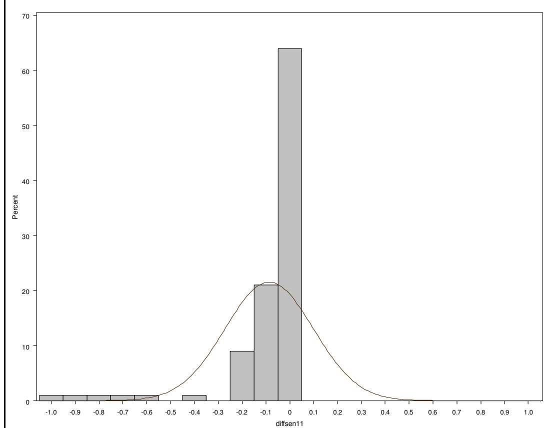
Panel h: Difference between non-filtered searches and the searches aided with the methods narrow filter and contents narrow filter



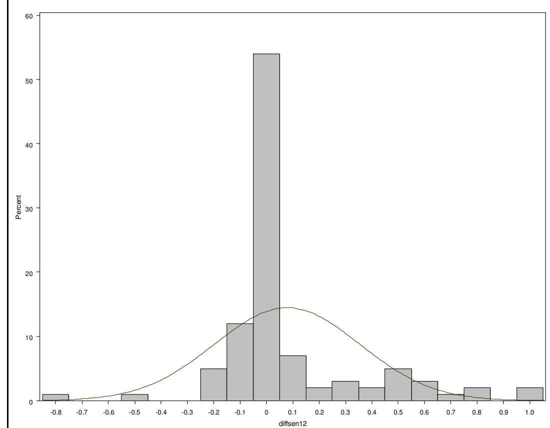
Panel i: Difference between non-filtered searches and the searches aided with the journal filter



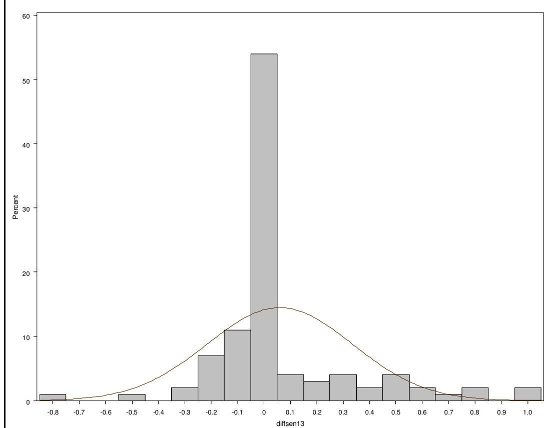
Panel j: Difference between non-filtered searches and the searches aided with the journal and methods broad filter



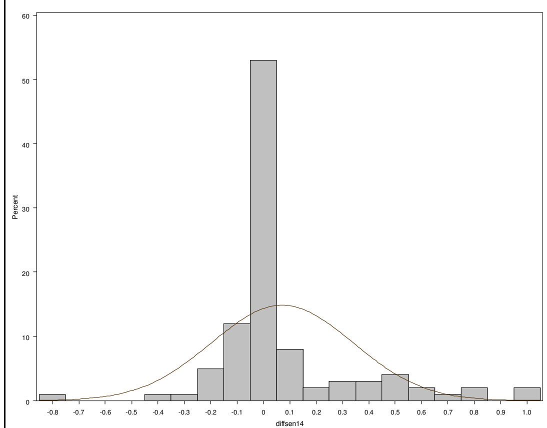
Panel k: Difference between non-filtered searches and the searches aided with the journal and methods narrow filter



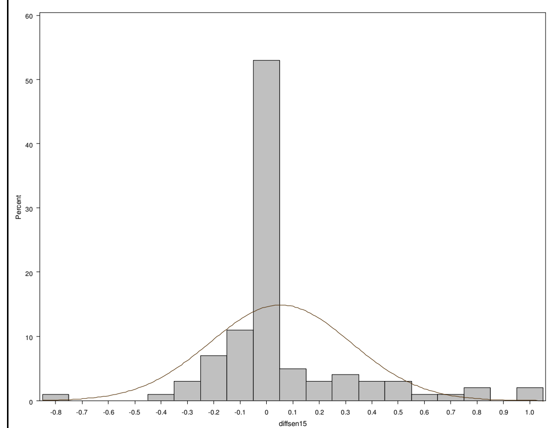
Panel l: Difference between non-filtered searches and the searches aided with the journal and content broad filter



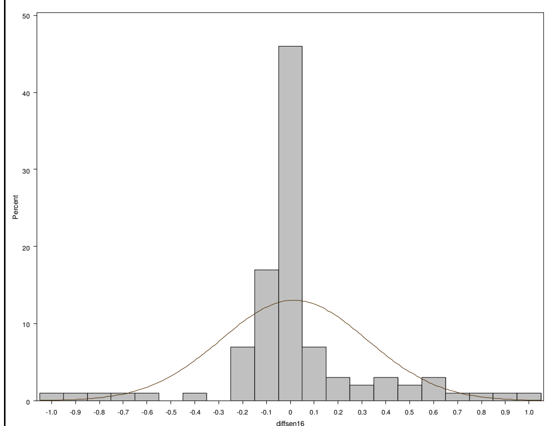
Panel m: Difference between non-filtered searches and the searches aided with the journal and contents narrow filter



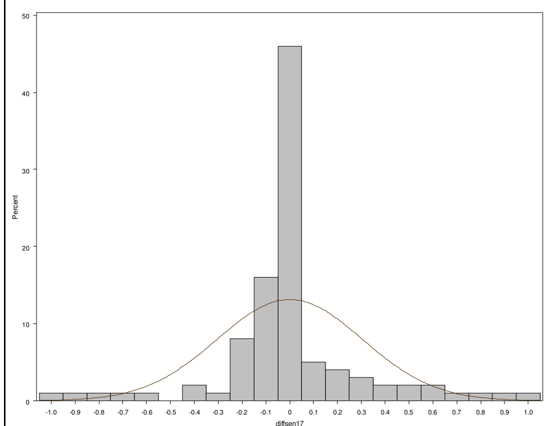
Panel n: Difference between non-filtered searches and the searches aided with the journal, methods broad filter and contents broad filter



Panel o: Difference between non-filtered searches and the searches aided with the journal, methods broad filter and contents narrow filter

