Retrieval of Sibling Studies for Clinical Randomised Controlled trials

Faten F. Hamad

Department of Information Studies
Aberystwyth University, UK

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy

2012

ABSTRACT

Aims and objectives. For a particular randomised controlled trial, it is often useful to retrieve associated siblings - qualitative research, process and economic evaluations done alongside the randomised controlled trial (RCT). This thesis examines both the effectiveness and efficiency of search strategies, and the productivity of different databases, in retrieving sibling studies for an RCT. **Methods.** Five seed studies from different clinical areas were selected. A range of Boolean searches with simple subject term combinations with authors' names, together with citation and similarity search strategies, were applied, on different databases that had different subject coverage and interests. Specialised search filters were combined with the simple search strategy and tested. The retrieval performances of the simple and sophisticated search strategies on PubMed were tested and compared using one of the seed studies as a case study. Recall, precision and odds estimators were used for all retrieval tests. Non-parametric statistical tests were used to test a set of hypotheses that set out to explore relationships underlying retrieval performance. **Results.** Neither one particular search strategy nor one database was an overall winner. The simple author-subject search provided a good recall with a readable retrieval size. The recall varied among seed studies and different databases. Search filters provided good recall for retrieving specific types of sibling, especially the qualitative filter. PubMed related articles strategy provided a good performance for some seeds, but not as good overall as the simple author-subject searching. Combining a similarity search with simple author-subject search provided complementary retrieval performance and therefore yields an optimal performance. Citing search did not perform well in terms of retrieving sibling studies. The simple author-subject search shows performance consistency, being the best search strategy among other strategies for all seed studies in terms of recall and precision. WoK and SCOPUS were the best databases for retrieving sibling studies. **Conclusions.** Simple author-subject search, especially when searching multiple databases, can yield an optimal performance in retrieving sibling studies.

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed .... (candidate)

Date ....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where ***correction services** have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references.
A bibliography is appended.

Signed .... (candidate)

Date ....

[*this refers to the extent to which the text has been corrected by others]

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate)

Date ....

ACKNOWLEDGEMENTS

It is a great pleasure to thank everyone who helped me write my dissertation successfully. For that I would like to express my sincere and earnest gratitude to my supervisor, Dr Christine Urquhart, for her outstanding support, guidance, patience. This thesis would not have been possible without Dr. Urquhart's help, as she gave me focused and relevant advice all the way through and above all she gave me friendship for which I am extremely thankful.

I would like to acknowledge the support of my parents, brother's and other family members, especially my cousin Fatema and thank them for their unequivocal support throughout. I am obliged to many of my friends who supported me.

And finally, I would like to thank the Department of Information Studies and my second supervisor Dr Pauline Rafferty for their encouragement and support.

TABLE OF CONTENTS

ABSTRACT ii

DECLARATION iii

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS v

LIST OF TABLES xi

LIST OF FIGURES xiii

ABBREVIATIONS xiv

Chapter One 1

Introduction 1

1.1 Systematic Reviews – an Overview 1
1.2 Beyond the RCT – Getting Evidence into Practice 2
1.3 Research Problem and the Definition of Sibling Studies 3
1.4 Research Questions and Hypotheses 7
1.5 Personal Perspective on the Research 9
1.7 Potential Contribution to the Body of Knowledge 10
1.8 Structure of the Thesis 10

Chapter Two 13

Systematic Review and Sibling Studies: Background and Examples 13

2.1 Introduction 13
2.2 Systematic Review Overview 14
2.3 Economic and Cost Effectiveness Research 15
2.3.1 Economics Evaluation: Summary of Characteristics 20
2.4 Qualitative Research Studies 22
2.4.1 Qualitative Studies: Summary of Characteristics 33
2.5 Assessment Studies / Process Evaluation 34
2.5.1 Process Evaluation: Summary of Characteristics 41
2.6 Summary 43

Chapter Three 44

The Context of Information Retrieval and Trends in IR Research 44

3.1 Introduction 44
3.2 Information Retrieval 45
3.2.1 What is an Information Retrieval System? 45
3.2.2 Information Retrieval: Fom Childhood to Maturation 47
3.2.3 Subject Representation and Indexing 49
3.2.4 Users' Needs and Interface 52
3.3 Information Retrieval Characterisation 53
3.3.1 Matching and Retrieval 53
3.3.1.1 Retrieval Models 53

3.3.2 Information Retrieval Research Trends 55
3.3.3 Information Retrieval Research for Systematic Reviews 57
3.4 Digital Libraries 63
3.5 Searching for Clinical Studies 66
3.5.1 Searching for Types of Clinical Study 66
3.5.2 Related Articles - Similarity Searching 69
3.5.3 PubMed Automatic Term Mapping (ATM) 70
3.5.4 Search Filters 71
3.5.5 Qualitative Research Retrieval 75
3.5.6 Clinical Trial Number and Clinical Registries 75
3.6 Retrieval Effectiveness Measurements 77
3.6.1 Recall and Precision 78
3.6.2 Relevancy 80
3.6.3 Database Performance 84
3.7 Text Mining in Health Research 87
3.8 Summary 91

Chapter Four 92

Methodology 92

4.1 Introduction 92
4.1.1 Resources to Locate Existing Literature 94
4.2 Seed Study Selection 95
4.3 Search Strategies and Databases 96
4.4 Phases of the Study - Overview 99
4.5 Metrics for Analysis 105
4.6 Phase I: Exploring the Proposed Search Strategies Performance - Pilot Study . 112
4.7 Phase II: Exploring Combined Search Strategies/Search Filters and Databases Performance. 113
4.8 Phase III: Direct and Indirect Siblings Identification 114
4.9 Phase III: Direct Sibling Identification 114
4.10 Research Hypotheses and Hypotheses Testing 115
4.11 Summary: Reflection on Methodology 120

Chapter Five 122

Relevant Retrieval: An Explanatory Phases for Siblings Retrieval 122

5.1 Introduction 122
5.2 Pilot Study 123
5.3 Second Stage Search Strategies 125
5.3.1 Diabetes Telemedicine (IDEATeL) Seed Study 125
5.3.1.1 Recall 125
5.3.1.2 Precision 125
5.3.1.3 Odds Estimator 126
5.3.1.4 Sibling Studies Retrieval by Study Type 127
5.3.2 Breast Cancer (Tamoxifen) Seed Study 130
5.3.2.1 Recall 130
5.3.2.2 Precision 130

5.3.2.3 Odds Estimator 130
5.3.2.4 Sibling Studies Retrieval by Study Type 131
5.3.3 Breast Cancer (Body Mind Spirit) Seed Study 133
5.3.3.1 Recall 134
5.3.3.2 Precision 134
5.3.3.3 Odds Estimator 134
5.3.3.4 Sibling Studies Retrieval by Study Type 135
5.3.4 In Vitro Fertilisation (hMP vs rFSH) Seed Study 137
5.3.4.1 Recall 137
5.3.4.2 Precision 138
5.3.4.3 Odds Estimator 138
5.3.4.4 Sibling Studies Retrieval by Study Type 139
5.3.5 Chronic Lung Disease (Dexamethasone) Seed Study 142
5.3.5.1 Recall 142
5.3.5.2 Precision 142
5.3.5.3 Odds Estimator 142
5.2.5.4 Sibling Studies Retrieval by Study Type 143
5.3.6 Summary: Reflection on Retrieval Performance of Search Strategies and Databases 146
5.4 Indirect Siblings 148
5.4.1 Diabetes Telemedicine (IDEATeL) Seed Study 148
5.4.1.1 Recall 148
5.4.1.2 Precision 149
5.4.1.3 Odds Estimator 149
5.4.1.4 Sibling Studies Retrieval by Study Type 151
5.4.2 Breast Cancer (Tamoxifen) Seed Study 153
5.4.2.1 Recall 153
5.4.2.2 Precision 153
5.4.2.3 Odds Estimator 153
5.4.2.4 Sibling Studies Retrieval by Study Type 155
5.4.3 Breast Cancer (BMS) Seed Study 157
5.4.3.1 Recall 157
5.4.3.2 Precision 157
5.4.3.3 Odds Estimator 157
5.4.3.4 Sibling Studies Retrieval by Study Type 159
5.4.4 Chronic Lung Disease (Dexamethasone) Seed Study 160
5.4.4.1 Recall 160
5.4.4.2 Precision 161
5.4.4.3 Odds Estimator 161
5.4.4.4 Sibling Studies Retrieval by Study Type 163
5.4.5 In Vitro Fertilisation (IVF) Seed Study 165
5.4.5.1 Recall 165
5.4.5.2 Precision 165
5.4.5.3 Odds Estimator 165
5.4.5.4 Sibling Studies Retrieval by Study Type 167
5.4.6 Summary: Reflection on Search Strategies and Databases Performance 168

Chapter Six **171**

Direct Sibling Retrieval **171**

6.1 Introduction 171
6.2 Diabetes-Telemedicine - IDEATeL Seed Study 171
6.2.1 Recall 172
6.2.2 Precision 172
6.2.3 Odds Estimator 172
6.2.4 Sibling Studies Retrieval by Study Type 173
6.2.5 IDEATeL Missing Siblings 176
6.2.6 IDEATeL Siblings Overall Retrieval Process 178
6.2.7 Sophisticated Search vs Simple Search 179
6.3 Breast Cancer (Tamoxifen) Seed Study 179
6.3.1 Recall 179
6.3.2 Precision 180
6.3.3 Odds Estimator 180
6.3.4 Sibling Studies Retrieval by Study Type 181
6.4.5 Breast Cancer (Tamoxifen) Overall Retrieval Process 184
6.4 Breast Cancer (BMS) Seed Study 185
6.4.1 Recall 185
6.4.2 Precision 185
6.4.3 Odds Estimator 185
6.4.4 Sibling Studies Retrieval by Study Type 187
6.4.5 Breast Cancer (BMS) Siblings Overall Retrieval Process 189
6.4.6 Checking Breast Cancer (BMS) studies 190
6.5 Chronic Lung Disease (Dexamethasone) Seed Study 190
6.5.1 Recall 191
6.5.2 Precision 191
6.5.3 Odds Estimator 191
6.5.4 Sibling Studies Retrieval by Study Type 193
6.5.5 CLD Siblings List 195
6.5.6 CLD Siblings Overall Retrieval Process 195
6.6 In Vitro Fertilisation (IVF) Seed Study 197
6.6.1 Recall 197
6.6.2 Precision 197
6.6.3 Odds Estimator 197
6.6.4 Sibling Studies Retrieval by Study Type 199
6.6.5 IVF Siblings Overall Retrieval Process 201
6.7 Research Hypotheses and Hypotheses Test 202
6.8 Sibling Studies Publication Time 205
6.9 Multiple Database Retrieval 206
6.9.1 Unique Retrieval and Overlap 207
6.10 Summary: Reflection on Retrieval Performance 212
6.10.1 Sibling Studies Retrieval 216

Chapter Seven 222

Discussion 222

7.1 Introduction 222
7.2 Search Strategies and Databases: Main Findings 223
7.2.1 Effective and Precise Search Strategy(s) 223
7.2.2 Database Performance Comparisons 226
7.2.3 Retrieval Performance for Particular Type of Siblings 228
7.2.3.1 RCT Siblings 228
7.2.3.2 Qualitative Siblings 230
7.2.3.3 Economic Evaluation Siblings 233
7.2.3.4 Process Evaluation Siblings 236
7.2.4 Sibling Studies Publication Time 237
7.2.5 Search Strategies for Systematic Review 238
7.3 Clinical Area and Research Type 239
7.3.1 Overview: Main Findings 239
7.3.2 Authorship and Group Authorship 241
7. 3.3 Derivation of Gold Standard According to Authors 245
7. 3.4 Reference Lists 246
7.4 Search Filters 246
7.5 Multiple Database(s) Search 248
7.6 Databases Selection Implication on Siblings Retrieval 250
7.7 Simple Search vs Sophisticated Search – With Reference to IDEATeL 252
7.7.1 PubMed Automatic Term Mapping (ATM) 255
7.8 Text Mining 257
7.9 Indirect Sibling Studies Retrieval 258

Chapter Eight 259

Conclusion 259

8.1 Introduction 259
8.2 Main Findings - Summary 263
8.3 Research Limitations 264
8.4 Contribution to Knowledge 266
8.5 Future Work and Recommendations 268

Appendix One 272

Search Strings 272

Appendix Two 285

PubMed Sophisticated Search String and ATM 285

Appendix Three 297

Siblings Retrieval Performance 297

Appendix Four 309

Normality test 309

Appendix Five 316
Pilot Study Retrieval Performance 316
Appendix Six 321
Sibling lists for all Seed Studies 321
Appendix Seven 343
IDEATel Publications 343
Appendix Eight 350
The Association Between Odds Estimator and Precision 350
Appendix Nine 351
Information Sheet 351
References 352

LIST OF TABLES

Table 1: Two by Contingency Table for Retrieval Performance Calculation --------------- 79
Table 2: Databases Search Dates---100
Table 3: Two by Two Table Showing Modified Odds Estimator Calculations -----------109
Table 4: Siblings Retrieval from PubMed Related Articles Search of the IDEATeL Seed Study---109
Table 5: Siblings Retrieval from CINAHL Author-Subject Search of IDEATeL Seed Study ---110
Table 6: Search Strategies and Databases Ranks (OR and OE Ranks)---------------------112
Table 7: Three Search Strategies and Databases (Pilot Retrieval of Relevant on Topic)-124
Table 8: The IDEATeL Search and Databases Retrieval-------------------------------------127
Table 9: IDEATeL Search Strategies and Databases Odds Estimator per Sibling Type--129
Table 10: Breast Cancer (Tamoxifen): Search and Databases Retrieval Performance ---131
Table 11: Tamoxifen Search Strategies and Databases Odds Estimator per Sibling Type133
Table 12: The BMS Search and Databases Retrieval --135
Table 13: BMS Search Strategies and Databases Odds Estimator per Sibling Type ------137
Table 14: The IVF Search Strategies and Databases Retrieval -------------------------------139
Table 15: IVF Search Strategies and Databases Odds Estimator per Sibling Type--------141
Table 16: The CLD Search Strategies and Databases Retrieval------------------------------143
Table 17: CLD Search Strategies and Databases Odds Estimator per Sibling Type ------145
Table 18: Top Three Search Strategies and Databases in Retrieving Relevant (on-topic) Studies---147
Table 19: The IDEATeL Search Strategies and Databases Retrieval------------------------150
Table 20: The IDEATeL Search Strategies and Databases Odds Estimator per Sibling Type ---152
Table 21: Breast Cancer (Tamoxifen): Search and Databases Retrieval Performance ---154
Table 22: The Tamoxifen Search Strategies and Databases Odds Estimator per Sibling Type ---156
Table 23: The BMS Search Strategies and Databases Retrieval------------------------------158
Table 24: BMS Search Strategies and Databases Odds Estimator per Sibling Type. -----160
Table 25: The CLD Search Strategies and Databases Retrieval------------------------------162
Table 26: CLD Search Strategies and Databases Odds Estimator per Sibling Type ------164
Table 27: The IVF Search Strategies and Databases Retrieval -------------------------------166
Table 28: IVF Search Strategies and Databases Odds Estimator per Sibling Type--------168
Table 29: Top Three Search Strategies and Databases Retrieval to Retrieve Direct and Indirect Siblings---170
Table 30: The IDEATeL Search Strategies and Databases Retrieval------------------------173
Table 31: IDEATeL Search Strategies and Databases Odds Estimator per Sibling Type 175
Table 32: Comparison of Simple and Sophisticated Search Strategies for Siblings Retrieval on PubMed---179
Table 33: The Tamoxifen Search Strategies and Databases Retrieval-----------------------181
Table 34: Tamoxifen Search Strategies and Databases Odds Estimator per Sibling Type183
Table 35: The BMS Search Strategies and Databases Retrieval------------------------------186
Table 36: BMS Search Strategies and Databases Odds Estimator per Sibling Type ------188
Table 37: The CLD Search Strategies and Databases Retrieval------------------------------192
Table 38: CLD Search Strategies and Databases Odds Estimator per Sibling Type ------194
Table 39: The IVF Search Strategies and Databases Retrieval -------------------------------198

Table 40: IVF Search Strategies and Databases Odds Estimator per Sibling Type--------200
Table 41: Test of Significance of Retrieval Measurements and Search Strategy ----------202
Table 42: Test of Significance of Retrieval Measurements and Database ------------------203
Table 43: Test of Significance of Retrieval Measurements and Clinical Area -------------203
Table 44: Recall and Clinical Area Dependency ---204
Table 45: Precision and Clinical Area Dependency --204
Table 46: Odds Estimator and Clinical Area Dependency ------------------------------------204
Table 47: Number of Siblings Retrieved and Clinical Area Dependency-------------------204
Table 48: Sibling Studies Publication Time ---205
Table 49: Multiple Databases Retrieval Performance---------------------------------------207
Table 50: Unique Contribution of Search Strategies and Databases-------------------------211
Table 51: Summary of Search Strategies and Databases Retrieval Performance ----------214
Table 52: Search Strategies and Databases Average Performance for Siblings Retrieval (Average Recall, Precision and OE) --216
Table 53: Summary of Top Three RecallThree Recalls of Search Strategies/Databases per Sibling Type --219
Table 54: Summary of Top Three Precision of Search Strategies/Databases per Sibling Type--220
Table 55: Summary of Top Three Odds Estimator of Search Strategies/Databases per Sibling Type --221
Table 56: Comparison Between Systematic Review and Sibling Studies-------------------239
Table 57: Some of the Selected Databases Pros and Cons ---------------------------------251