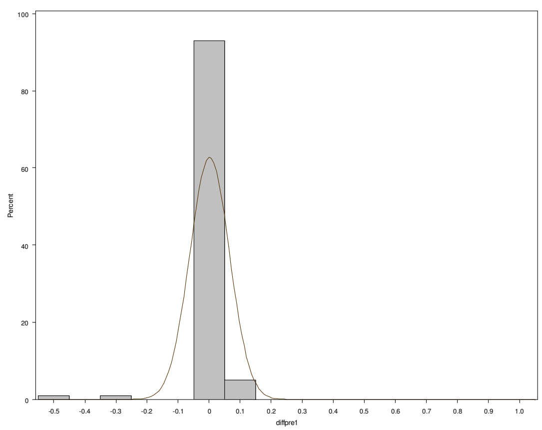


Panel p: Difference between non-filtered searches and the searches aided with the journal, methods narrow filter and contents broad filter

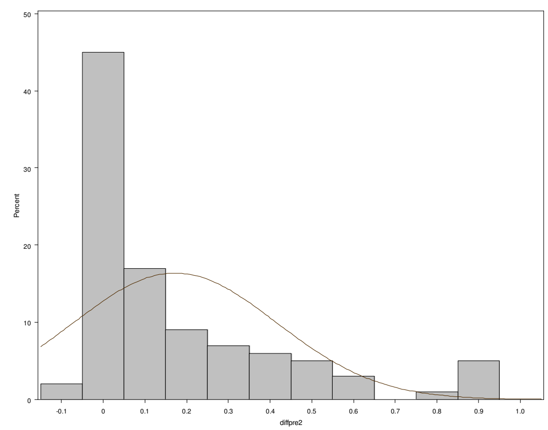


Panel q: Difference between non-filtered searches and the searches aided with the journal, methods narrow filter and contents narrow filter

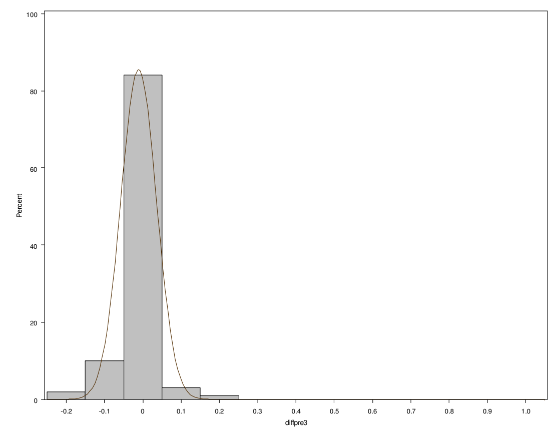
Figure 2 (a-q): Histograms of the differences in precision between filtered and non-filtered searches



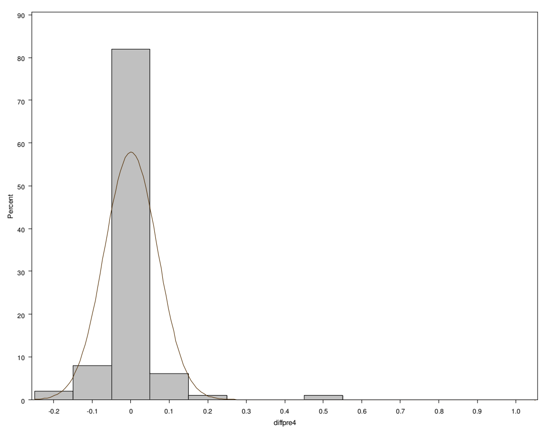
Panel a: Difference between non-filtered searches and the searches aided with the methods broad filter



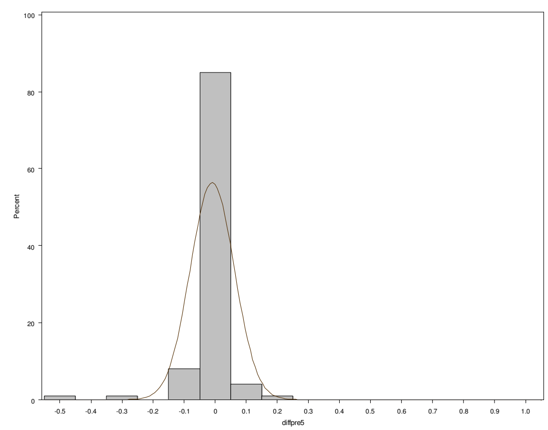
Panel b: Difference between non-filtered searches and the searches aided with the methods narrow filter



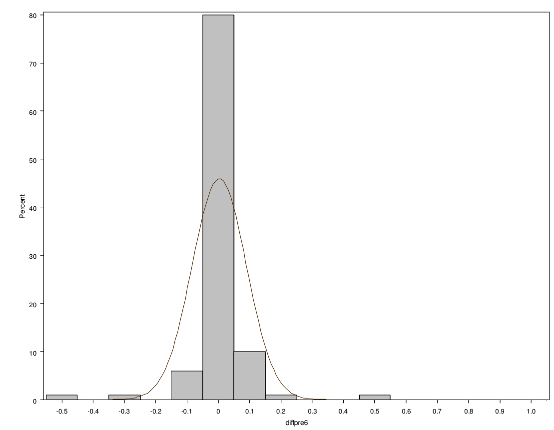
Panel c: Difference between non-filtered searches and the searches aided with the content broad filter



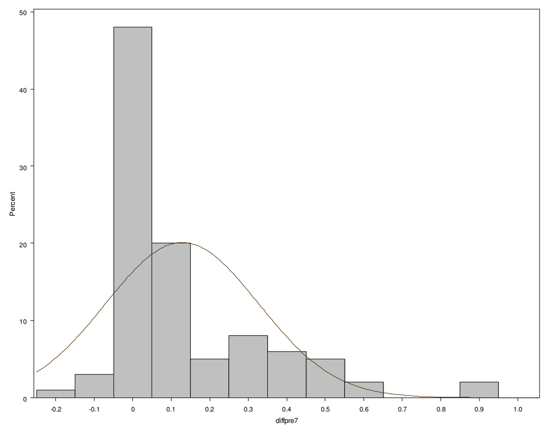
Panel d: Difference between non-filtered searches and the searches aided with the content narrow filter



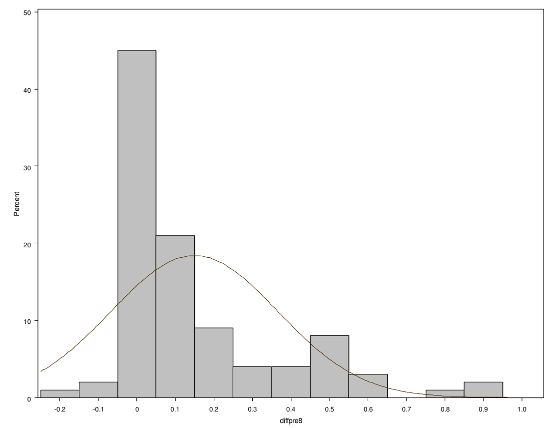
Panel e: Difference between non-filtered searches and the searches aided with the methods broad filter and contents broad filter



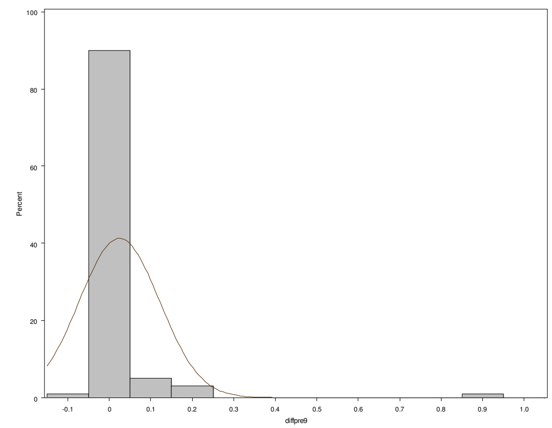
Panel f: Difference between non-filtered searches and the searches aided with the methods broad filter and contents narrow filter



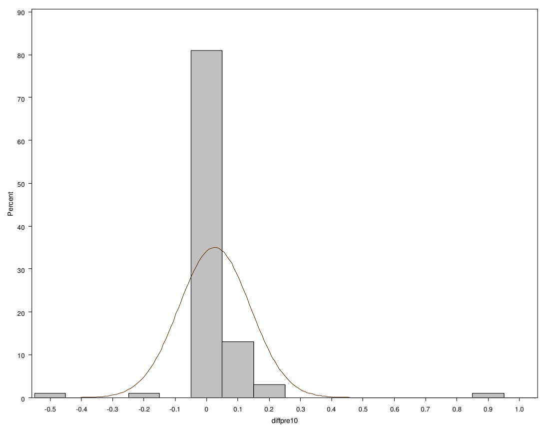
Panel g: Difference between non-filtered searches and the searches aided with the methods narrow filter and contents broad filter



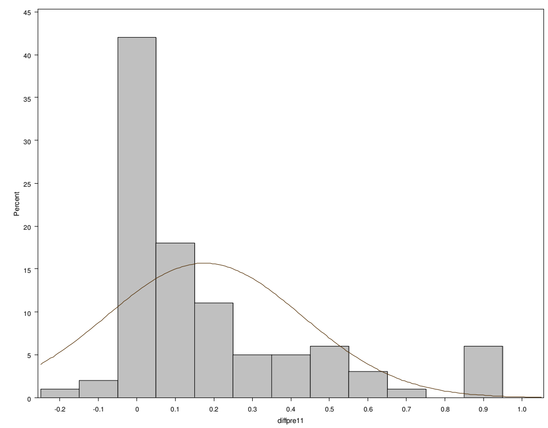
Panel h: Difference between non-filtered searches and the searches aided with the methods narrow filter and contents narrow filter



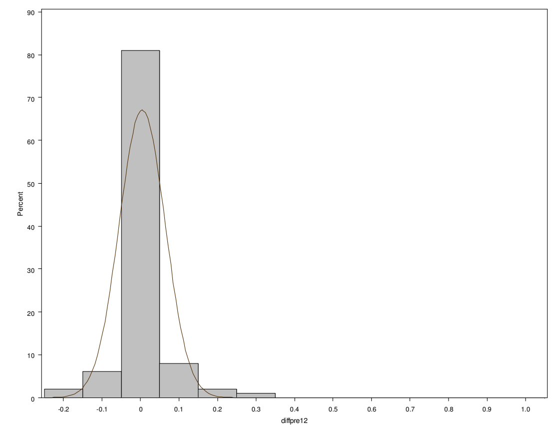
Panel i: Difference between non-filtered searches and the searches aided with the journal filter



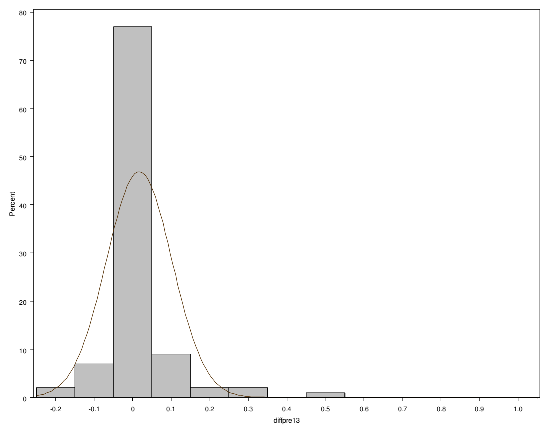
Panel j: Difference between non-filtered searches and the searches aided with the journal and methods broad filter



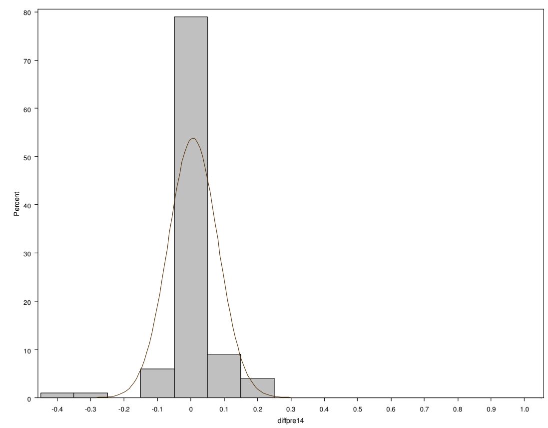
Panel k: Difference between non-filtered searches and the searches aided with the journal and methods narrow filter



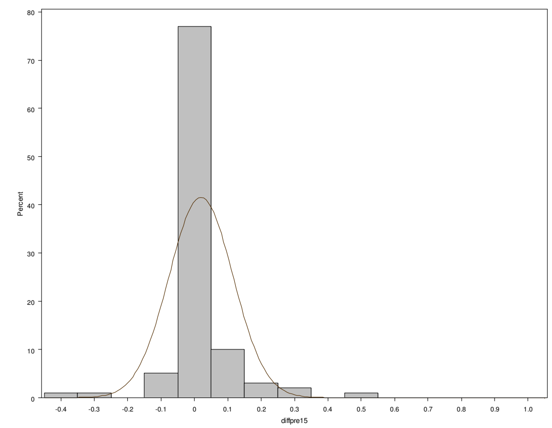
Panel l: Difference between non-filtered searches and the searches aided with the journal and content broad filter