LIST OF FIGURES

Figure 1: Retrieval Process of Qualitative Studies of RCTs. (Source: Lewin et al., 2009) 24
Figure 2: Information Retrieval Process (Source: Heimstra, 2009)--------------------------- 46
Figure 3: Information Retrieval Evolution and Improvements --------------------------------- 48
Figure 4: Screening Process for Sibling Studies Identification-------------------------------103
Figure 5: SPSS File for Data Set 1--119
Figure 6: SPSS File for Data Set 2--120
Figure 7: IDEATeL Siblings Retrieval--176
Figure 8: Overall Siblings Retrieval & Identification for IDEATeL Seed Study----------178
Figure 9: Overall Siblings Retrieval & Identification Process for Tamoxifen Seed Study --184
Figure 10: Overall Siblings Retrieval & Identification Process for BMS Seed Study ----189
Figure 11: Overall Siblings Retrieval & Identification Process for CLD Seed Study-----196
Figure 12: Overall Siblings Retrieval & Identification Process for the IVF Seed Study -201
Figure 13: Databases Retrieval and Unique Contribution for IDEATeL Siblings --------208
Figure 14: Databases Retrieval and Unique Contribution for Tamoxifen Siblings--------208
Figure 15: Databases Retrieval and Unique Contribution for BMS Siblings---------------209
Figure 16: Databases Retrieval and Unique Contribution for CLD Siblings ---------------209
Figure 17: Databases Retrieval and Unique Contribution for IVF Siblings ----------------210

ABBREVIATIONS

AHRQ	Agency for Healthcare Research and Quality
AS	Author-Subject
ASSERT	Automatic Summarisation for Systematic Reviews using Text Mining
ATM	Automatic Term Mapping
BMJ	British Medical Journal
BMS	Body Mind Spirit (A reference to the Liu et al. breast cancer seed study)
BNI	British Nursing Index
CINAHL	Cumulative Index to the Nursing and Allied Health Literature
CLD	Chronic Lung Disease (A reference to the Vermont Oxford Network Steroid Study Group seed study)
CONSORT	Consolidated Standards of Reporting Trials
CRD	Centre for Reviews and Dissemination
E-Library	Electronic Library
EMBASE	Excerpta Medica Database
HiRU	Health Information Research Unit
ESRD	End Stage Renal Disease
HSSS	Highly Sensitive Search Strategy
$HSSS_{12}$	Highly Sensitive Search Strategy (phases one and two)
$HSSS_{123}$	Highly Sensitive Search Strategy (phases one, two and three)
ICMJE	International Committee of Medical Journal Editors
ICTRP	International Clinical Trials Registry Platform
IDEATel	Informatics for Diabetes Education and Telemedicine
IDNT	Irbesartan in Diabetic Nephropathy Trial
IME	Indice Médico Español
IR	Information Retrieval
IRMA	International Microvascular Abnormalities
IVF	In Vitro Fertilisation
JAMA	The Journal of the American Medical Association
LMS	Library management systems
MARC	Machine-Readable Cataloguing
MEDLARS	Medical Literature Analysis and Retrieval System
MEDLINE	Medlars Online
MeSH	Medical Subject Heading
NaCTeM	The National Centre for Text Mining
NHS	National Health Service
NHS EED	NHS Economic Evaluation Database

NICE	The National Institute for Health and Clinical Excellence
NIST	US National Institute of Standards and Technology
NLM	National Library of Medicine
R/N	Relevant and not Retrieved
RR	Relevant and Retrieved
OR	Odds Ratio
OE	Odds Estimator
PICO	Population, Intervention, Comparison and Outcome
PRESS	Peer Review of Electronic Search Strategies
PST	Trials of Problem Solving Therapies
PsycINFO	Psychological Information Database
RCT	Randomised Controlled Trial
RR	Relevant and Retrieved
SCI	Science Citation Index
SPSS	Statistical Package for Social Sciences
SR	Systematic Review
TREC	Text Retrieval Conference
WoK	Web of Knowledge
WoS	Web of Science

Chapter One
Introduction

1.1 Systematic Reviews – an Overview

The popularity and importance of systematic reviews of the health research evidence have increased these days for many varieties of users, i.e. professional practitioners, managers, decision makers for planning and policy, and researchers. Primarily, systematic reviews aim to inform decision makers, to ensure decisions are made on the basis of the best available research evidence in the area of interest. However, conducting systematic reviews is considered to be challenging due to the massive amount of research studies that may exist in one research area or health topic, and moreover, each individual study has certain characteristics and settings that may differ from another such as study design, data collection and data analysis methods and reporting methods. As a solution, systematic reviews adhere to a specific protocol to deal with the diversity of studies, to fulfil the main objective of providing a high quality, trustworthy review. Even with a recognised protocol, systematic reviews are time-consuming, involving searching for the relevant studies, management of the search and screening, decisions about inclusion and exclusion, and data analysis of the included studies.

McGowan and Sampson (2005), two experienced information scientists involved in systematic reviews, note that a systematic review (SR) can be defined as:

> A review that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. (p.75)

In order to get a high quality systematic review, a SR should maintain the quality of the included studies in the SR, but before that, successful identification and retrieval of all related studies that address/answer the questions for that SR is the key to a reliable and authoritative SR.

The search strategy used in retrieval processes is a critical part of conducting a systematic review. Successful and unbiased retrieval of the entire evidence base relevant to an SR topic depends mainly on the search strategy and to what degree it is possible for the searches to maximise sensitivity, specificity and precision (Wilczynski et al., 2004; McGowan & Sampson, 2005; Wilczynski, Marks, & Haynes, 2007).

In addition, there are usually several databases that could be rich sources of most of the available (published) evidence, and, frequently, searching several databases is recommended due to the differences between databases' coverage of articles including selection procedures and indexing processes (Sampson et al., 2006b). In other words, to increase SR quality (and limit the bias) a comprehensive literature search is required to identify as much of the relevant literature as possible. Search terms and filters used should also be taken into consideration when carrying out the search, as filters may help to limit the search output to the clinical aspect of interest (e.g. diagnosis or prognosis). Each database may have its specific search method, filters and interface which searchers should examine when conducting the search process to avoid search errors that may adversely affect sensitivity, specificity and precision (Sampson & McGowan, 2006). Search filters have been devised by various groups – but principally the Hedges team at McMaster University in Canada (HiRU, 2007).

1.2 Beyond the RCT – Getting Evidence into Practice

Up until recently most emphasis has been placed on inclusion of quantitative research designs, particularly randomised controlled trials (RCTs), as these provide the least possible bias. RCTs help to answer whether one intervention is better, on average, than another intervention. However, there are other questions of interest to patients and policymakers in health research. Therefore, different types of studies may need to be clearly identified and retrieved to increase not only the robustness and reliability of the SR, but also the transferability of knowledge into practice.

These studies aim to address different yet relevant and important knowledge in the same clinical field as the topic of the RCT, such as economic issues, patients' attitudes or perspectives, professional attitudes to the interventions, as these associated factors could have direct or indirect effects on implementation and effectiveness of the intervention in a particular setting. That type of knowledge may help to provide information on how to implement an intervention in an efficient and effective manner. Thus such studies complement each other and complement the RCT.

A large body of research exists on search filters for retrieving RCTs, dealing with publication bias, but the retrieval research on non-RCT studies is much less advanced. Accordingly, this research will investigate the challenges that emerge when retrieving studies with research design and objectives that complement RCTs.

We know that RCTs may be accompanied, preceded or followed by research that is associated in some way with the RCT. The main objective is to investigate the performance of various search strategies, and different databases, where the main concern is to retrieve the complementary and associated studies. These could be other RCT publications, or the associated research that aims to provide knowledge about how an intervention should be implemented, or knowledge about the cost-effectiveness of the intervention.

1.3 Research Problem and the Definition of Sibling Studies

As indicated above, for guideline development and implementation, other studies with different research designs to the randomised controlled trials (RCTs) can provide information about the feasibility of the intervention. For example, qualitative evidence helps to ensure that professional and patient views of the experience of any intervention are considered. Economic and process evaluations help policymakers and health service managers to implement best practice and what the changes will cost (Harden et al., 2004). The Joanna Briggs Institute (2008) emphasises consideration of social, cultural and economical factors together with the clinical aspect of health care for stronger evidence.

According to Pawson et al. (2004) social interventions:

> ...are complex systems thrust amidst complex systems and are never implemented the same way twice. Non-equivalence is the norm. Realists envision interventions as whole sequences of mechanisms that produce diverse effects according to context, so that any particular intervention will have its own particular signature of outputs and outcomes (p. 33).

and implementing many health interventions requires both quantitative and qualitative evidence (Campbell et al., 2000).

Each clinical area may be different, or have different priorities, depending on the health technology involved. The following are the main types of studies which are associated with RCTs:

- Qualitative research or studies.
- Process evaluations (process assessments and outcome evaluations).
- Economic evaluations (cost analyses and economic effectiveness).
- And RCTs themselves, of course.

This doctoral research used five studies each of which will be referred to as the "seed study". Seed studies were chosen from different clinical areas in order to investigate the differences and influence the clinical area might have on retrieval performance.

The name sibling studies was the name provided by the Cochrane Information Retrieval Methods Group to describe the relationship among a set of related studies that would include a randomised controlled trial, and complementary qualitative, process, or economic evaluations. Accordingly, the term "Sibling studies" used (in this PhD research) refers to studies that are based on or emerge from the seed studies and aim to investigate other aspects that may interfere with, affect or explain the intervention output and employ a different or the same research design than the seed study. This doctoral research was done under the auspices of the Cochrane Information Retrieval Methods Group, and the main aim is to target the direct siblings for a particular RCT, defined as the studies that would be done by, or with some association to, the research team(s) involved in that RCT.

The original aim of this study was to retrieve sibling studies of an RCT. However as the research progressed and evolved it became obvious that identifying other relevant qualitative, process or economic studies would also be useful to practitioners and policymakers, and two type of siblings were differentiated, direct siblings and indirect siblings.

> **Direct siblings** use the same or different research design to investigate other factors that may interfere, affect or explain the intervention output. Direct sibling studies must share at least one of the seed study authors.
>
> **Indirect siblings** use the same or different research design to investigate other factors that may interfere, affect or explain the intervention output, but do not however share any author with the seed study authors.

It also became clear that there were difficulties when deciding on the boundary between a direct and indirect sibling, or even between an indirect sibling and other possibly relevant on-topic research, as the processes of scholarly communication and collaboration are complex, and not always visible. The influences, and researcher working relationships associated with a piece of research may or may not be acknowledged, therefore, the term indirect siblings has been chosen to describe the relevant studies that have a closer relationship with the seed study but are not direct siblings. In other words, they are not direct siblings (no shared authors with the seed study authors) but have some loose relationship, i.e. relevant ("on topic") and cite or are cited by the seed study.

Up to now most effort has been placed on search strategies to retrieve RCTs themselves. Other search filters have been devised, for example, to retrieve studies on therapy, or diagnostic studies. There are also some search strategies to retrieve qualitative research and some search strategies devised to retrieve economic studies. Some search strategies for each category have been developed by the Hedges Project at McMaster University (Wong et al., 2003; Wilczynski et al., 2004; Haynes & Wilczynski, 2004; Wong, Wilczynski & Haynes, 2006; Wilczynski, Marks, & Haynes, 2007; Wilczynski, McKibbon,& Haynes, 2007; HiRU, 2007).

Sibling studies do not appear to receive much attention in the existing literature. One can find studies addressing qualitative research and its important role in explaining how an intervention works and why it has those outcomes. Some studies are about economic issues of interventions, and others are concerned with delivery of the intervention on a large scale. Each category of these studies has its specific search terms and key words to identify the topic of the studies. Therefore, integrating the common factors between the siblings composing comprehensive search filter/terms to identify the siblings in one search would be helpful and convenient.

There are other approaches to finding related studies that do not rely on the usual combination of search filter (for research design) plus topic term set. The Science Citation Index may be used to find articles that cite a particular document published previously. PubMed has a related articles feature that provides a direct link to the set of articles that discuss issues related the original article directly or other issues that emerge from it. As both of these approaches are in tune with the concept of "sibling studies" such approaches are examined as well.

In addition, the Google effect cannot be ignored. Increasingly, searchers expect search outputs to provide the most relevant items first, and this research must also take account of probabilistic searching and associated metrics. Although purist information professional searchers may not be in favour of meta-database searching as it is much less precise than searching single databases, cross database searching (federated searching) has been possible for some time, and many university libraries offer MetaLib type searches. It seems sensible to include an evaluation of federated searching as well.

For a robust and unbiased systematic review it is necessary to conduct a search that is sensitive enough to capture all relevant studies on the one hand, but specific enough not to conceal them, inadvertently, through retrieval of a large number of non-relevant studies. The initial objective of this doctoral research was to retrieve studies that are related to a certain study, referred to as the seed study in this research. The search strategy is required to be comprehensive and sufficiently sensitive to retrieve all relevant sibling studies of the seed study. However, as noted, specificity is also important, to ease the burden of sifting through a large number of unwanted or non-relevant studies. This is an important consideration as such searches may require a much wider search than conventional systematic review searches to retrieve randomised controlled trials. Simple search strategies have been emphasised in this research using general subject terms that describe the seed study generally. No further details about the seed study such as publication type were used, to avoid restricting the retrieval to a specific research design. However, a later phase of the study compares the retrieval of the highly sensitive sophisticated search strategy with the simple search strategy (for PubMed).

There is some justification for a simple search (Chapter 3) as other researchers have found that using a simple search made no difference to the conclusion of systematic reviews which used this search approach when compared to the more exhaustive highly sensitive search strategies. Managers and policymakers are likely to value search strategies that are more likely to retrieve relevant material than irrelevant material - time constraints influence their satisficing strategies when information seeking (MacDonald, Bath, & Booth, 2011).

In conclusion, identifying the best evidence for a clinician needs integration of epidemiological and bio-statistical research with knowledge derived from pathophysiology and personal experience, ultimately incorporating meta-analyses of randomised trials into decisions about therapy, economic analyses and the use of decision analyses (Sackett & Rosenberg, 1995). Retrieval of RCTs has been a priority and progress has been made on the development of search filters that aim to reach an optimal ratio of sensitivity, specificity and precision, but what about the siblings?

1.4 Research Questions and Hypotheses

This research takes a different approach to the development and evaluation of comprehensive search strategies devised to retrieve relevant studies from various databases. It examines simple search strategies, as well as the sophisticated search strategies, that may be used to retrieve the sibling studies directly or indirectly associated with a particular RCT. The research also examines how sibling association can be determined.

Thus, the major research objective of this thesis is to explore the performance characteristics for various information retrieval approaches i.e. search strategies and databases. The main research question to be addressed is:

> *Is there one or more efficient search strategy(s) to retrieve qualitative, economic, and/or process studies that may be associated directly with the seed RCT?*

From that major research question, more specific research questions emerged:

1. How can sibling studies be identified? Are there common characteristics that make the studies siblings?
2. Are subject searching, author searching, related articles and citation searching search strategies effective, and if so, to what extent, in retrieving sibling studies?
3. Which database is considered to be more productive and comprehensive and which provides more unique or reliable studies within a specific time frame?
4. What metrics should be applied to measure retrieval performance and effectiveness of both search strategies and databases, including multiple databases simultaneously searched (federated searching)?
5. Does the clinical area affect the retrieval performance of search strategy or database?
6. Is there any pattern or information to associate the seed study and its siblings, i.e. clinical trial number?
7. How often are siblings published before/simultaneously/after the seed study?

Based on the above research questions the following hypothesis set was generated:

Set 1: Search strategies and databases performance efficiency relationship

H0: There is no difference between databases' precision.

H1: There is a difference between databases' precision.

H0: There is no difference between databases' recall.

H1: There is a difference between databases' recall.

H0: There is no difference between search strategies' precision.

H1: There is a difference between search strategies' precision.

H0: There is no difference between search strategies' recall.

H1: There is a difference between search strategies' recall.

Set 2: Search strategies, databases and siblings retrieval relationship

H0: Database X is not more likely to retrieve sibling studies rather than non-siblings.

H1: Database X is more likely to retrieve sibling studies rather than non-siblings.

H0: Search strategy X is not more likely to retrieve sibling studies rather than non-siblings.

H1: Search strategy X is more likely to retrieve sibling studies rather than non-siblings.

Set 3: Search strategy, databases and relationship with clinical area

H0: Search strategy performance is independent of the clinical topic.

H1: Search strategy performance is dependent on the clinical topic.

H0: Database performance is independent of the clinical topic.

H1: Database performance is dependent on the clinical topic.

1.5 Personal Perspective on the Research

I came from a computer science background, where systems analysis and performance analysis are involved. The idea of information retrieval and database performance did not surface until I enrolled in research training modules for the PhD. There it became much clearer how electronic search and use of databases are becoming the first and main choice of reference for people, with frequent use. Health information is being targeted by both the public and professionals. As information retrieval is all about the interaction between users and information, it offers many opportunities and challenges research which is worth study, and important to many groups of users. Accordingly, information retrieval systems and technology need continuous attention to maintain a high performance to meet the satisfaction of users. From a computer science perspective, one of the main principles that I find practical is to keep things simple. This is to develop the simplest and neatest solutions to a problem – good code is short and effective, long code often has bugs in it.

1.7 Potential Contribution to the Body of Knowledge

This doctoral research makes an original, novel contribution to the body of knowledge. It contributes to the field of information retrieval for systematic reviewing in two ways. Firstly, it explores the existence of a relationship between studies with different research design (siblings) that are related directly or indirectly to an RCT, and considers what factors govern the relationship. Secondly, the research explores the retrieval performance of different search strategies in retrieving siblings from different databases. Search filters have often focused on one research design or clinical aspect – a broader approach was required here. The use of an odds estimator was trialled as a way of providing a comparison metric that combines aspects of both recall and precision, the conventional metrics used (and reported in this research). As search strategies and databases may perform differently among the various seed studies, the research investigated how multiple databases and combination of strategies could provide optimal retrieval of sibling studies. Previous research has tended to examine only one clinical area or topic when coming to conclusions about comparative search strategy performance, whereas in this research study, several clinical areas were investigated, and this demonstrated how dependent performance might be on the clinical topic.

1.8 Structure of the Thesis

Chapter one: Provides a brief description of systematic review and the reasons why it is increasingly important for systematic reviewers and policymakers to consider research that complements the RCT, providing evidence about the economic, social and cultural feasibility of an intervention.

Chapter Two: Discusses systematic reviews in more detail. The chapter also provides examples for each category of siblings in order to provide a more focused picture of the siblings, their characteristics, the importance of these types of studies as well as how such studies complement systematic reviews based on meta-analysis of RCTs.

Chapter Three: Discusses the historical aspects of information retrieval and evaluation as well as addressing recent trends in information retrieval research. This chapter investigates how performance of search strategies and databases is assessed, as well as some approaches that might help in retrieval and identification of research relevant to evidence-based practice.

Chapter Four: Discusses the methodology of information retrieval and the implication this has on the methods used in this research. It addresses the main issues and trends in information retrieval from the perspective of user needs and user satisfaction, and how this is interrelated to systematic reviews and evidence-based practice. This chapter presents the methods used in this research.

Chapter Five: Presents phase one, two and three retrieval performance of search strategies and databases. Phase one is a pilot study to explore search strategies and databases' performance in order to plan the next phases (Appendix five presents the pilot study performance results). Phase two presents the performance of search strategies and databases in retrieving relevant "on topic" studies. Direct and indirect sibling retrieval performance is presented in this chapter (phase three).

Chapter Six: Presents the analysis of the performance for retrieving direct sibling studies. The comparative performances of search strategies and databases are analysed and investigated in phase four (direct siblings (only) retrieval analysis). This chapter presents the statistical test results of the research hypotheses.

Chapter Seven: Discusses the performance efficiency of search strategies and databases in retrieving sibling studies. This chapter provides an explanation of the retrieval performance of search strategies and databases as well as deducing the factors that might impose on specific performance.

Chapter Eight: Concludes the research, with consideration of the research limitations, further improvements for future research and recommendations for practice.

The appendices provide materials to supplement the main data and analysis methods used in this research. Search strings used for each seed study siblings retrieval are listed in **Appendix one**. **Appendix two** presents the PubMed Automatic Mapping for simple and sophisticated search strings. **Appendix three** provides tables demonstrating each sibling type retrieval performance (search strategies and databases). **Appendix four** presents normality tests for statistical tests. In **Appendix five** search strategies and databases performance from the pilot study are presented. **Appendix six** the lists of direct and indirect siblings that were created for each seed study. **Appendix seven** presents the list of siblings that were provided by contact authors and helpes to draw out a gold standard. In **Appendix eight** a figure to demonstrate the association between odds estimator and precision is provided. And finally, **Appendix nine** provides the information sheet that was provided to contact authors when contacted to advice gold standard. This thesis uses Harvard APA citation style to organise references which are alphabetically ordered and presented at the end of the thesis.

Chapter Two
Systematic Review and Sibling Studies: Background and Examples

2.1 Introduction

Sampson *et al.* (2006b, p. 461) and NHS Centre for Reviews and Dissemination (2009, p. 3) suggested several advantages of systematic reviews:

- Examining all related studies together can explore the consistency between different studies' results; therefore the available evidence will be more trustworthy.

- Moreover, variations in study settings and designs provide evidence of robustness and transferability of results to other settings. If the studies are inconsistent between settings, then the sources of variation can be examined.

- Finally, gaps in research evidence and interventions can be identified, letting the researchers to know where to start from.

This chapter aims to explore and investigate information retrieval issues in the creation of systematic reviews, as the purpose of all systematic reviews is to support high quality evidence based practice by identifying and retrieving high quality studies that address the systematic review (SR) problem.

For the doctoral research, one of the questions similarly concerned the type of search strategies required to search efficiently and effectively for sibling studies, and another question concerned how that can be measured, (Section 1.4).

The literature review in this chapter also considers the characteristics of each type of sibling. Examples are discussed in each section to illustrate the design and methodological issues, which affect how such siblings are published (if at all) and some of the important concerns discussed by specialists in the respective fields. This discussion helped in understanding how such studies complement RCTs in a broader systematic review or for policy around practical implementation, how they are to help decision making. The discussion also helped when compiling search strategies, and checking any existing filters.

2.2 Systematic Review Overview

Systematic reviews have become one of the most important tools to support evidence based practice and decision making. Searching the existing literature shows that systematic reviews depend heavily on RCTs. Other types of research design contribute useful knowledge for the implementation of an intervention, but their retrieval has not been so well assessed. Harden et al. (2004, p. 794) stated that "Systematic review methodology is well developed for trials, but the debate about systematic approaches to reviewing non-experimental research is in its early stages." Furthermore, Sackett and Rosenberg (1995) emphasise that clinical evidence based practice and other health care decisions should be based on the best 'patient-based' and 'population-based' as well as 'laboratory-based' evidence for better evidence based decisions.

Successful and unbiased identification and retrieval of all of the evidence base relevant to a SR topic depends mainly on the search strategy and to what degree it is possible for the searches to maximise sensitivity[♣], specificity[♥] and precision[♠], (Wilczynski, et al., 2004; McGowan & Sampson, 2005; Wilczynski, Marks, & Haynes, 2007).

Nowadays there is an increased consciousness of the different types of studies that play a role in the decision making process, and these studies are being sought, so that they provide more information and explanation for professional acceptance for certain interventions, to help open a door for more improvements, (Harden et al., 2004).

The Joanna Briggs Institute (JBI, 2008) urges the utilization of other types of research, particularly qualitative research, economic research and policy research when conducting systematic reviews. The emphasis on considering health care, social, cultural and economic

---

♣ *The proportion of high quality studies retrieved on specific topic.*

♥ *The proportion of low quality studies not retrieved.*

♠ *The proportion of retrieved studies that are of high quality.*

factors all together makes for an important pillar for strong evidence, and therefore sibling studies (Section 1.3) are to be sought out for analysis alongside the randomised controlled trial for a comprehensive view and understanding.

Each one of these siblings is concerned about, or discusses one of the main issues separately from the other research associated with one particular project problem. Pawson *et al*. (2004), states that an intervention may be viewed as;

> a complex system where intervention itself is a fragile creature that is delivered in a social system of interacting elements, such as an individual's capacity, interpersonal relationships, institutional setting and infrastructure (p.15).

i.e. political support and funding resources to support the intervention, so evaluation of such interventions require both quantitative and qualitative evaluation (Campbell, 2000).

2.3 Economic and Cost Effectiveness Research

The NHS Centre for Reviews and Dissemination (2009) defined economics studies as:

> An economic evaluation is a study in which both the cost and health outcomes of comparative technologies or interventions have been assessed to identify, measure, value and compare the cost and consequences of the alternative interventions being considered (p. 202).

Any intervention needs to be assessed not only in terms of its effectiveness, but also affordability - if it is economical and cost-effective. Considering economic aspects along with intervention effectiveness can make reviews more useful to health care decision-makers in order to achieve maximum health gain from limited resources. Many studies cover the cost effectiveness of intervention or/and intervention alternatives, as policy decision makers need to identify interventions that are cost effective before considering implementation (NHS CRD, 2009). In this section some studies that analyse economic aspects of interventions are considered as case studies to explore the type of designs used. Examples are taken from the clinical specialties considered for the sibling research, to demonstrate the main issues that this type of study aims to address in this clinical area, and how it has been handled.

Palmer et al. (2004) investigated the hypertensive patients with type 2 diabetes and renal disease aiming to examine the most cost effective point to initiate irbesartan treatment. The data used for this study was obtained from previous research findings (*IRMA-2; Irbesartan in Reduction of Microalbuminuria-2 and IDNT; Irbesartan in Diabetic Nephropathy Trial*). There were probabilities of the disease progression from *microalbuminuria to early overt nephropathy,* these were calculated and driven from *IRMA-2* study data and then were used as an entry threshold into the *IDNT*. (Palmer et al., 2004)

The research focuses on the analysis of the incremental costs of administrating *irbesartan* and the *ESRD* treatments only, where the U.S. Renal Data system, 2001 was the source of the costs of *ESRD* treatment and the Drug Topics Red Book, 2000 provided the information about the costs for *irbesartan.*

A simulation model was used, using a hypothetical data of patient with type 2 diabetes. The simulation model used three treatment models: 1) "Control" treatment in this stage a standard antihypertensive medications was administered targeting a certain point of blood pressure. *2)* "Early irbesartan" treatment: when patients were in the state of microalbuminuria a daily of 300 mg of irbesartan medication was started. And 3) "Late irbesartan" treatment: at this final stage when patients were in the states of microalbuminuria and early overt nephropathy the same treatment model used in the first phase was used, however when the patients reach the state of advanced overt nephropathy a 300 mg of irbesartan was administrated daily .

Sensitivity analyses were performed using *Second-order Monte Carlo simulation*. The analysis focused on: 1) calculating progression probabilities for the *irbesartan* treatment; 2) the entry point of which patients to enter the *IDNT* treatment based on *UAE* level analysis; 3) Life expectancy analyses (the probability of death in microalbuminuria state, overt nephropathy and advanced overt nephropathy) which were made based on different assumption of annual probabilities derived from the UK assessment (Palmer et al., 2004, pp. 1899-1900).

The findings of study shows that with the control arm, that either late or early administration of *irbesartan* resulted in reduction of *ESRD* and life expectancy improved accordingly, which led to overall cost savings. However, comparing the two initiation points revealed that the early use of *irbesartan* was most efficient in terms of both clinical outcomes and economical savings. (Palmer et al., 2004, p.1900).

There is another study by Palmer *et al.* (2008) which is based on the previous study of Palmer *at al* (2004). In order to evaluate the health economic impact of screening for nephropathy (microalbuminuria and overt nephropathy) followed by optimal renoprotective-based antihypertensive therapy in US settings. This example is a sibling study to the previous study (Palmer at al., 2004) and was conducted by Palmer and other authors in 2008. The early study aimed to assess an efficient time point to start the treatment, the later study aimed to assess cost effectiveness of screening procedure and treatment, and it was proved that screening and optimal treatment always improved the clinical outcomes and value for money. Accordingly, these two examples might be useful to provide an overview about the nature of the sibling relationship as both were done by same author and one was based on the previous study results.

In another example, it has been proven that improved blood glucose control is both cost effective and efficient procedure to reduce chances of progression of microvascular disease in patients with type 1 diabetes (The Diabetes Control and Complications Trial Research Group (DCCT) 1996 quoted in Gray et al., 2000). According to this, Gray et al., (2000) conducted a study to estimate the cost effectiveness of conventional versus intensive blood glucose control in patients with type 2 diabetes. This study was conducted alongside the randomised control trial.

The study used data from 5102 of patients aged 25-65 years that were recently diagnosed with type 2 diabetes from 23 hospitals as clinical study centres in UK. Moreover, data about each hospital duration and admission were recorded. In addition, data about non-hospital and outpatient resources use were obtained by a questionnaire which was distributed either to patients during their clinic visits, or posted to patients who did not visit clinics. The questionnaire mainly targeted homes, clinics, practitioners, nurses, dieticians and other specialists over 4 months. Non-compliant patient data were not considered to avoid bias. (Gray et al., 2000, p.1374).

The Gray et al. (2000, p. 1377) economic analysis information was based directly on clinical trial information of a randomised sample to lessen bias and uncertainty. Cost effectiveness was investigated using the sensitivity analysis method. The focus of the analysis was the change in visit pattern and the associated cost of visit and blood glucose test schedules. In general, cost per patient analysis shows no significant differences between the intensive and conventional management in terms of drugs and clinical visits. The study demonstrated some robust findings showing that increased therapy costs of intensive blood glucose control in type 2 diabetes, are largely counterbalanced by significantly reduced costs of complications.

This third study has a different nature form the previous examples, as it examines a surgical procedure for an IVF trial. Strandell, Lindhard and Eckerlund (2005) evaluate the cost-effectiveness of doing salpingectomy before to the first IVF cycle or after failed cycles among hydrosalpinx patients, as this had never been done previously.

Data for this study were obtained from a Scandinavian multicentre trial on salpingectomy prior to IVF. 204 patients were randomly allocated to one of two groups; laparoscopic salpingectomy (116 patients) or to have no intervention before their first cycle (88 patients). Observations about each pregnancy were kept (Strandell, Lindhard & Eckerlund, 2005, p. 3285).

Treatment and intervention cost calculation and analysis included medical costs as drugs costs for patient undergoing IVF and some complications, in addition to costs that are related to the pregnancy, such as costs for spontaneous abortion or delivery, hospital charges considering all visits and complications that require hospitalization were added to total costs, (Strandell, Lindhard & Eckerlund, 2005). The same sample size as the original study was used.

Study findings show that the live birth rates were higher in the intervention group (60.8%) than in the control group (40.9%). The average treatment costs per patient, including surgery and IVF, were higher in the intervention strategy compared to control group. The average cost per patient, including treatment and pregnancy-related costs, in the intervention group was also higher compared to control group. Based on results of salpingectomy, it has been recommended in several countries to undergo salpingectomy prior to IVF since the incremental cost to achieve a higher birth rate using that strategy seems to be reasonable, (Strandell, Lindhard & Eckerlund, 2005, pp. 3286-3289).

Economic evaluation studies have been accused of lacking generalisability and transferability. As is noted in the description of the studies above, several of the costing models use data that relate to health service delivery in one country only. The work of Boulenger et al. (2005) was designed to explore factors that affect transferability and generalisability. The study suggested some factors that it claims to be the main obstacles for transferability and generalisability:

- Variations in epidemiology, prices related to it, and health care resources availability, besides resources utilization will affect transferability. The methodology used in the study and the type of data being used to assess effectiveness as well as its sources, will influence the transferability and generalisability.

- Clinical practice patterns, besides the reporting method of the study information, i.e. study sample, impact on the effectiveness calculation of the results and the costs associated with resources consumptions and statistical analysis.

The second aim was to develop some sort of checklist to assess the reporting level required to assure transferability and being able to assess the level of transferability and generalisability, (Boulenger et al., 2005). The study recommended high quality, explicit reporting in order for the user to be able to judge if a specific study setting is relevant – to judge transferability and generalisability- as the main outcome for the study, and therefore, a clear and full reporting can help the users, and database administrators/indexers to assess the study relevancy more efficiently when a decision needs to be made. (Boulenger et al., 2005).

2.3.1 Economics Evaluation: Summary of Characteristics

The examples discussed above show that economic evaluation studies cover different types of intervention. They may be associated with diagnostic trials, preventive trials, screening, and some of these evaluations are conducted alongside the trial such in case 2 (Gray et al., 2000), while others are conducted after the trial results have been obtained such as case 1 (Palmer et al., 2004). Moreover, economic evaluation studies are being published in prestigious medical journals such as *BMJ, The European Journal of Health Economics* and *International Journal of Clinical Practice*, where these journals are concerned with the evaluation and development of clinical practice and guarantee high-quality, peer reviewed publications to help in the decision making process and improve the quality of clinical implementation and reporting. The economic part of the evaluation is obviously reported separately if done after the trial, but it is possible that some economic evaluation data may be reported within a (mainly) clinical article report on a trial.

These studies indicate that modelling and sensitivity analyses are integral to many economic evaluations. Baseline data on costs may be derived from standard manuals. Some of the studies obtain much data from patient records, retrospectively or prospectively, and others rely on additional questionnaire surveys. As indicated, the economic analysis may be done alongside or after the RCT, and may require data from several sources – there do not appear to be ways of predicting the timing.

Uncertainty has a major influence on decision making, therefore a greater consideration of sensitivity analyses should be investigated more deeply. A NICE committee emphasised the value of sensitivity analysis as a critical part of the economics evaluation process highlighting the value of analysed parameters that can influence the decision as a high level of uncertainty will lead to a negative decision (Andronis, Barton, & Bryan, 2009).

Moreover, uncertainty is heavily associated with the modelling structure that is being employed by the sensitivity analyses i.e. the choice of parameters to investigate, where economists can choose different values of different parameters at different time point in order to justify their model results and resolve the uncertainty (Taylor, 2009).

The other concern about economics studies is generalisability and/or transferability where this is a major debate. According to economists, there are several factors that lead to variation in economic evaluation and analyses such as the severity of the disease, the price unit and the availability of health resources, therefore the need for guidelines to ease and adapt these variations has been established. Drummond et al. (2009) stated that the study starting point, experimental technology, comparators, patient population, the context in which the treatment has been delivered and the price unit should be relevant to decision making request (question(s) or problem) in order to assess the possibility for the evaluation adjustment and hence its transferability according to guidelines. Consequently, checklists and analytical modelling approaches have been advised to assess the existence and pattern of heterogeneity and to assess the presence or absence of core incremental units (costs and effects) that may affect a certain jurisdiction, as attempts to deal with economics evaluation and transferability of cost effectiveness findings.

There are several methods for transferability. A decision analytic model is the most widely used model to address transferability issues, the decision model provides a framework where evidence from a range of resources, i.e. effectiveness data from meta analysis from international trials, resource use data from single observational studies and unit costs from a particular jurisdiction can be pooled and adapted for a specific decision making of interest. Essers et al. (2010) study is an example of such transferability, in their analyses they estimated the cost effectiveness of a certain treatment to be adopted in Netherlands using a UK model as the basis. There was a need for parameter adaptation and to substitute missing data with corresponding data from the UK setting. Cost effectiveness results proved to be transferable between the two settings, despite the above limitations that may be considered to be a source of an expected bias. However the results recommended a more transparent reporting of methods, results and analytical model employed as well as the data about health care resources.