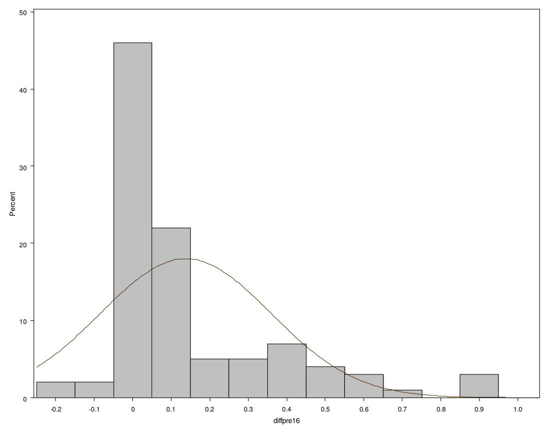
Panel m: Difference between non-filtered searches and the searches aided with the journal and contents narrow filter



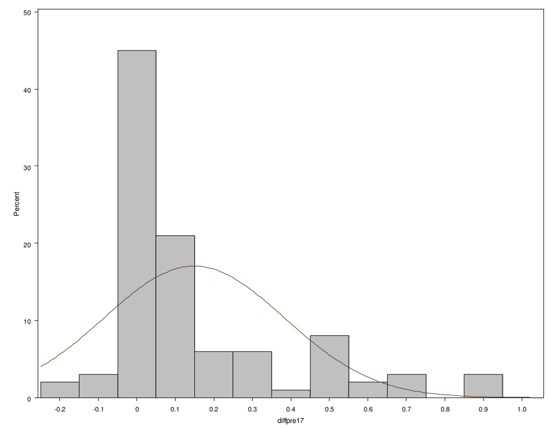
Panel n: Difference between non-filtered searches and the searches aided with the journal, methods broad filter and contents broad filter



Panel o: Difference between non-filtered searches and the searches aided with the journal, methods broad filter and contents narrow filter



Panel p: Difference between non-filtered searches and the searches aided with the journal, methods narrow filter and contents broad filter



Panel q: Difference between non-filtered searches and the searches aided with the journal, methods narrow filter and contents narrow filter

Appendix 20: SAS code used for Objective 2

```
%macro printOutUnivariateSummary(datain, varnm, filename);
    %do i=1 %to 18;
        proc univariate data=&datain noprint;
            var &varnm&i;
            output out=temp mean =mean median=median  min=min max=max;
        run;

        %if &i=1 %then %do;
            data out;
                set temp;
            run;
        %end;

        %if &i>1 %then %do;
            data out;
                set out temp;
            run;
        %end;
    %end;

    PROC EXPORT DATA=OUT
            OUTFILE= "P:\My Documents\SALIMAH\PROJECTS\Survey -
Filters Results\June
2010\Analyses\Secondary\Original\AllRanks\&filename"
            DBMS=tab REPLACE;
    RUN;
%mend;

%macro printOutUnivariateStatistics(datain, varnm, filename);
      %do i=1 %to 17;
            proc univariate data=&datain noprint;
                  var &varnm&i;
                  output out=temp mean=mean msign=signStatistic
probm=singPvalue probs=signedRankPvalue signrank=signedRankStatistic
probt=ttestPvalue
                  t=ttestStatistic;
            run;

            %if &i=1 %then %do;
                  data out;
                        set temp;
                  run;
            %end;

            %if &i>1 %then %do;
                  data out;
                        set out temp;
                  run;
            %end;
      %end; /*end do*/

       PROC EXPORT DATA=OUT
```

```
            OUTFILE= "P:\My Documents\SALIMAH\PROJECTS\Survey -
Filters Results\June
2010\Analyses\Secondary\Original\AllRanks\&filename"
            DBMS=tab REPLACE;
    RUN;
%mend;

proc import
      datafile = "P:\My Documents\SALIMAH\PROJECTS\Survey - Filters
Results\June
2010\Analyses\Secondary\Original\AllRanks\allSRsOriginalSummary.xls"
      out=fullSet dbms =excel replace;
run;

/*BEGIN Total Number of Included Found*/
%printOutUnivariateSummary(datain=fullSet, varnm=Search,
filename=outNumIncludedStudiesFound.txt);
/*END Total Number of Included Found*/

/*BEGIN Total Number of Citations Found: mean, median and rage*/
%printOutUnivariateSummary(datain=fullSet, varnm=SearchTotal,
filename=outNumTotalStudiesFound.txt);
/*END Total Number of Citations Found*/

/*Calculate the Sensitivity for all 18 Searches (Doc + 17 filters)*/
data allSen(keep=sen:);set fullSet;
  array Search{18} Search1 - Search18;
  array sen{18} sen1 - sen18;
  if _includedstudies>1 then
      do i=1 to 18;
            sen{i}=Search{i}/_includedstudies;
      end;
run;

/*BEGIN Sensitivity: mean, median, range*/
%printOutUnivariateSummary(datain=allSen, varnm=sen,
filename=outSen.txt);
/*END Sensitivity*/

/*Calculate the difference in Sensitivity between the 17 Filters and
the Doc search*/
data diffSen( keep=diffsen:);set allSen;
 array sen {18} sen1 - sen18;
 array diffsen {17} diffsen1-diffsen17;
   do i=1 to 17;
      diffsen{i}=sen{i+1}-sen1;
   end;
run;

/*Draw Histograms of the differences in Sensitivity to test for
normality*/
proc univariate data=diffSen noprint;
histogram diffsen1-diffsen17 / cfill=ltgray
                        midpoints=0 to 1 by 0.1
                        normal;
run;
```

```
/*get ttest results for the Difference in Sensitivity with an alpha of
0.0015*/
/*ods trace on;*/
/*BEGIN Diff Sen Confidnece Intervals*/
ods output BasicIntervals=outdiffSen BasicMeasures=outdiffSenRest;
proc univariate data=diffSen cibasic (alpha=.0015) cipctldf (TYPE =
ASYMMETRIC alpha=0.01);
      var diffsen1-diffsen17;
run;

PROC EXPORT DATA= WORK.OUTDIFFSEN
            OUTFILE= "P:\My Documents\SALIMAH\PROJECTS\Survey -
Filters Results\June
2010\Analyses\Secondary\Original\AllRanks\outDiffSenCI.txt"
            DBMS=tab REPLACE;
RUN;
/*ods trace off;*/
/*END Diff Sen Confidnece Intervals*/

/*BEGIN Diff Sen Rest*/
%printOutUnivariateStatistics(datain=diffSen, varnm=diffsen,
filename=outDiffSenRest.txt);
/*END Diff Sen Rest*/


/*calculate precision*/
data allPre(keep=pre:);set fullSet;
  array search{18} search1 - search18;
  array searchtotal {18} searchtotal1 - searchtotal18;
  array pre{18} pre1 - pre18;
  if _includedstudies>1 then
      do i=1 to 18;
            if searchtotal(i)=0 then searchtotal(i)=1;
      pre{i}=search{i}/searchtotal{i};
      end;
run;

/*BEGIN Precision: mean, median, range*/
%printOutUnivariateSummary(datain=allPre, varnm=pre,
filename=outPre.txt);
/*END Precision*/

/*calculate difference in precision*/
data diffPre( keep=diffpre:);set allPre;
 array pre {18} pre1 - pre18;
 array diffpre {17} diffpre1-diffpre17;
   do i=1 to 17;
      diffpre{i}=(pre{i+1}-pre1);
   end;
run;

/*Draw Histograms of the differences in Sensitivity to test for
normality*/
proc univariate data=diffPre noprint;
histogram diffpre1-diffpre17 / cfill=ltgray
                        midpoints=0 to 1 by 0.1
                        normal;
```

```
run;

/*get ttest results for the Difference in Precision with an alpha of
0.0015*/
/*ods trace on;*/
/*BEGIN Diff Pre Confidnece Intervals*/
ods output BasicIntervals=outdiffPre BasicMeasures=outdiffPreRest;
proc univariate data=diffPre cibasic (alpha=.0015) cipctldf (TYPE =
ASYMMETRIC alpha=0.01);
      var diffpre1-diffpre17;
run;
PROC EXPORT DATA= WORK.OUTDIFFPRE
            OUTFILE= "P:\My Documents\SALIMAH\PROJECTS\Survey -
Filters Results\June
2010\Analyses\Secondary\Original\AllRanks\outDiffPreCI.txt"
            DBMS=tab REPLACE;
RUN;
/*ods trace off;*/
/*END Diff Sen Confidnece Intervals*/

/*BEGIN Diff Pre Rest*/
/*%printOutUnivariateStatistics(datain=diffpre, dataout=outdiffpre,
varnm=diffpre, filename=outDiffPreRest.txt);*/
%printOutUnivariateStatistics(datain=diffpre, varnm=diffpre,
filename=outDiffPreRest.txt);
/*END Diff Pre rest*/
```

Appendix 21: Objective 2 - Details of search results (all results)

Methods		Content		Journal	Number of Included Studies Found			Number of Total Citations Found			Sensitivity (S) and Precision (P)				
Broad	*Narrow*	*Broad*	*Narrow*		*mean*	*median*	*max*	*mean*	*median*	*max*	*measure*	*mean*	*median*	*min*	*max*
Physician-provided search (unaided)					5	3	61	37270	113	3389033	P	5.3%	1.1%	0.0%	66.7%
											S	37.5%	25.0%	0.0%	100.0%
x					5	3	61	16279	92	1461893	P	5.5%	1.5%	0.0%	50.0%
											S	36.7%	25.0%	0.0%	100.0%
	x				4	2	55	1431	11	125523	P	22.5%	8.8%	0.0%	100.0%
											S	31.5%	18.6%	0.0%	100.0%
		x			7	6	53	21269	339	693068	P	4.2%	0.8%	0.0%	77.8%
											S	50.2%	53.6%	0.0%	100.0%
			x		7	4	43	12157	260	369942	P	5.4%	1.0%	0.0%	100.0%
											S	48.0%	48.5%	0.0%	100.0%
x		x			7	6	53	8940	283	203678	P	4.4%	1.2%	0.0%	50.0%
											S	49.5%	50.0%	0.0%	100.0%
x			x		7	5	43	5297	228	116794	P	5.6%	1.7%	0.0%	100.0%
											S	47.3%	45.8%	0.0%	100.0%
	x	x			6	4	48	621	38	12047	P	18.1%	6.4%	0.0%	100.0%
											S	42.4%	33.3%	0.0%	100.0%
	x		x		6	4	38	380	27	7116	P	20.3%	8.5%	0.0%	100.0%
											S	40.7%	33.3%	0.0%	100.0%

Methods		Content		Journal	Number of Included Studies Found			Number of Total Citations Found			Sensitivity (S) and Precision (P)				
Broad	***Narrow***	***Broad***	***Narrow***		***mean***	***median***	***max***	***mean***	***median***	***max***	***measure***	***mean***	***median***	***min***	***max***
				x	5	2	57	6901	63	531959	P	7.7%	1.3%	0.0%	100.0%
											S	34.0%	20.0%	0.0%	100.0%
x				x	5	2	57	4502	55	364660	P	7.9%	1.8%	0.0%	100.0%
											S	33.1%	20.0%	0.0%	100.0%
	x			x	4	2	53	613	9	49282	P	22.8%	8.1%	0.0%	100.0%
											S	28.9%	14.3%	0.0%	100.0%
		x		x	7	4	49	8949	181	241923	P	5.8%	1.2%	0.0%	77.8%
											S	45.4%	40.8%	0.0%	100.0%
			x	x	6	4	39	6269	171	164658	P	7.0%	1.6%	0.0%	100.0%
											S	43.5%	33.3%	0.0%	100.0%
x		x		x	7	5	49	4494	167	94711	P	6.0%	1.7%	0.0%	50.0%
											S	44.7%	36.7%	0.0%	100.0%
x			x	x	6	4	39	3211	143	66935	P	7.3%	2.3%	0.0%	100.0%
											S	42.8%	33.3%	0.0%	100.0%
	x	x		x	6	3	46	444	30	8214	P	19.1%	6.6%	0.0%	100.0%
											S	39.0%	29.3%	0.0%	100.0%
	x		x	x	5	3	36	308	24	5571	P	20.2%	7.6%	0.0%	100.0%
											S	37.5%	28.6%	0.0%	100.0%

Appendix 22: Objective 2 - Details of search results (when restricting the results to the top 40 citations)