According to NHS CRD (2009) guidance, it is possible to do a meta-analysis of economic evaluations. However the process is not straightforward, meaning help from an experienced health economist would be required due to variations in reporting among economic evaluation studies. The NHS EED database is part of the CRD scope of activities and one of the major databases in UK that focuses on the economic evaluation reports about interventions in order to help decision and policy makers to understand the impact of an intervention and its effectiveness and make the decision accordingly. In addition, AHRQ (Agency for Healthcare Research and Quality) aims to improve the quality, safety, efficiency, and effectiveness of health care services in America by providing research and information for policy and decision makers for more informed decision-making.

2.4 Qualitative Research Studies

This section discusses the major role that qualitative research plays in providing evidence that complements what is gained from other study types, in particular quantitative and/or RCTs and how qualitative research fits alongside systematic reviews. Several examples of qualitative studies were selected to shed light on typical qualitative research objectives, settings and outcomes.

Both qualitative and quantitative approaches under certain circumstances are considered to complement each other. One can say that qualitative can be used either as a preparation stage before using quantitative methods, i.e. refining the research question qualitatively, or as an interpretation method to explain the intervention (quantitative) outcomes (Dixon-Woods et al., 2001).

Therefore, qualitative research is another type of sibling study which uses a different research philosophy to the quantitative trials, concerned primarily with perspectives and attitudes that may affect the interventions' implementation as well as outcomes. These studies have started to receive greater attention over the past ten years, since it helps in understanding and exploring the social and cultural settings that may have a great impact on an intervention. For example an intervention may considered very effective and successful in a small scale implementation but fail or result in degraded outcomes on a larger scale. Qualitative research may help to answer questions about what is really happening, why and how. These are the questions to which qualitative studies can provide answers, as means of enhancing the link between evidence and practice, (Thomas & Harden, 2008). Simply put, Dixon-Woods et al. (2001, p. 126) believe that "Qualitative research has an especially valuable role to play in answering questions that are not easily addressed exclusively by experimental methods."

Furthermore, qualitative and quantitative research can be combined with each other in single study framework, when a mixed research design is used in a single study (Dixon-Woods et al., 2001 & Pope, Royen & Baker, 2002).

There have been many debates about the role of qualitative research in systematic reviews, but the earlier discussions merely focused on outlining the rationale for its inclusion in systematic reviews beside the assessment of quality in qualitative research, rather than on delineating more precisely on the contribution of qualitative approaches to a systematic review, (Dixon-Woods et al., 2001).

Consequently, Lewin, Glenton and Oxman (2009) investigated the contribution of qualitative methods when conducted alongside the randomised controlled trial of complex health interventions. For this purpose they used a random sample of 100 RCTs published between 2001 and 2003 from the Cochrane Effective Practice and Organisation of Care Review Group register. They aimed for the sample to comprise the most recently published trials of complex interventions, assuming that 1) the use of multiple methods has increased recently, 2) RCTs should have been published before associated qualitative studies are published, and finally, 3) the Cochrane Effective Practice and Organisation of Care Review Group register has a greater focus on complex interventions and therefore the possibility of locating more relevant complex interventions would be greater than in other databases such as MEDLINE and EMBASE.

Several approaches were used in order to locate the qualitative studies that are associated with the sampled RCTs. These included PubMed related articles search, searching for studies with the same author(s), citation search searching the reference list of the RCTs, papers citing the RCTs and finally authors were contacted for more information. The results show that only 30 RCTs from the sample had associated qualitative studies, where only 19 were published. Analysing the characteristics of those thirty RCTs revealed that 27 used specific qualitative methods for data collection and analysis while the remaining seven used qualitative data collection methods without further information about the qualitative analysis method being used. Qualitative studies can be conducted at any time. However these results showed that qualitative studies usually preceded RCTs, (see Figure 1).

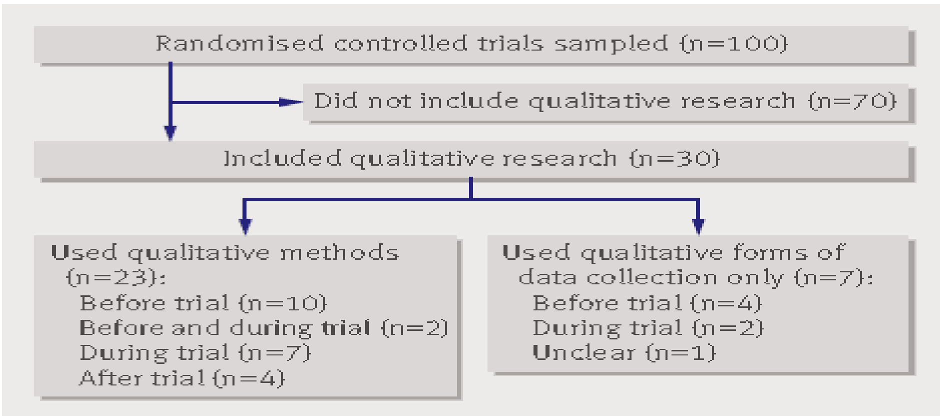


Figure 1: Retrieval Process of Qualitative Studies of RCTs. (Source: Lewin et al., 2009)

The time to carry out the qualitative research depends mainly on the objectives of the research as at each time point it can provide different results:

- Before the RCT:
  - Helps to uncover the contextual and research issues in addition to the surrounding circumstances prior to the RCT.
  - Accordingly, it can help in organising the RCT settings, generate hypotheses and select the appropriate RCT measurements.

- Alongside the RCT:
  - Explore the RCT implementation process by providing a description of the delivery process and whether the delivery criteria were met and if possible changes could be made.

- Investigate and assess attitudes and perspectives of people involved in the RCT.

- After the RCT:
  - Explain the RCT findings and variations in effectiveness of the RCT within the sample.
  - Investigate the RCT theory and procedures appropriateness and generate more hypotheses if necessary.

Methodological heterogeneity between the identified studies was noticed. Some did not specify a specific approach, some mentioned the approach used such as grounded theory, some studies reported the use of several qualitative methods while others used only one approach. Moreover, the sampling method and data analysis used were poorly reported, and therefore, quality assessment and integration between qualitative studies and RCTs could not be fully realised. It was concluded that the relationship between the RCT and its associated qualitative study(s) is not a straightforward one to detect, although the researchers found that half of the qualitative studies (16) had shared authors with the RCTs, while some of the remaining studies (9) specified an explicit association or link (by stating the RCT it is associated with) between the qualitative study and the RCT.

The authors note that concurrent qualitative studies were less frequent than might be thought. Lewin et al., (2009) suggest that this is due to the evaluation of the intervention using 'linear models', in such cases the need for the qualitative approach is being marginalised to explore external factors since in this model the RCT runs through multiple evaluation phases in order to ensure the trial's effectiveness. Additionally, the study authors, when contacted, mentioned that researchers encountered restrictions on resources and poor access to relevant expertise. In addition, many of the qualitative studies were conducted prior to the RCT (and did not seem to be of high quality).

RCTs with an explicit theoretical basis identified seem more likely to have qualitative studies associated with them. For example, a systematic review of health promotion interventions for preventing loneliness among older people comments on whether the intervention used a theoretical framework. Although this review only included the high quality quantitative studies in the analysis, the review notes the accompanying 12 qualitative studies, and in fact comments that the effective interventions often included some form of process evaluation (Cattan et al., 2005).

In another study by Glenton, Lewin and Scheel (2011) the importance of qualitative research, when conducted alongside trials to explore factors and processes that might have influenced intervention outcomes, has been established. They attempted to find any qualitative research associated with 82 randomised trials used in their study on lay health workers. To locate as much as possible of the qualitative research (if available) they contacted authors of the trials, checked reference lists, searched PubMed using related articles search for related studies and other studies by same authors. Glenton et al. (2011) found that more than half of the trials (63%) had no qualitative research linked to them. 12% of the studies (trials) had qualitative data collection referred to in the trial paper or in authors' e-mails (when contacted) but was unavailable. Only 7% of studies have qualitative research carried out before the trial and this was available as either published or unpublished reports. 17% of the studies have qualitative research carried out before or after the trial and were available in the same paper as the trial, or as a separate paper. In conclusion, they suggested that qualitative research is still less common than expected and the associations between qualitative studies and trials that are reported separately are still unclear and need to be explicitly linked to one another to facilitate retrieval. However, there findings are limited to lay health worker programmes for maternal and child health and infectious disease control intervention delivery. But it can help to explain some sort of retrieval patterns for other clinical areas.

Much effort has been made trying to improve public services or tackling social issues, and taking experiences of patients and carers and their views into account may have a great impact in improving the quality of research and effectiveness of interventions. NICE, for example, has a Citizen’s Council that aims to ensure that the views of public are input into the decision-making processes. Patients with specific conditions may also be represented through patient support groups, as such groups may become registered “stakeholders” in the development of guidelines for that condition.

Furthermore, the previous section (Section 2.3) examined some of the approaches to economic evaluation of diabetes screening and treatment. Diabetes requires a specific treatment and adjustments in life style in order to achieve the required results. Adherence to medication and life style modification are the main factors that are associated with successful treatment. However, poor adherence is considered to be a serious problem in public health that reduces the benefits of current medical care. Quantitative studies have often failed to verify factors which play a major role in explaining adherence or non-adherence to treatment recommendations with diabetes type 2 patients. Therefore, qualitative studies may be used to help to identify some issues that might explain such behaviours. This is important as the economic evaluations are often based on particular modes of treatment being followed.

For these purposes, Vermeire et al. (2007) tried to investigate the barriers to adherence in living with type-2 diabetes using a focus group approach and this may give a strong indication about the importance of qualitative assessment of interventions, as this was a large international study.

To assess health attitude and barrier to treatment recommendation adherence, Vermeire et al. (2007) conducted focus group interviews, using the same set of questions, in seven European countries. Sample information was included in the study report. 'Grounded theory' was used to analyse the data in order to derive primary themes provided in native languages, and later the researchers from each country presented their primary data to all other researchers.

There were several difficulties when the researchers attempted to combine their data: mainly cultural and language differences and the problem of losing some value in content, due to the problems of translating the codes and themes into one common language. Meta-ethnography was deemed the most appropriate method to capture common key themes and ensure transferability between countries. One of the strong points about meta-ethnography techniques is the ability to make a comparative analysis of qualitative findings from different situations (Vermeire et al., 2007, p. 27).

Vermeire et al. (2007, pp. 28-29) used inductive analysis of data generated from research studies to detect similarities and differences between studies, development of hypotheses and hypotheses testing. This process required recursive analysis of the new generated themes and the original data to ensure that the results can go beyond the original studies to explain broader phenomena. Themes derived from the first study (Belgium study) were considered as the standard to compare other studies' themes against. To ensure data transferability and coherence between the original data and the translated ones two additional interpretation phases were conducted.

The study, explained the adherence and non-adherence behaviour of diabetes patients, where the data collection and data analysis were found to be sufficient to provide robust data analysis and findings. The sample size was adequate for qualitative research and so the researchers were confident with their results. Accordingly, the qualitative study succeeds in uncovering the problems which the quantitative studies failed to handle.

The next example discusses research on health care providers, rather than patients. Improving health care quality often needs qualitative studies to assess attitudes, experiences and behaviours of professionals within organizations and healthcare teams as these factors have a great impact on health care improvement. Qualitative assessment of service quality requires an understanding of the context, environment and the manner in which the treatment is delivered. Successful recognition of patients, providers, politicians and the public serves in identifying what is considered as a good quality care.

Accordingly, Pope, Royen and Baker (2002) investigated the factors that might improve health care quality, using qualitative methods to help explain variations in health care. This study was based on interviews and observation (Pope et al., 2002, p. 148).

The sample size for interview was between 30 – 50 respondents (a typical figure for qualitative studies, unlike RCTs with thousands of participants). Often, for observational studies data, this might well be based on a single study focusing on one organizational setting. This reflects the nature of qualitative studies as they tend to be exploratory in approach and so the sampling strategies used in qualitative research are "purposive or theoretical rather than representative or probability based" (Pope et al., 2002, p. 149). Coding analysis was conducted focusing on the words, context and experiences of participant to generate hypotheses and theoretical framework (Pope et al., 2002, p. 150).

Pope et al. (2002, p. 151) findings proved that qualitative research methods helped in identifying what really matters to patients and care providers. For example, it showed the cultural and social factors that obstruct or encourage service use. Furthermore, it helped to explain some behaviours, and reasons for resistance to change. Moreover, qualitative methods can highlight the process of policy implementation, uncovering the causes of failure or success and suggest solutions accordingly. In the end, Pope et al. (2002) recommended that qualitative methods can complement other research. The qualitative research can help to illuminate different facets and inform quantitative approaches to researching health care.

An argument put forward by Dixon-Woods, Fitzpatrick and Roberts (2001) about the crucial importance of both qualitative and quantitative research to address reservations in many health care areas. They call attention to the role of this approach as a preliminary to quantitative research, during or after the trial, aiming to provide explanations to the process and outcomes and as a means of enhancing the link between evidence and practice. Moreover, in the current state of affairs, a systematic review may seem inattentive to patients' perspectives and views in their conclusions, discussion or implications for future work which might help in creating a robust and focused systematic review questions, (Dixon-Woods et al., 2001, pp. 125-127).

Understandably, there have been changes since 2001, and the 16th Cochrane Colloquium (2008) focused on the consumer perspective, and the application of non-RCT designs. For example, Berkman and Viwanathan (2008) discussed the development of a tool to evaluate the quality of observational studies, and Glenton et al. (2008) discussed the benefits of combining a review of trials with a review of qualitative studies and how this can provide more understanding of the intervention's effectiveness. In the same context, Hansen (2008) presented a lecture discussing qualitative research methods and highlighted patients' perspectives, the value of these and the influence on health technology assessment.

There are other ways of using qualitative research findings. In one example, qualitative methods were used to investigate the effect of psychological support to mothers in first time labour, and although the results of qualitative methods did not contribute to the quantitative synthesis, Bayesian statistics can help in synthesising both quantitative and qualitative to support and strengthen the outcomes of the quantitative trial (Dixon-Woods et al., 2001). In conclusion, qualitative research is now generally considered as a good or at least useful evidence resource in systematic reviews.

However, qualitative research may still be marginalised by some, despite the rich contributions it offers to systematic review. Debates about what could be considered as evidence can be fiery, and there is resistance from quantitative researchers in accepting qualitative results on one hand, while on the other hand qualitative researchers mistrust quantitative methods.

Yet, techniques for including and synthesizing qualitative evidence remain under-developed compared to synthesis of quantitative evidence. NHS Centre for Reviews & Dissemination became aware of this problem and recommendations about developing formal methods for qualitative methods and quantitative methods have been made. Moreover, locating qualitative research may still be difficult and frustrating, although the effort for developing search filters for qualitative research were noted, (Dixon-Woods et al., 2001, p. 130) (and work has continued since then).

Qualitative synthesis is considered the most challenging issue when qualitative studies are to take a part in systematic review creation. Gough (2007) said:

> We can use what we know from different sorts of knowledge collected and interpreted in different ways to develop theories, test theories, and make statements about (socially constructed) facts (…) and the challenge is to develop a language to represent this different sort of knowledge to enable debate at the level of synthesis of knowledge rather than at the level of individual studies (p. 3-4).

Data synthesis range from ‘statistical meta-analyses’ to ‘narrative synthesis’, and methods include meta-ethnography, depending on the nature of review questions under investigation. Of course, this usually depends on quality and relevance assessment to decide to what extent a study findings answer the review questions (usefulness of the contribution to meta-synthesis) and assign a weight accordingly, (Gough, 2007 & Dixon-Woods et al., 2001).

According to Zimmer (2006), meta-synthesis methodologies varied according to the objective the analysis aims to achieve, the focus of the research and the nature of the data to be analysed. For example *ethnography* examines the cultural context and sittings in order to generate an explanatory theory to describe specific phenomena while *Grounded theory*, creates a theory based on the original source data.

Another argument was raised about the possibility of losing the epistemological commitments and core intent during meta-synthesis. Zimmer (2006, p.316) stated that for efficient and coherent synthesis, language, context, time, contradiction, and epistemological intent of the primary study should be carefully explored. Moreover, participants' attitudes; the researchers' situational issues; and how to bring the previous interpretation all together into one synthesised final summary should be deeply investigated.

In the same context, Thomas and Harden (2008) recognised the increased and valuable contribution of qualitative research in health care evidence and systematic reviews. This study was intended to develop methods to bring together and integrate qualitative findings from multiple studies within the systematic review, using thematic synthesis.

In order to undertake thematic synthesis, a search was conducted to locate all relevant studies which will add a value to the systematic review. Doyle (2003, cited by Thomas and Harden (2008, p. 3)) emphasise that:

> Like meta-analysis, meta-ethnography utilizes multiple empirical studies but, unlike meta-analysis, the sample is purposive rather than exhaustive because the purpose is interpretive explanation and not prediction.

This means that conceptual synthesis depends mainly on ‘conceptual saturation’ and retrieving heterogeneous studies in contrast to meta-analysis.

Since qualitative researches are difficult to locate depending on electronic searches only, hand searching was used for searching grey literature to locate the studies (Thomas & Harden, 2008, p. 4).

Furthermore, Thomas and Harden (2008, p. 4) urge that extracting data from a qualitative study or determining the key concepts is not an easy task to perform. This is due to the variation in representing and reporting the data. Moreover, data itself are different from published reports. Data are empirical based on participant point of view, observations or experiences while the findings represented here are based on derivation method being used, external data sources and researchers’ conclusions and implications.

Thomas and Harden (2008, p. 8) checked that themes can be translated from one situation to another, integrating facilitators and barrier of such translation with different study contexts. The study’s ability to answer the research question was prioritized over study design which was the basis for including and excluding criteria. Sensitivity analysis revealed that poor quality studies, compared to high quality studies, made moderately little contribution in the synthesis where unique and developed themes were not found.

Recommendations were made by NHS CRD (2009) to use qualitative approaches to help in evaluating intervention implementation process in depth and its outcomes:

> qualitative studies are an effective tool which help in understanding the mechanisms behind effectiveness or ineffectiveness, understanding heterogeneous results, identifying factors that impact on the implementation of an intervention, describing the experience of people receiving the interventions, and providing participants’ subjective evaluations of outcomes (p. 221).

2.4.1 Qualitative Studies: Summary of Characteristics

So far the importance and the major support qualitative studies provide for systematic reviews robustness and policy/decision making process has been established. Qualitative approaches used in this type of studies focus more on the context and circumstances surrounding the intervention besides exploring the perspectives of the receivers of this intervention, providing a deeper explanation of the intervention validity. Interviews, questionnaires and focus groups were the main approaches used in such studies for data gathering, followed by qualitative data analysis.

Qualitative synthesis was of importance as the qualitative studies themselves are not as easy to integrate as the results of randomised controlled trials. Several approaches with their particular characteristics have been explored, noting the consensus about the challenging nature of meta-synthesis.

Nevertheless, there are some debates about the inclusion of qualitative studies in systematic reviews and making decisions based on recommendations from qualitative research. The criticism that qualitative studies receive according to (Pope, van Royen & Baker, 2002 & Mays and Pope, 1995) is that the research is strongly subject to researcher bias; qualitative studies are difficult to replicate; they lack generalisability beside the fact that qualitative methods tend to describe a small number of settings with a huge amount of textual detail, but there are often inadequacies in the reporting methods.

The other view tends to be supporting, and Pope, van Royen and Baker (2002, p. 150) argue that qualitative research can provide data with good internal validity, and can explain and explore phenomena more precisely. However, reliability is still complicated, as well as questions of transferability and generalisability. Moreover, there are still some arguments about the appropriateness of combining qualitative studies with different qualitative approaches based on different assumptions and theory, (CRD, 2009).

The example discussed above (Section 2.4) demonstrates some characteristics and nature of qualitative research. Of interest for retrieving sibling studies is the observation that study reporting may be inconsistent, which makes searching and locating qualitative research more complicated.

One of the examples (Lewin, Glenton and Oxman, 2009) discussed above discussed some search strategies used to locate and identify all published and unpublished qualitative studies about RCTs. They used citations from the RCTs themselves and other citation indexes for the papers citing the RCTs, RCT authors' other publications, and PubMed related articles to identify qualitative studies. This supports the choice of search strategies to be examined in this research (Section 4.3). As might be expected, there is little consensus on search filters for locating qualitative research, as the CRD list indicates (http://www.york.ac.uk/inst/crd/intertasc/qualitat.htm).

2.5 Assessment Studies / Process Evaluation

Reviewing intervention evaluation reports shows that some interventions fail while others succeed, and process evaluation can help to distinguish between interventions which are based on a faulty theory base from those which are merely poorly delivered (Oakley et al., 2006). Accordingly, Hulscher, Laurant & Grol (2003) emphasise that process evaluation is an important means of describing the intervention implementation, settings, and intervention target (recipients) as well. This type of research aims to explore and investigate the mechanism and processes, assess the result and detect variations and reasons for variations, together with implications. Thus, process evaluation does not only mean measuring the effectiveness, it also means understanding the 'workability' and integration of interventions in dynamic and complex settings. Moreover, process evaluation inherits some of the characteristics of qualitative research and can be considered accordingly a special type of qualitative study.

RCTs may be considered as experiments aimed to enhance lifestyle and/or decrease mortality. However they have a social part as many interventions have an impact on lifestyle and the treatment process may require a group of professionals to work together. Therefore the entire intervention may be composed of several interactional parts that may affect each other and affect the intervention outcomes in the end. Several studies such as Oakley et al. (2006) and May et al. (2007) emphasise the importance of process evaluation as an integral part of any RCT implementation, specially for complex interventions - *health service interventions that are not drugs or surgical procedures, but have many potential 'active ingredients'* – because of their nature. According to them, interventions are meant to be implemented in social contexts which may well interact with the intervention and may affect the intervention course and outcomes. In conclusion, the social environment is a complex system of elements which interact with each other in different, undetectable patterns; consequently, the environment will change or affect any experiment outcomes.

May (2006) defined a model called Normalization process model for process evaluation which helps in detecting the factors that affect intervention implementation. Greenhalgh et al., (2004, cited in May 2006, p. 2) stated that understanding implementation potentials of a new practice required understanding the 'whole system' under which the practice was implemented. This idea emerged from studies that have aimed to explore an organisations' capacity to implement new systems of practice and build theories and models accordingly; to facilitate understanding the intervention's outcomes and associated behaviours of staff and patients.

The model has two stages, the first stage, *formative analyses of qualitative data* (Italic in original), based on analysing four groups of qualitative studies. The analysis focused on 1) exploring the relationship between professionals and patients in terms of their social organisational framework; 2) investigating implementation and delivery of new modalities in health care; 3) "the social construction and production of evidence" which analyses data from several/different settings, comparing and contrasting specific theoretical interpretations of data items from different parties; and finally 4) the changing organization of clinical work around chronic illness in primary care, (May, 2006, p. 3).

Building a higher level model (Italic in original) was the second stage, to develop a general set of propositions based on the formative analysis results from the first stage. It consists of four interpretive theory building activities: 1) identification of components: extracting the 'core' component from the set of result from formative analyses; 2) retention and rejection of components, where further analysis for the cores was conducted (the retention criteria concern validity, to reflect a generalised representation of interaction processes and clinical practice); 3) building a set of propositions and constructs to be evaluated against a known outcomes of specific services to ensure validity; and 4) and finally, an informal validation of the constructs and propositions by circulating them to an informal reference group to ensure the propositions had face validity for other researchers in the field, and that they were practically applicable into specific research contexts, (May, 2006, p. 4).

This model was interested in strategies that are able to make an intervention effective and integrate it in practice using constructs and propositions as a scoring system to assess the interventions, and four constructs and propositions construct the model, (May 2006, pp. 5 - 8):

- *Interactional workability* (Italic in original): the first constructs are concerned with the conditions and situations under which the intervention is operationalised. It is characterised by two dimensions; *Congruence* (Italic in original)*;* the order of interactions, co-operation, role and conduct of agents in a complex intervention and *Disposal;* how the interaction between agents affect the complex intervention (outcome patterns).
  P_1 → *"A complex intervention is disposed to normalization if it confers an interactional advantage in flexibly accomplishing congruence and disposal of work."* (Italic in original).

- *Relational integration* (Italic in original): refers to how clinical and social relationships have been established, and how the complex intervention might influence such relationships. It is characterised by two dimensions: *Accountability* (Italic in original)*;* refers to the internal reliability of knowledge and practice that an agent has and *Confidence* (Italic in original)*;* refers to the external reliability of knowledge, practice, and technologies which contributes to complex intervention delivery (understanding each other actions).
  P_2 → *"A complex intervention is disposed to normalization if it equals or improves accountability and confidence within networks."* (Italic in original).

- *Skill-set workability* (Italic in original): this refers to how complex intervention knowledge and tasks are distributed among a team of workers in health care settings. It is characterised by two dimensions: *Allocation* (Italic in original)*;* how the intervention related tasks and activities are allocated and *Performance* (Italic in original); how effective the agent handles and delivers the assigned activities.

  P_3 → "*A complex intervention is disposed to normalization if is calibrated to an agreed skill-set at a recognizable location in the division of labour.*" (Italic in original).

- *Contextual integration* (Italic in original): this final construct refers to how the organisation understand the implementation of a complex intervention and if the existing resources can accommodate the new modalities and to what capacity.

  P_4 → *"A complex intervention is disposed to normalization if it confers an advantage on an organization in flexibly executing and realizing work."* (Italic in original).

These propositions disclose that interventions are a complex system of integrated components that interact with each other and affect intervention implementation in all aspects. It provides a conceptual framework for understanding the processes in which complex interventions become embedded in practice, and thus sets out a rational framework for complex intervention evaluation. The focus of the model was the interactions within and between processes of practice.

There are different definitions of complex intervention. New guidance from the Medical Research Council (Craig et al., 2008) updates their 2000 draft guidance. They point out (Craig et al., 2008, p.7) that:

> Complex interventions are usually described as interventions that contain several interacting components. There are, however, several dimensions of complexity: it may be to do with the range of possible outcomes, or their variability in the target population, rather than with the number of elements in the intervention package itself. It follows that there is no sharp boundary between simple and complex interventions. Few interventions are truly simple, but there is a wide range of complexity.

May et al. (2007) applied that developed theoretical model (Normalized Process Model) to understand and evaluate the implementation of complex interventions and to explore the capability of the model to evaluate different intervention with different sittings.

The model was applied to two different complex interventions; ‘the delivery of problem solving therapies for psychosocial distress’ (psychosocial intervention), where the organisation and activity allocation had to be made, and ‘the delivery of nurse-led clinics for heart failure treatment in primary care’ (organic disease intervention), where activity allocation and work structure needed to be altered as the work will be assigned mainly to nurses rather than physicians.

May et al. (2007) believes that process evaluations need to be concerned deeply and strongly with their relative workability, embedding and integration. However, the model was limited and did not describe how the complex intervention was formed and how participants have been chosen, which called for refining the existing model resulting in normalisation process theory which has a wider scope compared to its predecessor.

Theory components are coherence, cognitive participation, collective actions and reflexive monitoring. Coherence emphasises that intervention is an ensemble of beliefs, behaviours and cognition of participants that define and organize the objects of intervention. Moreover, the meaning of an intervention is formed by participants’ own apprehension and continuous contribution. Cognitive participation which means that intervention is shaped by enrolments, engagements and investments of human actors, which position them for the interactional and material work of collective action of reshaped behaviours and action and reorganised relationships and contexts in goal oriented manner. This component employs all four propositions recommended by the normalisation process model: interactional workability; relational integration; skill-set workability; and contextual integration for organising and enacting an intervention with continuous efforts. Finally, reflexive monitoring emphasises everyday and collective understanding of an intervention, it involves continuous evaluations and judgements about utility and effectiveness of an intervention. From this it can be concluded that the theory provides a comprehensive analysis about the implementation of the intervention, as it focuses on production and reproduction of the implementation, embedding (or not), and continuing integration of material practice (May, 2009).

In another simpler example of process evaluation, the work of Flottorp, Håvelsrud and Oxman (2003) discussed the importance of providing supplementary information to describ the trial procedures and conditions, also to generate hypotheses describing why an intervention was successful or failed to change practice. Process evaluation and the trial were conducted simultaneously to explore factors that may explain why the outcomes varied from those expected. The RCT aimed "to assess the effectiveness of tailored interventions to support the implementation of guidelines for the management of urinary tract infections in women, and sore throat", (Flottorp et al., 2003).

Flottorp et al. (2003, p. 334 - 335) used four qualitative data collection methods to collect the required information about the trial process and surrounding circumstance.

Finally, the results suggested that the inadequate communication and time within the practice might be the most obvious obstacles that affect the outcomes, and practice acceptance of change. Practices that do not have routines for discussing guidelines and managing change for common problems find it difficult to integrate such projects into already existing system. Time and sufficient support, were found to be necessary in order to achieve the required change in practice or implementing guidelines. (Flottorp et al., 2003, p. 338).

There were a variety of quality improvement interventions that have helped in health care promotion; some of these interventions were successful while the others failed to achieve their purpose. Process evaluation can shed a light on the mechanisms and processes for intervention development and which have an influence on the results. Based on these facts, Hulscher, Laurant and Grol (2003) meant to explore the value and the purpose of process evaluation on quality improvement interventions besides addressing the issue about what to measure and how to measure.

According to Hulscher et al., (2003, p. 41), process evaluation can be applied at any phase of quality improvement intervention. Therefore, process evaluation can be applied to pilot studies or small scale improvement projects where the purpose here is to explore the possible changes, feasibility and applicability of conducting the intervention.

Additionally, process evaluation in controlled quality improvement studies is used mainly to investigate whether the implementation method is valid and applicable in a controlled environment. In this type of process evaluation, the study helps to explore the causes of different outcomes, following a standard implementation plan that will detect the source of the problem that may be responsible for the failure as well as clarifying participants' roles in success or failure of the intervention. (Hulscher et al., 2003, p. 41).

Process evaluation for large scale quality improvement programme effectiveness can focus on intervention goals and if these goals are realized and to what extent. It provides information about the actual intervention, situational and contextual and experience with the intervention, (Hulscher et al., 2003, p. 41).

Hulscher et al. (2003, p. 41 - 42) stated that a decision about what data to measure and how to measure these data should be made when process evaluation was decided on. Regarding what to measure, researchers need to decide what are the 'key features' of the intervention that are required to be included in the process description such as features to support uniform performance and participant exposure to the intervention.

Deciding on a suitable measurement method to use depends on the research question and nature of the process, so either a qualitative approach or quantitative approach can be used to gather the data. The data collection method has to take existing circumstances, as practical issues, the homogeneity of the data and privacy and confidentiality into account. Measurement methods should be simple, user friendly, however detailed comprehensively. And finally, a representative and valid population sample should be investigated by a qualified person, (Hulscher et al., 2003, pp. 43-44).

Information about the intervention can be gathered from documentations of the study plan, the programme proposal and minutes of meetings. Description about intervention implementation sequence can be gathered by interviewing implementers and participants. However, data reliability was an issue as it decreases as intervention complexity increases, and for long interventions the framework contains many features that the respondents might not be aware of during the implementation. Thus it has been recommended to gather information during the process implementation, (Hulscher et al., 2003, p. 44).

People's experience plays a great role in explaining the feature that may be considered as the one influencing intervention outcomes. During intervention it is useful to focus on barriers and facilitators so it will be easy and feasible to revise the intervention (Hulscher et al., 2003, p. 44).

2.5.1 Process Evaluation: Summary of Characteristics

As explained earlier, process evaluation can help to build a distinction between interventions which are based on a faulty theory base from those which are badly delivered. As aforementioned, an intervention is composed of several parts those interact with each other and may affect each other, and which can affect the intervention outcomes in the end. Thus process evaluation aimed to explore the contextual and environmental issues which may influence the intervention outcomes. If we look at the nature of these studies we can sense a resemblance between them and qualitative studies as process evaluation is based on qualitative data collection methods, as can seen in Flottorp, et al. (2003) study. Of course data collection method selection depends heavily on the circumstances under which the intervention had been implemented, also the components it aims to examine. The May normalization model for example focus on a set of components which are qualitative by nature and cannot be measured another way. "An evaluation model that asks what people do to make a complex intervention workable and to integrate it in practice." (May 2007, p. 2).

Both the Normalisation process model and Normalisation process theory can help by identifying possible barriers to implementation of new services, thereby allowing implementers to focus efforts on addressing areas likely to be particularly problematic. But the theory puts more weight on people's cognitions, involvement in, and appreciation of, the intervention which may affect the way they perceive how the intervention changes and why, therefore, one of the components of the theory was collective action and reflexive monitoring that emphasise continuous investments and assessments of implementation and practices, (May, 2009 & Morrison & Mair, 2011).

Process evaluation values and purposes were detailed in Hulscher et al. (2003, p. 40) as the following:

- 'Intervention as planned' description, this information is useful to help practices to understand how to adopt and implement new changes into the target population.

- Actual exposure, this information will clarify the implementation details which will reflect on the causes of intervention success or failure, so the intervention can be revised accordingly.

- 'Intervention as performed', these details make the intervention replicable in future and helps compare studies and *meta* analysis of crucial features of effective interventions.

- Perspectives and experiences of people who are involved in the intervention help in revising research question(s) and may provide explanation about influencing factors, which as result help in improving the intervention or suggesting future work and recommendations when the intervention is to be replicated.

Process evaluation can be applied to any quality improvement intervention at any stage. Process evaluation studies are used mainly to investigate the effectiveness of the implementation method under standardized circumstances. Data required about the intervention itself can be gathered from implementers and intervention related reports, from participants and receivers of the interventions. Furthermore, the evaluation employs qualitative evaluation methods and can be done alongside the intervention or after the intervention is finished.

Strategies for finding process evaluation studies could simply focus on "process evaluation", but review of the examples illustrates that themes such as feasibility, workability, practical implementation, complex intervention, skill substitution/skill sets, critical success factors, barriers and enablers also need to be considered, where qualitative research methods were mainly employed for either data collection or analysis. MeSH uses "Process assessment" as a MeSH subject term which can be used to searching for process evaluation study, however the retrieval performance in such cases depends on authors' descriptions and reporting of process studies. Exploring the Hedges team work (Section 3.4.2), it appears that process evaluation research did not receive as much attention as the qualitative and economics research.

2.6 Summary

Any intervention should be assessed not only in term of its effectiveness, but also its cost-effectiveness. Qualitative research helps to explain what is happening and how and therefore it can provide answers that clarify the link between evidence and practice. Identifying barriers to performance changes can play an important role in healthcare quality improvements. Qualitative research studies are indexed differently in different databases with no standard terms to identify these studies, which make the identification and retrieving of qualitative studies complex. Moreover, intervention experiments and/or implementation reports can explain why some interventions fail while others succeed. Process evaluation can describe the intervention implementation, settings, and intervention target (recipients), as well as exploring and investigating the mechanism and processes, assessing the results and detecting variations and their reasons.