Methods		Content		Journal	Number of Included Studies Found			Number of Total Citations Found up to 40			Sensitivity (S) and Precision (P)				
Broad	***Narrow***	***Broad***	***Narrow***		***mean***	***median***	***max***	***mean***	***median***	***max***	***measure***	***mean***	***median***	***min***	***max***
Physician-provided search (unaided)					1	0	9	30	40	40	P	4.6%	0.0%	0.0%	66.7%
											S	9.3%	0.0%	0.0%	100.0%
x					1	0	10	28	40	40	P	5.1%	0.0%	0.0%	50.0%
											S	10.3%	0.0%	0.0%	100.0%
	x				3	2	22	17	11	40	P	22.9%	12.5%	0.0%	100.0%
											S	22.7%	13.7%	0.0%	100.0%
		x			1	0	8	34	40	40	P	3.9%	0.0%	0.0%	77.8%
											S	8.3%	0.0%	0.0%	87.5%
			x		1	0	10	33	40	40	P	5.1%	0.0%	0.0%	100.0%
											S	9.8%	0.0%	0.0%	100.0%
x		x			1	0	10	34	40	40	P	4.0%	0.0%	0.0%	50.0%
											S	9.7%	0.0%	0.0%	100.0%
x			x		1	0	10	32	40	40	P	5.3%	0.0%	0.0%	100.0%
											S	10.9%	0.0%	0.0%	100.0%
	x	x			3	2	21	25	38	40	P	18.4%	6.3%	0.0%	100.0%
											S	23.4%	11.6%	0.0%	100.0%
	x		x		3	2	23	23	27	40	P	20.5%	7.5%	0.0%	100.0%
											S	25.1%	14.3%	0.0%	100.0%

Methods		Content		Journal	Number of Included Studies Found			Number of Total Citations Found up to 40			Sensitivity (S) and Precision (P)				
Broad	***Narrow***	***Broad***	***Narrow***		***mean***	***median***	***max***	***mean***	***median***	***max***	***measure***	***mean***	***median***	***min***	***max***
				x	1	0	9	27	40	40	P	7.0%	0.0%	0.0%	100.0%
											S	9.8%	0.0%	0.0%	100.0%
x				x	1	0	9	26	40	40	P	7.7%	0.0%	0.0%	100.0%
											S	11.1%	0.0%	0.0%	100.0%
	x			x	3	1	21	16	9	40	P	23.2%	10.0%	0.0%	100.0%
											S	21.6%	11.1%	0.0%	100.0%
		x		x	1	0	9	32	40	40	P	5.6%	0.0%	0.0%	77.8%
											S	10.6%	0.0%	0.0%	100.0%
			x	x	1	0	9	31	40	40	P	6.8%	0.0%	0.0%	100.0%
											S	11.0%	0.0%	0.0%	100.0%
x		x		x	1	0	9	31	40	40	P	5.8%	0.0%	0.0%	50.0%
											S	11.5%	0.0%	0.0%	100.0%
x			x	x	1	1	9	30	40	40	P	7.3%	2.5%	0.0%	100.0%
											S	12.6%	3.8%	0.0%	100.0%
	x	x		x	3	1	21	23	30	40	P	19.3%	7.5%	0.0%	100.0%
											S	22.9%	11.6%	0.0%	100.0%
	x		x	x	3	1	21	22	24	40	P	20.5%	7.6%	0.0%	100.0%
											S	23.7%	12.3%	0.0%	100.0%

Appendix 23: Objective 2 - Details of additional analyses (using modified search queries)

Analyzing all returned citations

The mean and median sensitivity and precision values of the 18 different searches are presented in Table 1. Descriptively, physician-provided search queries exhibited a median sensitivity of 42% (half the search queries retrieved over 42% of the relevant articles) and a median precision of 2% (1 in 50 articles retrieved by the searches were considered relevant). After applying the filters, median sensitivity ranged from 25% to 58% and median precision ranged from 1% to 10%.

Table 2 presents the mean and median differences in sensitivity and precision between the physician-provided searches and the filter aided searches. When considering the filters alone, sensitivity was most improved after applying the renal ‘content’ broad filter, while precision significantly decreased. Precision was most improved after applying the ‘methods’ narrow filter (median difference 10%, 99% CI: 4% to 15%), while sensitivity significantly decreased. The combination of ‘methods’ narrow filter and the ‘content’ narrow filter produced the best improvement in search performance; a 7% median improvement in precision (99% CI: 3% to 14%) while sensitivity remained unchanged. Expressing this improvement in precision another way, the ratio of relevant to non-relevant articles went from 1 in 50 with the unfiltered search to 1 in 10 when both filters were used in combination. No filters produced significant simultaneous improvements in both sensitivity and precision. The addition of the journal filter did not produce noteworthy improvements over the methods and content filters.

Table 1: Search performance of modified physician-provided searches and searches aided by filters

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
Physician-provided search (unaided)					P	6.0%	1.5%
					S	45.9%	42.3%
x					P	6.1%	1.9%
					S	45.2%	41.4%
	x				P	22.8%	11.5%
					S	37.9%	27.6%
		x			P	4.8%	1.0%
					S	54.5%	58.3%
			x		P	6.0%	1.5%
					S	52.4%	53.6%
x		x			P	4.8%	1.6%
					S	54.0%	57.7%
x			x		P	6.1%	2.0%
					S	51.9%	52.7%
	x	x			P	19.4%	8.1%
					S	45.9%	42.9%
	x		x		P	21.6%	10.0%
					S	44.2%	40.8%

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
				x	P	8.6%	2.0%
					S	41.4%	33.3%
x				x	P	8.7%	2.6%
					S	40.6%	31.4%
	x			x	P	23.2%	9.8%
					S	35.2%	25.0%
		x		x	P	7.2%	1.6%
					S	49.3%	50.0%
			x	x	P	8.5%	2.1%
					S	47.5%	48.5%
x		x		x	P	7.4%	1.9%
					S	48.7%	48.5%
x			x	x	P	8.7%	2.4%
					S	47.0%	46.6%
	x	x		x	P	20.5%	8.2%
					S	42.5%	33.3%
	x		x	x	P	21.6%	9.3%
					S	41.1%	33.3%

Table 2: Change in search performance between filtered and non-filtered modified physician-provided searches