There are debates about the ways in which qualitative, economic and process evaluation research should be integrated, or set alongside the results from RCTs. Different researchers place different stress on the theoretical approaches to be used (different methods of meta-synthesis for qualitative research, normalisation process models or simpler approaches to programme evaluation). Clearly the work involved in some process evaluations is immense, and perhaps it is to be expected that process evaluation in many trials focuses on the questions of satisfaction and acceptability (i.e. falling under qualitative research, and considered outside a large scale process evaluation framework).

Chapter Three
The Context of Information Retrieval and Trends in IR Research

3.1 Introduction

The previous chapter introduced three types of studies with research design and characteristics different to RCTs, but which can complement RCTs. Qualitative or economic or process evaluation studies done prior, alongside or after – and associated with an RCT in some way, may be understood as direct siblings for that RCT. Understanding each type of sibling study's properties and purpose will help when searching for these research designs.

The contribution that this research is trying to make is to identify the sibling studies for a particular set of studies, by recognizing the linking characteristics between them. Prior to that identification stage, different searches have to be used to retrieve a pool of possible relevant items, and then comparing the effectiveness of each strategy alongside the databases being searched. This also involves deciding on an effective measurement to measure retrieval performance and efficiency (Section 3.6).

This chapter aims to explore and investigate information retrieval issues and trends, as there have been recent changes with the advent of Google (and Google Scholar), SCOPUS and federated search (Meta-lib/e-library searching across several databases simultaneously). Most systematic review search research has focused on strategies for particular databases and consequently this chapter will introduce a discussion of the efforts made to design search filters, with emphasis on filters for particular types of research study, including those of the Hedges team research (Section 3.5.4). It is easier to design useful search strategies for users if users' needs (and possible associated contribution of particular siblings) for a review process are appreciated. Retrieval and management of references for a systematic review is a time consuming process, and other approaches have been proposed to assist systematic reviewers, and information professionals supporting policymakers. Text mining techniques appear to be a promising approach that can assist systematic review process creation, (Ananiadou et al., 2007) (Section 3.7). There are also different, non-Boolean subject searching approaches to finding relevant studies (Sections 3,5,2 & 4.3) as "related articles" in PubMed and "Citation reference". More novel approaches may get around the problem identified in Chapter 2 of qualitative research being hard to find due to inconsistent or incomplete reporting.

The structure of the chapter is as follows. Section 3.2 introduces an overview about information retrieval organisation and document representation based on user needs. Section 3.3 discusses information retrieval characterisations. In this section matching procedures (section 3.3.1), information retrieval trends (section 3.3.2) and information retrieval role in systematic reviews (section 3.3.3) will be addressed. In section 3.4 digital libraries will be investigated and federated search will be introduced and discussed. Search strategies and approaches for evidence based practice information retrieval will be introduced in section 3.5. In this section, the PubMed most featured functions and procedures, related articles and ATM (automatic term mapping), will be addressed (sections 3.5.2 & 3.5.3 respectively) Then, search filters will be discussed (section 3.5.4). Retrieval of qualitative research will be presented in section 3.5.5. Clinical trial registers will be discussed in sections 3.5.6. In section 3.6 retrieval effectiveness measurements, recall and precision, relevancy and databases performance (sections 3.6.1; 3.6.2 & 3.6.3 respectively), are addressed. Section 3.7 will introduce text mining as a new trend in information retrieval aims to help in systematic review development. And finally section 3.8 will provide a summary of this chapter.

3.2 Information Retrieval

3.2.1 What is an Information Retrieval System?

Manning, Raghavan and Schütze (2009) defined Information retrieval as:

> Information retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers) (p. 1).

While Liddy (2005) defined it:

> Information Retrieval as the computerized process of producing a list of documents that are relevant to an inquirer's request by comparing the user's request to an automatically produced index of the textual content of documents in the system (p. 1).

Both definitions provide an adequate description of the retrieval process, as they capture and highlight all factors that involve information retrieval process, but the former can be considered a simpler, and less limiting definition than the latter one which refers to the presence of an automatically produced index. The retrieval process initiates from end-user interest in acquiring certain knowledge (information need) from source of knowledge (document), leading to the search process in an attempt to locate the appropriate knowledge source from a wide range of available sources.

Therefore, information retrieval systems are a part of computer applications that aim to meet end users' satisfaction by providing the required information that matches their initial query via, in most cases, delivery of documents that contain some pertinent information. In addition, an information retrieval system can be considered as an interactive medium that links the users to their required information using specific methods and algorithms to meet that goal. Thus, the main components of any information retrieval system will be the system's users, indexing process and matching and retrieval algorithm (Ingwersen, 1992). The flow chart in Figure 2 demonstrates the entire information retrieval process (Heimstra, 2009):

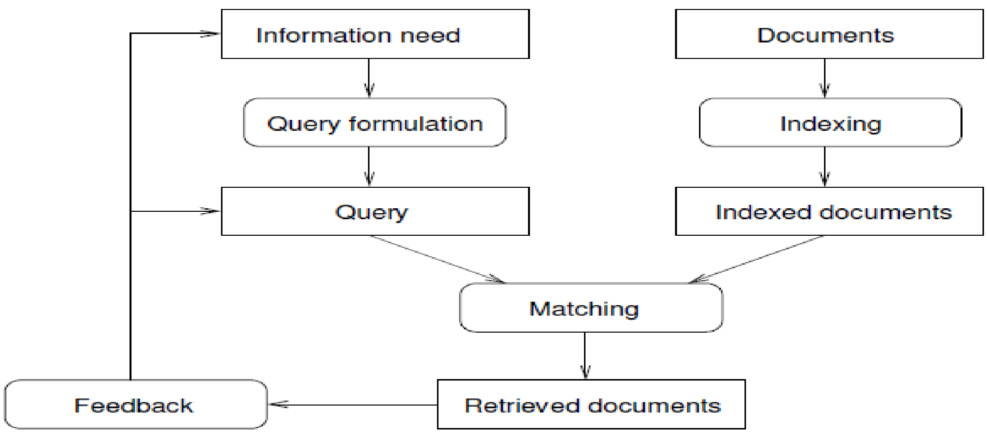


Figure 2: Information Retrieval Process (Source: Heimstra, 2009)

3.2.2 Information Retrieval: Fom Childhood to Maturation

According to Neufeld and Cornog (1986) and Kagolovsky and Moehr (2003) information retrieval began to receive increased attention in the 1950s due to two major factors. First, a large amount of documents in various formats were released after World War Two and became available to the research community, which demanded a process for organising the documents for more efficient retrieval, with use of indexing. Second, the appearance of computers gave some the idea of using computers to help organise, index and retrieve documents. It was a difficult task at that stage to accomplish as computers were slow, and required careful programming to ensure tasks could be completed, but once set up, the advantages of doing repetitive, "batch processing" type of tasks were apparent (Singhal, 2001).

The 1960s was the time when information retrieval as a concept started to emerge and develop (Neufeld & Cornog, 1986; Lesk, 1996; Singhal, 2001). At this stage computers were perceived to have better capabilities than humans in processing of documents in terms of consistent quality and quantity. Bibliographic databases started to appear in the mid-1960s with information being stored on magnetic tapes, i.e. MEDLARS database - the first medical database – and context-based access control (CBAC). While this was considered a great development, the search was still batch searching which was complicated, not flexible and non-user friendly as well. In the late 1960s, online services began to be available, but limited to a few countries and organisations (Convey, 1992; Lesk, 1996).

Later in the 1970s, on-line services started to be widely available to more users. Whereas options had been limited to requests handled remotely by batch searching, developments in telecommunications allowed users to access the databases directly (via teletype terminals, acoustic couplers). The dial-up access provided more flexible searching - users could access the database when they wished - for the users in USA and later in Europe, but interrogating the databases required use of command language, and there was (for reasons of costs of access) an emphasis on development of neat and efficient search strategies (Neufeld & Cornog, 1986; Convey, 1992). Database numbers (numeric and full-text) continued to increase as nearly all organisations transformed their data storage into computerized versions, increasing the amount of information available. However, searching was mostly still done by professional intermediaries, as dial-up access cost by the amount of time spent searching and few organisations were well enough resourced to allow their staff open-ended access to the databases. A little later, CD-ROM databases appeared in the 1980s (Convey, 1992). These (partly a result of their costing structure) allowed libraries to offer their users direct access to the databases to search for themselves (end-user searching). There had been previous attempts by the National Library of Medicine (GRATEFUL-Med) to reach end users directly, but it was not until CD-ROMs came in, that this process really accelerated. The following diagram summarises the phases of information retrieval over the time:

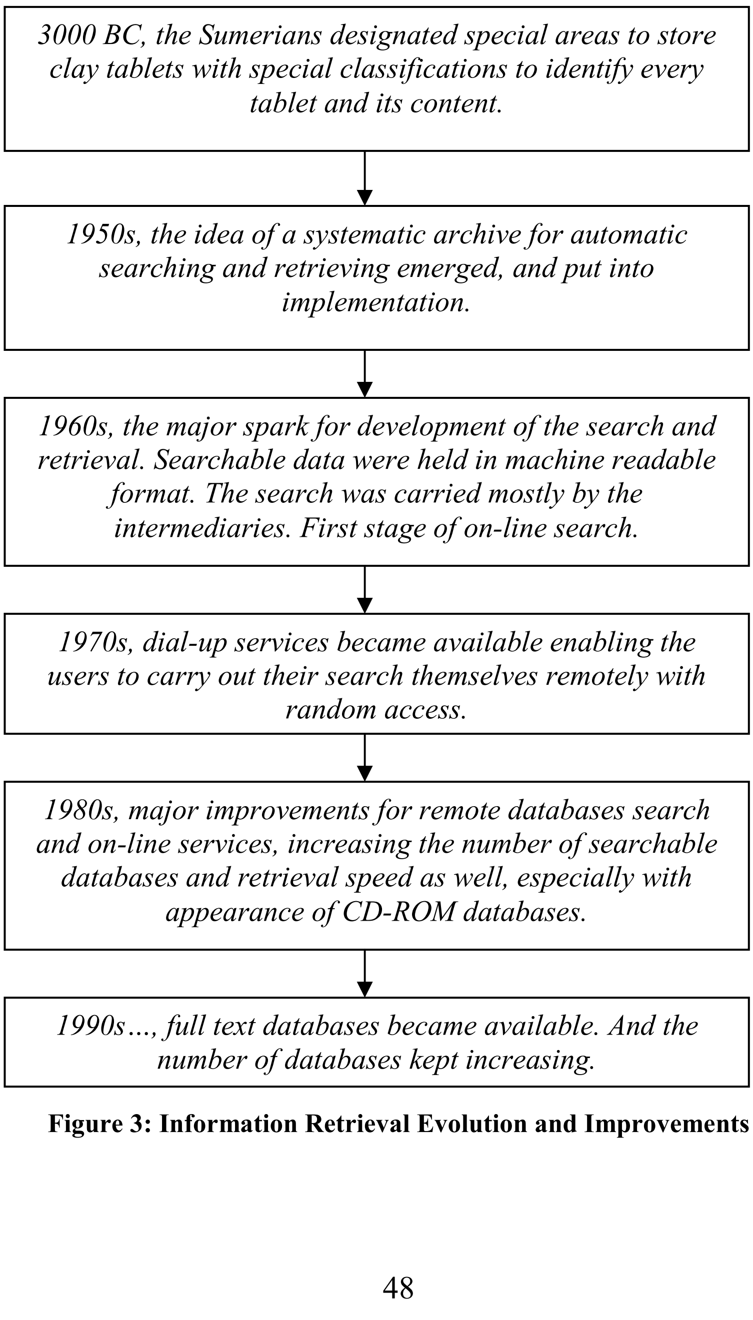


Figure 3: Information Retrieval Evolution and Improvements

This brief outline of the history of information retrieval helps to explain the epistemological perspective – what should be regarded as acceptable knowledge about information retrieval. Clearly at the start of information retrieval in the 1960s the emphasis was on technical knowledge, an engineering approach of develop and trial and test. The system was a tangible object, and the aim of research and development was to manipulate the system to perform more efficiently, by organising the records in inverted file structures. As systems representation and perception moved from physical realm to the logical one the necessity to understand information retrieval systems was a crucial demand, therefore, data organisation and representation and human-computer interaction became more important for the purpose of understanding information retrieval systems, their potential and possible problems for the users of those systems.

3.2.3 Subject Representation and Indexing

All information retrieval systems aim to retrieve the documents and information that best match the users' requirements. Therefore, document organisation and/or representations are considered to be fundamental within the retrieval systems. Representing each document using appropriate and descriptive words that identify the subject and the contents of that document is referred to as indexing and consequently indexes are considered to be the core of any information retrieval system (Liddy, 2005). Indexing emerges from metadata, where metadata in general means data about data, whereas in information and content management it means information about objects where objects refer to the documents with different format and design stored in the system. Metadata should provide descriptions, extract properties and any information that can uniquely identify an object within the system (Garshol, 2004).

However, Croft stated that the best ways to represent document content or the user's need are still unclear; moreover, relevance judgements are still a matter of huge debate (Croft, cited in Ingwersen, 1992). Therefore it is difficult to know how to represent a document precisely, to construct a matching procedure (Section 3.3.1) and to make a relevancy judgement. Indexing theory revolves around two main concepts; controlled vocabulary and natural language or a mixed approach that involves both.

In controlled vocabulary indexing, the index is constructed by assigning specific terms from a list of standardised terms to a document; in this case the users are expected to consult this list (thesaurus) in order to build their search query. Accordingly, thesauri provide a standard vocabulary for indexes and search queries, with the option to broaden or narrow the search query according to the user's need (Baeza-Yatez and Robeiro-Neto, 1999). However, with controlled vocabulary indexing, indexers' inconsistency is still a problem in most digital libraries and databases, and a user's familiarity with the search terms and query (user's perception of the 'aboutness') varies from one user to another (Ingwersen, 1992).

The natural language approach is based mainly on the author's view of the *aboutness* of the document. Indexing and matching follow an algorithmic approach, using title, abstracts and full text to represent a document. Although this can eliminate the indexer's inconsistency, it leads to author's inconsistency. Each approach has its merits compared to the other. However, research and tests have failed to favour one over another, so the general suggestion has been to combine both approaches, thus creating hybrid systems (highly recommended) (Chowdhury, 2004 & Muddamalle, 1998). Controlled vocabulary revolves around standards and uniformity, while the natural language approach provides the users with more flexibility.

Automatic indexing became available, replacing or supplementing manual indexing, where document analyses and indexing are performed automatically. The process involves meta-tags, controlled vocabulary and subject headings. The computerised indexing systems set the rules for the human indexer making the manual indexing process partially automated, and thus it performs more accurately in retrieval and controlling the indexing process (Sykes, 2001). Automatic indexing can be fully automated based on a terms weighting algorithm, where all terms in a document are assigned a weight in relation to the entire document, based on term frequency (occurrence) in the document and pre-specified frequency threshold. Terms with a weight that exceed that threshold are the keywords that represent a specific document.

Recently, a free structured approach for classification has been advised, by allowing the users to their own labels to categorize and cluster information. This approach is often referred to as social tagging and became associated mainly with web resources, e.g Delicious, as a collaborative tagging system for web bookmarks. According to Hammond et al. (2005) tagging has become a useful way for users to recall information sources for later use as well as to communicate interesting nuggets of information to other users.

Social tagging is a new trend in information retrieval to allow users to store, organise, manage and retrieve data they intend to share with other users using their own freely chosen metadata to describe their information. It does not follow the traditional hierarchical structure of a controlled vocabulary (broader and narrower terms), and being produced by different users the classification scheme will reflect personal categories. However, the main problem with it might be a problem in differentiating between general categories, which may exaggerate the system's fuzziness depending on individuals' interpretations of the tags and such differences may raise conflicts between different parties (Hammond, 2005; Golder & Huberman, 2006; Huang & Chuang, 2009). Therefore, the need for a systematic approach to social tagging is increasing, in order to make the representation of objects more structured without the loss of the human factor in the whole process. This area still needs intensive research in order to create a framework for designing social tagging systems for a better understanding and interpretation of tags.

Furnas et al. (2006) believe that social tagging can associate users – a community of users, tag writers and system designers – with documents due to the fact that tagging is done by them reflecting their cognitive state and perception of document contents when assigning a descriptive tags, and that this will open the opportunity for information, thoughts and experience to be exchanged between users (Hammond et al., 2005; Chi & Mytkowicz, 2008).

The interest in natural language versus controlled vocabulary indexing (or combination of controlled plus natural language additions) focused on the technical efficiency and effectiveness of system performance, weighing up the costs (associated with human input) and benefits (representations that might fit the user query better). More recently, the interest in social tagging has increased pointing the potential this new approach offers, despite the complexity and ambiguity it may offer.

3.2.4 Users' Needs and Interface

Users are the essential part of every information retrieval systems. Users vary in their nature in general, so do their information needs and requirements. Several studies have tried to categorise users based on the different criteria, for example some categorise users based on the library they usually use, while other categorise them according to the type of activity they are carrying out (Devadason & Lingam, 1996).

Users' needs depend heavily on the environment in which they are active, and as far as there are different types of users, it is expected that their needs will vary. In addition, users' information seeking behaviours vary as well, depending on personality, the characteristics of the information system they are using, their educational levels and background, how experienced they are with search methods and more importantly the type of environment the user is in. Consequently, full understanding of the system, users natures and their needs is essential for information retrieval systems. This can be achieved by investigating each major environment (organisation/institution in which the users are engaged) and its user groups as well as main characteristics of individual users as well (Chowdhury, 2004).