Methods Filter		**Content Filter**		**Journal Filter**	**Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)**					
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean difference***	***% of queries improvement seen***	***Median difference (99% CI)***	***p-value Wilcoxon***	***p-value Sign test***
x					P	0.1%	75	0.31 (0.14 to 0.55)	**<0.0001**	**<0.0001**
					S	-0.7%	2	0.00 (0.00 to 0.00)	0.2031	0.1797
	x				P	16.9%	73	9.81 (3.72 to 15.15)	**<0.0001**	**<0.0001**
					S	-8.0%	1	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
		x			P	-1.1%	25	-0.05 (-0.29 to 0.00)	**0.0001**	**0.0005**
					S	8.6%	30	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
			x		P	0.1%	46	0.00 (-0.02 to 0.04)	0.9325	0.5154
					S	6.5%	26	0.00 (0.00 to 0.00)	0.0030	0.0113
x		x			P	-1.1%	42	0.00(-0.09 to 0.03)	0.4608	0.9142
					S	8.0%	28	0.00 (0.00 to 0.00)	**0.0005**	0.0095
x			x		P	0.2%	61	0.19 (0.00 to 0.56)	0.0132	0.0002
					S	5.9%	24	0.00 (0.00 to 0.00)	0.0219	0.2682
	x	x			P	13.5%	76	5.96 (2.80 to 12.31)	**<0.0001**	**<0.0001**
					S	-0.1%	24	0.00 (0.00 to 0.00)	0.7094	0.2892
	x		x		P	15.6%	76	7.40 (3.35 to 14.17)	**<0.0001**	**<0.0001**
					S	-1.7%	22	0.00 (0.00 to 0.00)	0.3127	0.0869

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)					
Broad	*Narrow*	*Broad*	*Narrow*		*Measure*	*Mean difference*	*% of queries improvement seen*	*Median difference (99% CI)*	*P-value Wilcoxon*	*p-value Sign test*
				x	P	2.6%	74	0.48 (0.17 to 0.86)	**<0.0001**	**<0.0001**
					S	-4.6%	0	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
x				x	P	2.7%	76	0.83 (0.33 to 1.75)	**<0.0001**	**<0.0001**
					S	-5.3%	1	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
	x			x	P	17.3%	72	7.97 (3.95 to 16.48)	**<0.0001**	**<0.0001**
					S	-10.8%	1	-2.66 (-8.33 to 0.00)	**<0.0001**	**<0.0001**
		x		x	P	1.3%	55	0.01 (0.00 to 0.37)	0.0390	0.0178
					S	3.3%	25	0.00 (0.00 to 0.00)	0.5022	1.0000
			x	x	P	2.5%	63	0.18 (0.00 to 0.85)	**0.0002**	**<0.0001**
					S	1.6%	22	0.00 (0.00 to 0.00)	0.9374	0.4011
x		x		x	P	1.4%	63	0.29 (0.01 to 0.86)	**0.0009**	**0.0001**
					S	2.8%	24	0.00 (0.00 to 0.00)	0.6686	0.5831
x			x	x	P	2.7%	67	0.49 (0.06 to 1.52)	**<0.0001**	**<0.0001**
					S	1.0%	21	0.00 (0.00 to 0.00)	0.7599	0.1690
	x	x		x	P	14.5%	75	6.32 (2.91 to 12.18)	**<0.0001**	**<0.0001**
					S	-3.5%	22	0.00 (-5.56 to 0.00)	0.0939	0.0125
	x		x	x	P	15.6%	74	5.98 (3.33 to 14.42)	**<0.0001**	**<0.0001**
					S	-4.8%	20	0.00 (-6.67 to 0.00)	0.0352	0.0026

<u>Analyzing the top 40 returned citations</u>

The mean and median sensitivity and precision values of the 18 different searches, when restricting the results to the top 40 citations are presented in Table 3. Descriptively, physician provided search queries exhibited a median sensitivity of 0% (half the search queries retrieved none of the relevant articles within the first 40 citations) and thus, a median precision of 0%. After applying the filters, median sensitivity ranged from 0% to 17% and median precision ranged from 0% to 13%.

Table 4 presents the mean and median differences in sensitivity and precision between the physician-provided searches and the filter aided searches when restricted to the top 40 returned citations. When considering the filters alone, sensitivity and precision were maximally improved after applying the ‘methods’ narrow filter (sensitivity median difference: 0%, 99% CI: 0% to 14%; precision median difference 10%, 99% CI: 3% to 17%). The combination of ‘methods’ narrow filter, the ‘content’ narrow filter and ‘journal’ filter produced the best combined improvement; 8% median improvement in precision (99% CI: 3% to 15%) while sensitivity also significantly improved (median difference 2% 99% CI: 0 to 16%). This was comparable to the combined effect of the ‘methods’ narrow and ‘content’ narrow filters which resulted in a 9% median improvement in precision (99% CI: 3% to 15%) and 0% median improvement in precision (99%: 0% to 17%).

Table 3: Search performance of modified physician-provided searches and searches aided by filters, when restricted to the top 40 returned results

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
Physician-provided search (unaided)					P	5.5%	0.0%
					S	12.7%	0.0%
x					P	5.7%	2.5%
					S	13.3%	1.8%
	x				P	23.1%	12.5%
					S	26.1%	14.6%
		x			P	4.5%	0.0%
					S	10.0%	0.0%
			x		P	5.7%	0.0%
					S	11.4%	0.0%
x		x			P	4.5%	0.0%
					S	11.3%	0.0%
x			x		P	5.7%	1.3%
					S	12.5%	0.7%
	x	x			P	19.7%	8.8%
					S	25.7%	13.8%
	x		x		P	21.7%	11.3%
					S	26.8%	16.7%

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Search performance (P=precision; S=sensitivity)		
Broad	***Narrow***	***Broad***	***Narrow***		***Measure***	***Mean***	***Median***
				x	P	8.1%	1.3%
					S	13.2%	0.7%
x				x	P	8.3%	2.5%
					S	14.0%	5.8%
	x			x	P	23.6%	12.5%
					S	25.5%	14.3%
		x		x	P	7.0%	0.0%
					S	12.4%	0.0%
			x	x	P	8.2%	2.5%
					S	13.0%	2.5%
x		x		x	P	7.1%	2.5%
					S	13.2%	4.4%
x			x	x	P	8.6%	2.5%
					S	14.6%	5.9%
	x	x		x	P	20.8%	7.5%
					S	25.1%	14.3%
	x		x	x	P	21.8%	11.3%
					S	25.7%	16.3%

Table 4: Change in search performance between filtered and non-filtered modified physician-provided searches, when restricted to the top 40 returned results