User interfaces serve as the bridge that connects the users to the information source. They provide an environment for the users to search and browse the information resource, and display the search results. The interface is an important part of the information system as it is the visible part of the system that the user interacts with, and therefore it is important to design an interface that is efficient and easy to understand and manipulate. Interfaces may help the user with query formulation using either Boolean searching or natural language searching, and then they may provide the search results with relevancy ranking and sometimes with categorisation, thus enabling the user to evaluate the search results and review them accordingly. Moreover, visual representation can communicate more meanings and explanation than the usual methods do (Baeza-Yatez & Robeiro-Neto, 1999 & Chowdhury, 2004).

Visualisation is a new trend in information retrieval systems, exploring how users interact with systems. For example, it may evolve around the idea of displaying results with a connectivity feature from a broader set of documents (domain A) to another narrower set of documents (domain B). Moreover, it has the capability of reconstructing a query based on the content of the document/s being previously retrieved, offering more flexibility in search and retrieval strategy (Alhenshiri, Shepherd & Watters, 2010). For example, the EBSCO database introduced a visual search tool permitting users to view and explore search results efficiently by providing a series of blocks or columns where data are stacked by publication or subject and sorted based on relevancy order or date – according to users' display preference – and when a relevant document is located it can be easily dragged to a collect area to be printed or saved. All of these actions are done in one interface saving the users from having to navigate a few to hundreds of pages to locate relevant documents (EBSCO, 2011).

3.3 Information Retrieval Characterisation

3.3.1 Matching and Retrieval

If the information is indexed with a proper representation within the information retrieval system; the user is ready to conduct the search for retrieving the required information. This section considers how research questions around searching and retrieval have developed.

3.3.1.1 Retrieval Models

The most complicated task for all information retrieval systems is to determine which document is relevant to the user's query. Accordingly, information retrieval models and ranking algorithms are needed to make this decision by providing ranking of the retrieved documents. There are several classic retrieval models: the Boolean model, Vector model and probabilistic model (Baeza-Yatez & Robeiro-Neto, 1999; Chowdhury, 2004; Manning, Raghavan & Schütze, 2009).

The Boolean model is the basic and the simplest retrieval model used. This model works by determining the absence or presence of index terms in the document, and then assigning a binary weight to the terms, either relevant or non-relevant. The output in this case does not provide a ranked list of documents. Moreover, the Boolean model has other limitations, since the user has no hints or restrictions on formulating the search statement, which in turn may be too narrow or too broad, with only the binary ranking as well, leading to either too little retrieval or too much retrieval (Baeza-Yatez & Robeiro-Neto, 1999; Chowdhury, 2004; Manning et al., 2009).

Using a proper weighting for index terms should enhance the retrieval performance, and therefore a vector space model has been suggested to overcome the binary weighting process and its limitations. The vector model assigns weights to index terms in query and documents as well, and then uses these weights to compute the degree of similarity between documents and the user query. The weight assignment is based on two measurements; a) computing term frequency within the document (terms frequency *tf*) and b) computing the inverse of the frequency of the terms (inverse document frequency *idf*) where the terms that occur in many documents are deemed not suitable to be a distinguishable factor, (Baeza-Yatez & Robeiro-Neto, 1999; Heimstra, 2009 (cited in Goker & Davies (date)); Manning et al., 2009). Therefore, the documents that match the search query even partially will be retrieved unlike the Boolean model. Although the advantages of the vector model include some degree of simplicity and speed, it has some limitations. According to Baeza-Yatez and Robeiro-Neto (1999) the vector model does not take into account the locality dependency of the index terms and assumes that the terms are independent, which may harm the overall system performance.

The Probabilistic model is based on probability theory. The system has to make the decision about the documents being relevant based on the representation of documents within the system. A document is represented as a vector of attributes/descriptive units or terms describing its contents. The probabilistic model revolves around three basic principles: the probabilistic indexing model; the probabilistic retrieval model; and the probabilistic ranking model (Sparck , Walker, & Robertson, 2000).

The probabilistic retrieval model can be considered a binary model, as the document retrieved is either relevant or non-relevant (depending on the bar set for relevancy), but, in any case, the query term is present or absent in the document. To implement probabilistic retrieval model there is a need to estimate the term occurrence in the document and how that term contributes to document relevancy, therefore, term frequencies, documents frequencies (the number of documents that contain the term) and document length need to be explored. Each term is assigned a weight according to its occurrence in the document. Term weighting can be considered as a relation between the number of times in which the term *k* occurs in the document, *tf(k,d);* the number of documents in the collection that contain that term *k*, *df(k)*; and the number of documents in the collection *N* (Sparck et al., 2000; Hiemstra, 2009).

The actual document's relevancy information is unknown in the beginning. Therefore, the probability of the document relevancy is based on the estimated probability of terms in terms of the incidence vector, and consequently all documents get a probability score, then the documents are ordered in a descending order in the set accordingly (Chaudhuri et al., 2004). After the initial guessing of the relevancy probability, the process becomes iterative as the system can benefit from user feedback of relevancy, enabling the system to calculate a new set of weights reflecting the importance of each query term more accurately and hence leading to more enhanced retrieval performance (Baeza-Yatez & Robeiro-Neto, 1999; Chowdhury, 2004; Chaudhuri et al., 2004; Manning et al., 2009). The main limitation of this model is that the initial relevancy judgement can be simply a guess.

3.3.2 Information Retrieval Research Trends

Information retrieval is obviously an active field. It is evident that information retrieval and retrieval systems have been improved in terms of quantity, and hopefully quality for users, particularly the end-users. There are identifiable trends in the research, but priorities have changed over time.

The Cranfield experiments started a trend in research on documentation representation, and efficiency and effectiveness in information retrieval. Perhaps the debate on 'relevancy' really started here. TREC experiments provide a large-scale evaluation of text retrieval methodologies, encouraging the communication among different parties and opening the opportunity to exchange of research ideas.

As for the Digital libraries, a main theme is the interoperability (See Section 3.4), mainly through co-operative cataloguing and metadata standardization (i.e. MARC format) that is created either by librarians or by authors - to support Web scalability – (Mathes, 2004). Library management systems provide different interfaces that can provide help and support for the users, and user interfaces began to be important areas of research, particularly with retrieval of non-text, non-numeric information that became possible with the integration of image and music data into the library management system.

Furthermore, and as the Internet grows continuously as well as the amount of data it provides, search engine performance became a topic of research (Section 3.3). Whether for general or specific purposes, search engines are always under a continuous evaluation, changing and evolving their mechanisms and their user interfaces in order to maintain the quality and capability to meet users' needs (Atsaros, Spinellis & Louridas, 2008).

That concept, of meeting users' needs has remained a constant undercurrent in much of information retrieval research. Ideally, search engines need to capture and observe users needs and information seeking behaviours to develop their mechanisms, but information seeking behaviour is neither predictable nor regular as it is human behaviour (Section 3.2.4). Kim (2009) affirmed that information seeking behaviours are correlated with the task being performed by the searchers and are influenced by it, due to wide variation in tasks and activities (professional activities, research activities, personal activities, etc…) being done by users. Information seeking behaviours are complex to define unless we are able to understand the context in which the task is taking place and the nature of information needs as well. Consequently information systems structure and design might be developed bringing a better performance in term of satisfying users' needs (Donald, 2007; Kim, 2009). It is clear that the user interface affects the search result and accordingly human-computer interaction receives much attention in order to maintain search system performance (Gwizdka, 2009).

Yet, there is another notion that is related to and strongly affect users' need and information seeking behaviour, that is relevance judgement and criteria, and (since Cranfield) much research has tried to address and discuss relevancy (judgement and criteria) in more depth, but up to now there is no stable and definitive criteria of how to define the relevancy. Most research has concluded that relevancy has some certain dynamic characteristics that reflect the human nature of thinking, as the user's cognitive state grows and evolves during the search (Vakkari & Hakala, 2000), and hence the search becomes more focused, but the topicality factor, perhaps, remains the crucial criterion for relevancy (Section 3.6.2).

Generally speaking it is obvious that information retrieval science research revolves around:

- Information resources indexing and organisation (content representation) (Section 3.2.3).
- Information systems search and services, systems inter-operability and Interfaces.
- Information formats, retrieval methods and retrieval ranking.
- User behaviour, information needs and relevancy criteria.

3.3.3 Information Retrieval Research for Systematic Reviews

As indicated earlier, some of the early information retrieval developments occurred in the health sector, with the development of MEDLARS, and the developments in bioinformatics and genomic information continue this trend. Dee (2007) described the historical development of the MEDLARS system, noting that there was debate about the need for evaluation of the MEDLARS bibliographic searching, in case the evaluator (Lancaster) produced negative findings. In fact, the Lancaster evaluation in 1965 provided an early example of recall and relevance data and the evaluation was not just important for the development of MEDLARS, it provided a prototype for future evaluations of bibliographic retrieval systems.

However, within healthcare itself, the wealth of research evidence, and the choice of treatments available have affected the type of information retrieval research that is required. The main focus is the need for health professionals and policymakers to find solutions to the problem of coming to a well-informed decision about the clinical efficacy of an intervention for a particular patient or group of patients.

Systematic reviews of the research evidence have become very popular and important in these days for many varieties of users, i.e., professional practitioners, managers, decision makers for planning and policy, and researchers. Primarily, systematic reviews inform decision making, to ensure decisions are made on the basis of the best available research evidence in the area of interest. However, conducting systematic reviews is challenging due to the massive amount of research studies in one research area or health topic (sections 2.2 & 3.3.3).

In a systematic review search strategies are the main concerns in the retrieval process due to the importance of systematic reviews as a rich source of 'synthesized knowledge' for evidence based consumers (Yoshii et al., 2009). Therefore, researchers sought to develop a search strategy with an optimal retrieval performance, probably with more emphasis on finding all relevant items (optimising recall at the expense of precision). Early reporting of search strategies for Cochrane Reviews was not consistent (Yoshii et al., 2009), but more recently there is an increased demand for more transparent reporting with attention paid to the quality of the reported search strategies so that the reader can judge the credibility of a systematic review. More rigorous search strategies should help in future attempts to update the systematic review by replicating the search to maintain consistency in the systematic review (but this depends partly on consistencies in vocabulary and its usage). Furthermore, reporting a search strategy explicitly can provide general knowledge about the retrieval process in certain clinical area (Yoshii et al., 2009). Although Yoshii et al. (2009) found that none of the analysed systematic reviews complied with the Cochrane Handbook guidelines on search strategies reporting, they used a small number of systematic reviews which were all published in the same time, leading to suggestions about some bias around the conclusion drawn from this study. McGowan, Sampson and Lefebvre (2010) describe the operation of the PRESS forum, for peer review of search strategies, to help improve the quality of the searching.

Electronic search strategies need to be devised for systematic review searching. These need to be comprehensive but also efficient search statements to be used in retrieval process. In addition, several databases are usually available as a resource and/or a repository of most of the available (published) evidence. Consequently, in order to find all, or at least most of the relevant studies, searching several databases is necessary due to the differences between databases' coverage of articles including selection procedures and indexing processes (McGowan & Sampson, 2005). Therefore, taking advantage of the differences in indexing across databases will increase the chances of retrieving relevant items that may be in different databases, but easier to find in one database than in another with the chosen search strategy (Section 3.2.3). In other words, to increase SR quality (and limit the bias) a comprehensive literature search is required to identify as much of the relevant literature as possible. Search terms and filters used should also be taken into consideration when carrying out the search (Section 3.5). Each database has its specific search method, filters and interface which searchers should examine when conducting the search process to avoid search errors that may affect sensitivity, specificity and precision negatively (Sampson & McGowan, 2006). Search filters have been devised by various groups – but principally the Hedges team at McMaster University in Canada (Section 3.5.4).

In this context the Cochrane Information Retrieval Methods Group aims to support the Cochrane collaboration's information retrieval activities. The group provides policy advice on information retrieval monitors the quality of search reporting methods and conducts methodological research. Members of the group have published research on updating methods (Moher et al., 2007) and search reporting methods (Sampson et al., 2008a).

And recently, Sampson conducted research focusing on information retrieval issues in updating health systematic reviews. This research aimed at combining a high level of precision with a moderate level of recall to locate the best and most relevant new evidence. The 'Related articles' search on PubMed in MEDLINE, 'Citing references' and 'Boolean search' were tested in this research study to establish the patterns and performance efficiency of each search strategy. Recall and precision were used as the main measurements of this study (Sampson, 2009). This research found that each search strategy's performance varied depending on the clinical area being searched. Similarity searches (PubMed related articles (Section 3.5.2)) outperformed all search strategies showing higher precision over all other searches, but both Boolean search and related article searches seemed to complement each other producing (together) high recall for new evidence, implying that a combination of more than one search strategy was required (Sampson, 2009). Citing reference search performed poorly in this study, as, according to most recent evidence will not have time to be cited by other studies, therefore, due to the short period of time that this study was done, using cited reference search was not worthwhile (Pao & Worthen, 1989).

Moreover, Sampson *et al.* (2006a) researched the possibility of taking advantage of search engines in order to help in systematic review screening process. Search engines can provide a ranked search output according to search terms used in search query formulation and where these terms appears in a document and their frequency (Section 3.3.1). Therefore, if it is possible to profit from this functionality from search engines, especially for Boolean search queries to improve the relevancy ranking algorithm placing the most relevant records in a smaller ranked list, then time and effort for screening search result for its eligibility to contribute in systematic reviews can be reduced.

The Ultraseek search engine, due to its ability to handle meta-data was selected and configured in this study to investigate relevancy ranking of Boolean search results functionality, and the study finds that relevance ranking to speed up the creation of systematic reviews is technically feasible to implement (Sampson et al., 2006a). However this depends mainly on a careful selection of search terms as well as the order in which these terms are entered. This limited the generalisation of the findings of this study, although it seems that enhancing relevance ranking for systematic reviews would improve performance.

Sampson *et al*. (2009) asserted that complex and highly sensitive electronic literature search strategies are required for systematic reviews, thus search strategies need to focus on those elements that will negatively impact search performance such as subject headings used in the database being searched, and the choice of logical operator to combine the search terms. It should be kept in mind that the indexing process differs from database to another and different authors may not use the same or standardised subject terms in their publications.

Furthermore, the Cochrane handbook has recommended that the searches for systematic reviews should be as extensive as possible in order to ensure the inclusion of as many relevant studies as possible. However, this is inevitably associated with lower precision (Higgins & Green, 2011). With a higher precision, screening time can be reduced and accordingly, Taljaard et al. (2010) conducted a study aimed to investigate the possibility of enhancing precision rate while preserving the recall rate. In this study several well-known search strategies were evaluated in term of sensitivity (recall) and precision against a gold standard created manually, and then these search strategies were combined with clusters designed of related search terms. The results showed that precision can be improved by combining cluster design-related terms with regular search strategies, yet there are some limitations that suggested the further investigation for future improvement in general.

Furthermore, over time, greater attention has to be paid to the process of updating systematic reviews to maintain the systematic review's quality and currency. Sampson's 2009 doctoral thesis focused on this issue of updating systematic reviews and explored several search strategies' performance for this purpose, using recall and precision as the major performance measurements used. In conclusion, it is obvious that all systematic reviews research aims to keep the quality of systematic reviews as high as possible. This covers the production of the initial systematic review, as well as the updating process. However, there is increasing interest in the contribution of economic studies and qualitative research to the systematic review process, particularly for policy decisions about implementing the best research evidence – knowing how best to implement an intervention which has been demonstrated to be effective.

Recently, evidence mapping has emerged to complement systematic review in a broader clinical context; it can link the gaps and strengthen the systematic reviews, but it requires a wide ranging overview of the literature, and could be considered to inform the implementation of any intervention (or set of interventions) in the broader clinical context. According the Cochrane handbook recommendation it is essential to use a highly sensitive search though this will results in dramatic degradation of precision (Higgins & Green, 2011). And therefore Parkhill et al. (2011) examined the retrieval performance of the recommended search strategy for evidence mapping searching. However, when employing the highly sensitive search - in the first and second reviews in their study - it retrieved an unworkable number of references to be handled with the available number of staff and therefore in the final stage of their research they planned a highly specific and a highly sensitive search strategy which they referred to as evidence-based mapping. They compared its retrieval performance with the highly sensitive search on MEDLINE.

The comparison criteria which were used in this research were: the number of studies retrieved; the number of included studies; the number of missing studies and staff time. They used six questions from existing traumatic brain injury evidence maps which were selected randomly to test and compare both search strategy performances. For each one of the six questions a new search string was formed for both search strategies.

The results suggested that evidence-based mapping strategy is more precise than the highly sensitive search as the proportion of included studies to non-relevant studies retrieved is higher than the highly sensitive strategy. The evidence-based mapping search retrieved 1818 studies and the highly sensitive search retrieved 2599, 43% more total studies than the evidence-based search. The number of included studies from the evidence-based mapping search was 120 (77%) studies compared to 125 (81%) studies eligible to be included. The evidence-based mapping search missed 35 studies while the highly sensitive search missed 30 studies, and 30 studies were missed by both search strategies. Of the five studies which the evidence-based mapping strategy missed, three were not written in the English language and were found to be marginally related to the clinical topic and therefore they were tagged as studies of possible interest, which were discovered later to be indexed incorrectly in MEDLINE. Moreover, in terms of staff time, the analysis showed that reviewing the additional 781 retrieved by the highly sensitive search needed about 19.5 hours for the conventional review process.

In conclusion, Parkhill et al. (2011) found that the highly sensitive search strategy produced greater recall, as would be expected. This was, however, compared with a new search method of using a simple search string; that of the evidence mapping strategy. Comparing its performance with the traditional HS search strategy, the results here showed that the simple search strategy employed with evidence mapping is unlikely to miss the key articles for decision making, although this result is based on one clinical area, of course. In the end, in the context of evidence mapping, the benefit of producing an evidence map within time and budget restraints outweighs a small risk of missing marginally important studies and therefore simple search strategies are favoured over more highly sensitive search strategies.