Methods Filter		Content Filter		Journal Filter	Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)				
Broad	*Narrow*	*Broad*	*Narrow*		*Measure*	*Mean difference*	*Median difference (99% CI)*	*p-value Wilcoxon*	*p-value Sign test*
x					P	0.1%	0.00 (0.00 to 0.00)	**0.0006**	**<0.0001**
					S	0.6%	0.00 (0.00 to 0.00)	0.0579	0.0074
	x				P	17.6%	10.00 (2.50 to 16.94)	**<0.0001**	**<0.0001**
					S	13.4%	0.00 (0.00 to 13.64)	**<0.0001**	**<0.0001**
		x			P	-1.0%	0.00 (0.00 to 0.00)	0.0019	0.0019
					S	-2.8%	0.00 (0.00 to 0.00)	0.0057	0.0059
			x		P	0.1%	0.00 (0.00 to 0.00)	0.6685	1.0000
					S	-1.3%	0.00 (0.00 to 0.00)	0.2614	0.7011
x		x			P	-1.0%	0.00 (0.00 to 0.00)	0.2472	0.7493
					S	-1.4%	0.00 (0.00 to 0.00)	0.3779	0.5716
x			x		P	0.2%	0.00 (0.00 to 0.00)	0.2413	0.0961
					S	-0.3%	0.00 (0.00 to 0.00)	0.7910	0.5966
	x	x			P	14.1%	7.50 (2.50 to 14.19)	**<0.0001**	**<0.0001**
					S	12.9%	0.00 (0.00 to 11.11)	**<0.0001**	**<0.0001**
	x		x		P	16.2%	8.75 (2.50 to 15.00)	**<0.0001**	**<0.0001**
					S	14.1%	0.00 (0.00 to 16.67)	**<0.0001**	**<0.0001**

Continued on following page…

Methods Filter		Content Filter		Journal Filter	Difference in performance between filtered and non-filtered physician-provided searches (P=precision; S=sensitivity)				
Broad	*Narrow*	*Broad*	*Narrow*		*Measure*	*Mean difference*	*Median difference (99% CI)*	*P-value Wilcoxon*	*p-value Sign test*
				x	P	2.6%	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
					S	0.5%	0.00 (0.00 to 0.00)	0.0427	0.0066
x				x	P	2.8%	0.00 (0.00 to 0.00)	**<0.0001**	**<0.0001**
					S	1.3%	0.00 (0.00 to 0.00)	0.0235	**0.0009**
	x			x	P	18.1%	10.00 (2.50 to 17.50)	**<0.0001**	**<0.0001**
					S	12.8%	0.00 (0.00 to 14.29)	**<0.0001**	**<0.0001**
		x		*x*	P	1.4%	0.00 (0.00 to 0.00)	0.0450	0.0079
					S	-0.4%	0.00 (0.00 to 0.00)	0.6876	0.2962
			x	x	P	2.6%	0.00 (0.00 to 2.05)	0.0055	**0.0009**
					S	0.2%	0.00 (0.00 to 0.00)	0.4776	0.1325
x		x		x	P	1.5%	0.00 (0.00 to 2.14)	0.0143	0.0022
					S	0.5%	0.00 (0.00 to 0.00)	0.1280	0.0237
x			x	x	P	3.0%	0.00 (0.00 to 2.50)	**<0.0001**	**<0.0001**
					S	1.9%	0.00 (0.00 to 0.00)	0.0270	**0.0005**
	x	x		x	P	15.2%	7.50 (2.50 to 14.71)	**<0.0001**	**<0.0001**
					S	12.3%	0.00 (0.00 to 15.38)	**<0.0001**	**<0.0001**
	x		x	x	P	16.3%	7.50 (2.50 to 15.00)	**<0.0001**	**<0.0001**
					S	13.0%	1.76 (0.00 to 16.00)	**<0.0001**	**<0.0001**

CHAPTER 9: CURRICULUM VITAE

Name:	Salimah Shariff
Post-secondary Education and Degrees:	University of Western Ontario London, Ontario, Canada 2006-2011 Ph.D. Epidemiology & Biostatistics
	University of Waterloo Waterloo, Ontario, Canada 2000-205 BMATH Computer Science/Bioinformatics
Honours and Awards:	Canadian Institutes of Health Research Doctoral Research Award 2009-2011
	Ontario Ministry of Training, Colleges and Universities Ontario Graduate Scholarship 2007-2008
	University of Western Ontario Schulich Graduate Scholarship for Medical Research 2006-2011
	University of Western Ontario Western Graduate Research Scholarship, Science & Technology 2006-2007
Related Work Experience:	Bioinformatics Analyst University of Toronto 2005-2006
	Research Analyst University Health Network 2003-2004
	Clinical Data Management Specialist Pharmacia Canada 2001-2002
Research Support:	*Filtering Medline for Renal Information: Renal Sub-Filters* Canadian Institutes of Health Research Operating Grant Co-applicant; 03/2011 – 02/2012; $59,951
	Impact of MEDLINE filters on evidence retrieval Canadian Institutes of Health Research Operating Grant Co-applicant; 04/2009 – 03/2010; $85,142

Publications:

Shariff SZ, Bejaimal SAD, Sontrop JM et al. Searching for medical information online: A survey of Canadian nephrologists. *Journal of Nephrology* (in press): 2010.

Shariff SZ, Cuerden MS, Haynes RB et al. Evaluating the impact of MEDLINE filters on evidence retrieval: study protocol. *Implement Sci.* 2010 Jul 20;5:58.

Iansavichus AV, Haynes RB, **Shariff SZ** et al. Optimal search filters for renal information in EMBASE. *Am J Kidney Dis*. 2010 Jul;56(1):14-22.

Garg AX, Iansavichus AV, Wilczynski NL, Kastner M, Baier LA, **Shariff SZ**, Rehman F, Weir M, McKibbon KA, Haynes RB. Filtering MEDLINE for a clinical discipline. *BMJ* 2009 Sep 18;339:b3435. doi: 10.1136/bmj.b3435.

Horvat LD, **Shariff SZ**, Garg AX. Global trends in the rates of living kidney donation. *Kidney Int* 2009 May;75(10):1088-98.

Shariff SZ, Cuerden MS, Jain AK, Garg AX. The secret of immortal time bias in epidemiologic studies. *J Am Soc Nephrol* 2008 May;19(5):841-3.

Burgetz IJ, **Shariff S**, Pang A, Tillier ER. Positional homology in bacterial genomes. *Evol Bioinform Online* 2006;2:77-90.