Accordingly, this current research continues some of the past and current areas of interest in information retrieval research. This research aims to assess and investigate different search strategies' performance, over different databases, and the retrieval efficiency of each database. The focus is novel, as the area of research is in retrieving the sibling studies of specific RCTs (studies of different types and study designs that are directly or indirectly related to that RCT) that can contribute to systematic review production.

3.4 Digital Libraries

Digital libraries have then emerged, an idea to make use of computer capabilities and various functionalities, for information retrieval by a wider range of users, with a wider range of library content, making full use of library management systems (LMS). Accordingly, the librarians' responsibilities expanded from managing the library collections that are physically located in one location to managing and providing access to a wider range of digital resources within and outside the library (Tedd, 2006). Consequently, Baeza-Yatez & Robeiro-Neto, (1999) and Chowdhury (2004) claimed that from an information retrieval perspective digital libraries can be considered as an extended information retrieval system. They provide access to different information resources, with diversity of information formats and perhaps different languages, located in different places over the world and available to users with different requirements and needs.

The notion of digital library and databases is to accommodate different information resources in one place, regardless of being distributed over different computer systems in different locations (Borgman, 1999 & Chowdhury, 2004). Consequently, interoperability has different types and levels, i.e. systems interoperability, software interoperability and linguistic interoperability. Therefore, professionals suggested that this can be solved by using some standards to represent information within the system such as user interfaces, metadata format, networking protocols (communication protocols) and information retrieval protocols (e.g. Z39.50) in an effort to alleviate systems differences and making them compatible for working together (Chowdhury, 2004). However, with current LMSs, most of these obstacles were eliminated or lessened as LMSs became portable with a more powerful inter-connectivity and more user friendly interfaces. The employment of links from the LMS to the Internet made digital libraries more capable (Tedd, 2006).

Users can search databases from anywhere using local computers to communicate with other systems elsewhere, expecting the most relevant results to be presented first, as now expectations are based on Google searching (Clarke, Cormack & Tudhope, 2000). Online search was the facility introduced to search enterprise data resources with the introduction of the Web and the Internet, and along with it a wide range of databases and services became available. Searching the Web is considered to be a complex task due to the vast size of the web and the information it holds. Various types of search engines are available to assist the users with their search, and although each search engine has its own characteristics and capabilities, all engines use software called *spider or crawler* (my italics). This program traverses the web using tree representations and algorithms, and following the links available on the web it moves from one page to another. Selecting and ranking the web pages based on the search query differs from one search engine to another (Clarke et al., 2000). For example, Google uses citation analysis techniques to determine the web page importance by counting the number of times the document/web page has been cited, giving it a page rank (Chowdhury, 2004). For obvious reasons, many of the commercial information retrieval products and services do not give details about the way their algorithms work. But according to Yatez, all search engines in general use Boolean and vector space as their ranking models, based on their indexing method and properties (Baeza-Yatez & Robeiro-Neto, 1999).

In conclusion, libraries and web can be considered as synonyms where both aim to satisfy user information needs, both the digital library and web provide access or links to each other. As aforementioned, digital libraries have different formats from which the user can choose based on his/her information needs. Some libraries provide users with the capability of searching across several information resources and services on one interface saving the time costs of having to move from one interface to another. This is called federated search or meta-search which has become a practical tool in information retrieval systems. However this will increase digital library complexity. Such searching accesses different information resources with different interfaces and contents representation, maintaining interoperability between them and returning the search result to a single virtual interface without users' full awareness of the mechanisms involved. Aberystwyth University meta-search – used in this research – and Google Scholar, are examples of federated and meta-search approaches.

However, despite the advantages that federated searches have in allowing users to search several resources at the same time - saving time and effort in repeating the same search across several resources individually - this does not come without some difficulties. Different information resources have different data representation and indexing, different interpretations of the search query as well as the differences in relevancy ranking criteria employed. Data redundancy is to be expected, as the same document may be stored on different databases or data sources (in slightly different ways). However, a critical problem is how to integrate multiple retrievals from different resources into one unified ranked list of retrievals to the user.

Consequently, it is very critical for the meta-search engine to select the appropriate resource to search and retrieve relevant result to the initial query. The Bayesian Inference Network Model of information retrieval is a technique employed by most of the ranking algorithms to rank data sources. It helps to decide which source provides representative information and how many relevant documents it can provide in order to be selected to be searched. Besides, it is a good model to manage uncertainty (Acid et al., 2003). Relevance based ranking is a highly recommended technique to assign relevance score for documents from different sources in order to provide a final ranked list of retrieval according to the user query(Si & Callan, 2005).

In summary, the Web links different data resources stored on millions of computers across the world. Different search engines use different retrieval techniques in searching for web pages (web crawling), indexing them in their database and assigning weights and relevancy scores (probabilistic criteria of some sort, usually). The information about algorithms and procedures used in search engines are exclusive information which are not made public to the users, although Google is an exception (partial) as it provides information about the techniques employed in their systems. Digital libraries provide a similar type of probabilistic retrieval to resources that may require a subscription to access. Interoperability of specialist resources and databases has been a major concern.

3.5 Searching for Clinical Studies

3.5.1 Searching for Types of Clinical Study

Retrieving and processing all siblings may be a desirable aim when conducting a policy-based systematic review. A comprehensive coverage of the subject from all perspectives will provide a better understanding of the subject, thus leading to more robust and reliable results and conclusions, particularly if the studies are related in some way.

Successful retrieval of required and/or relevant studies is the specific and crucial demand of almost all research fields and systematic reviews; however, identifying all relevant studies is a difficult process due to several issues such as heterogeneity of indexing terms and keywords used to describe studies among different databases as well as the variation in terms used to describe subjects (Section 3.2.3) (Goss et al., 2007). Many efforts have been made in order to make search and retrieval processes more efficient and productive, so search strategies and filters have been developed (Dixon-Woods et al., 2001), (Section 3.5).

The most popular searches used in retrieval of the medical literature are: subject search using medical subject terms or headings to retrieve relevant articles, author search, citation search which means searching for article that cited a specific article that is known to be of a high quality study, and related articles search within the same specifications as the citation search, (Sampson et al., 2008b).

Other studies have examined strategies for located health services research that is not purely clinical. Papaioannou et al. (2009) found that subject search provided 73% of relevant studies when searching for social science studies, though the studies were obtained by searching multiple databases, and therefore, searching multiple databases appears to be essential to achieve and optimal performance (for non-clinical topics). This reflects the multidisciplinary nature of social science research, with a need to plan search strategies and selection of databases. Moreover, Papaioannou et al. (2009, p. 119) stated that "sensitive systematic review searches are not always exhaustive and unique references are identified via supplementary search techniques". In the same context, Relevo (2012) attributed retrieval performance to reporting and indexing factors. He stated that even highly sensitive searches would still miss relevant items and therefore searching multiple sources is imperative.

Other studies examine how to locate specific aspects of clinical interest. Golder et al. (2006) aimed to assess four subject search strategies performances: 1) searching for specified adverse effects using suitable indexing terms; 2a, b) searching with adverse effects subheadings with two variations *subheadings linked to drug name indexing term*, and *subheadings alone ('floating')*; 3) text word searches for synonyms of 'adverse effects' and finally 4) related terms and searching with indexing terms for 'adverse effects', using measures of recall and precision on MEDLINE and EMBASE. Golder et al. (2006, p. 6) created a gold standard of 84 adverse effects records pooled from the records retrieved from MEDLINE and EMBASE. In addition to the pooled list, there were records recommended by reference lists, clinical experts, the effectiveness searches, and submissions from drug companies.

Floating subheadings was the best approach with the highest sensitivity in MEDLINE. The highest precision was achieved by using *Subheadings linked to drug name indexing term*, (Golder et al., 2006, pp. 6-7). With EMBASE, using *Subheadings linked to drug name indexing term* was the best approach in terms of recall, the precision in EMBASE was lower than MEDLINE.

To further enhance the performance, combination of different search strategies were advised. In MEDLINE, a combination of specified adverse effect terms, floating subheading and text word gave the highest sensitivity (approaches 1, 2*b* and 3). For

EMBASE, specified adverse effect terms with text word searching was the winning combination (approaches 1 and 3) (Golder et al., 2006, pp. 6-8).

In a later study, Golder and Loke (2009) investigated and evaluated search strategies aimed to identify methodological studies on adverse effects. The authors assumed that in order to identify as much as possible of the literature on this type of study the search should be carried out on multiple bibliographic databases. However, identifying methodological studies had proved to be difficult because of differences in indexing this type of study. Therefore, Golder and Loke (2009) claimed that to overcome this barrier it is important to develop a pragmatic search strategy based on terms from titles and some free text words. The study findings indicated that use of floating subheadings (subheadings which are not attached to any indexing terms) has the potential of achieving a high sensitive search filter in MEDLINE. However, the performance of the EMBASE search strategy recommended that using subheadings derived from the drug name for the intervention can provide a good performance. Moreover one of the included studies emphasises the value of using text words in its search strategy. A highly sensitive search had been achieved with a very low precision. One of the included studies in this study is also an author of this systematic review which may cause bias. However, as the analysis was conducted by two researchers this bias might be lessened.

A difference in database performance was also observed, which is considered to be a strong indicator of databases' indexing differences. Sometimes the reporting method influences the indexing process which may indicate the differences in reporting quality, as well. In addition both subheading search and text word search have proved to perform well although with the case of text word search more investigation may be required to make the search process more focused. This study focused on one topic/subject area, making the generalisability to other clinical areas somewhat dubious.

Agoritsas et al. (2012) argued that searches that are applicable and easy to use in clinical practice are favoured by practitioners who seek rapid answers rather than sophisticated strategies used for performing systematic reviews. Accordingly, he proposed to utilise intervention methodological framework (PICO) and the clinical query narrow filter for a more likely retrieval of relevant clinical trials within readable outputs. Interestingly, this research indicates the need for retrieval indicators that combine recall with precision.

3.5.2 Related Articles - Similarity Searching

PubMed related articles search is the most well-known non-Boolean search option. The related articles algorithm works on measuring the similarity between documents by assessing the words in common between documents (textwords and MeSH terms). Each document has a list of words that represent the document (NLM, 2009). Each word has a different value and weight depending on the frequency of the term. Weight assigning is done automatically within the system, depending on three types of information:

1. *The number of different documents in the database that contain the term.*

This information is used to weight the term at the entire database level giving it 'global weight'. The more frequently the term occurs in the database, the lesser global weight it is assigned. The rationale of this assignment is that common terms that occur often tend to be misleading and provide less information about individual documents, while terms with less frequent occurrence can be considered as signature terms to a limited set of documents in which they occur. Therefore more information can be obtained from the term giving it a higher weight.

2. *The number of times the term occurs in a particular document.*

   and

3. *The number of term occurrences in the document.*

This information (from 2 and 3) is used to generate the local weight of the term in a particular document, assessing its importance in that specific document. The local weight mechanism works in reverse compared to global weight, thus more frequent terms represent the document's content. But, the length of document should not be allowed to affect the weight, for example, longer documents will result in higher frequency of one term giving it greater weight leading to making it a more important document (simply because the document was long). Thus to limit this from influencing the weighting criteria, there is a threshold for the term frequency.

In case the term occurs in two documents, its weight will be the product of its global weight and its two local weights in the two documents it has occurred in, i.e. (local wt1 × local wt2 × global wt). To compute the similarities between two documents for example, all the weights of all of the terms the two documents have in common will be added and a final similarity score for both documents will be obtained. This procedure will be repeated until all of the similarity scores of that specific document in relation to each of the other documents in the database have been computed, and the highest score is considered as the most similar. It should be mentioned that this is a pre-computed score so when a search is performed a list of its related articles appear alongside with it.

Lin et al. (2008) attempted to examine the use of the PubMed related articles algorithm, observing users' search behaviour (search logs) over a week in June 2007, which revealed that a fifth of non-trivial PubMed users used related articles search at least once. They found that this feature has become a crucial part of PubMed searching patterns. Similarity search approach might enrich the retrieval results and can be satisfactory for the users (Lin et al., 2008). In another study Liu and Altman (1998) revealed that a PubMed related articles search can achieve a recall of 75%, a strict precision of 32% and a partial precision of 42% when trying to update a bibliography using an incremental approach. However, a better performance can be achieved if multiple seed studies were combined to benefit from nearest neighbour conjunction (Bernstram, 2001).

3.5.3 PubMed Automatic Term Mapping (ATM)

PubMed Automatic term mapping is a procedure/algorithm that matches entries without tags in the search box against a MeSH translation table, a Journal translation table, the Full author translation table, Author index, the Full investigator translation table and an investigator translation table. Each one of the translation tables has its own specific contents to match the search against in order to do the translation properly. For example, the MeSH translation table has MeSH terms, entry terms for MeSH term, MeSH subheadings, publication types and other entry type, such as substance and their synonyms.

The search terms (untagged) will be matched against theses tables for its match search in PubMed. If a match is found then the term will be searched using the matched tag found. In other words if the term match is found in the MeSH translation table then the term will be searched as MeSH term with any specific terms associated with it in the MeSH hierarchy as well as all fields. The same procedure occurs with Journal translation table where the entry term will automatically be mapped to the journal abbreviation that is used to search journals in PubMed and all fields as well.

PubMed automatic term mapping performance was investigated by Lu, Kim and Wilbur (2009), using TREC 2006 and TREC 2007 topics as performance assessment criteria. In this work there were no search tags assigned to the original search query issued by the researchers. After exploring the mapped search query translated by PubMed Automatic Term Mapping, all search terms were associated with either [MeSH Terms], [Text Words] or [All Fields], where each search term can be assigned more than one tag. Lu et al. (2009) concluded that Automatic Term Mapping can expand the retrieval results with a greater possibility of retrieving more relevant documents. However, the process appeared to be more in favour of recall rather than precision and therefore it will not appeal to users interested in first ranked retrieval. In other words, if users are only looking for the first 20 or 30 records then PubMed Automatic Term Mapping will not be that much use (Lu et al., 2009).

3.5.4 Search Filters

Search filters are available for economics and qualitative research (and therefore such siblings). For example the Hedges team at HiRU[♣] has developed a search filter for different types of study for key clinical databases. Each search filter focuses on a different type of intervention (e.g. diagnostic studies) or different research design. Filters can improve the retrieval process of scientifically sound and clinically relevant study reports from large, general purpose, biomedical research bibliographic databases[♠].

♣ 'McMaster University and funded by the US National Library of Medicine'. For more information see: http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx.

♠ Hedges team created search filters for MEDLIN (Ovid and PubMed), Embase and PsycINFO. CRD created filters for CINAHL as well. See http://www.york.ac.uk/inst/crd/intertasc/index.htm.

The purpose of the Hedges project was to: 1) provide effective clinical search strategy(s) for health care providers; 2) retrieve all important relevant citations on health care problems; 3) enable the librarian to help clinicians to construct their own searches by providing them with the required resources; and finally 4) provide a form of guideline for indexers of databases about their indexing processes and the organisation of their databases (HiRU, 2007).

According to the Hedges team (HiRU, 2007), the main obstacles that prevent the ultimate enhancement in the search process are:

> Indexing and retrieval in large databases, and the widespread and rapidly increasing direct use of these databases by clinicians, researchers, educators, administrators, lawyers, journalists, patients, and the general public whose need to retrieve the relevant studies to take the appropriate decision as possible as they can.

This reflects the differences in indexing process between databases; differences in language and experience of users. It is rather difficult if not impossible to provide indexing terms that would provide all the needed access points to ensure that individual needs can be met, with some acceptable precision. Professionals who are familiar with specific medical terminology, and professional searchers, who have a good understanding of such terminology and its relationships can be more precise, more specific in their use of search terms compared to the non-professional who will use different language and terminology to describe the same ideas – some of these may be seen in the "entry terms" listed for each MeSH subject heading, but the query itself may be difficult to translate from non-professional phrasing to a professional way of expressing those ideas.

Haynes and Wilczynski (2004, pp. 1040-1041) carried out a study for search strategy development which focused on retrieving sound clinical studies on the diagnosis of health disorders in MEDLINE. To develop a search strategy for MEDLINE database, a gold standard of studies on prevention or treatments was created by hand searching 161 journals for the year 2000 which were in the database. Each journal title was searched alongside the full MEDLINE records (including citation, abstract, MeSH terms, and publication types).

Then search strategies were created and tested for their efficiency in retrieving articles indexed in MEDLINE, comparing the retrieval with the gold standard.

The authors found that single or two terms search strategy provided a good performance and sometimes it was better than the multiple terms search. The search filters depend on combining as possible as much of descriptors and key terms in the field. The more terms are joined with OR the more non relevant items are likely to be retrieved (Haynes & Wilczynski 2004).

Haynes et al. (2005) emphasise the development of search filters to improve retrieval of clinically relevant and scientifically sound treatment studies from MEDLINE and similar bibliographic databases, and according to the authors, search filters can be built using a combination of medical disease content term and subject headings, publication type, methodological key term (text-word), explosion and subheadings. A list of MeSH terms, and text-words from the included studies, were selected and then validated by consulting clinicians and librarians to construct search strategies. OR was used to combine the terms, AND was not used because it may compromise sensitivity. A gold standard was created following the same process used by Haynes 2004 earlier research.

Sensitivity, specificity, precision (defined earlier) and accuracy were the measurements used to assess the search filter reliability. Accuracy refers to "articles that met criteria and were retrieved plus articles that did not meet criteria and were not retrieved, divided by all articles in the database", (Haynes et al. 2005, pp. 1-2).This search finds that a search filter using single or multiple search terms for randomised controlled trials can achieve high sensitivity (99.3%).

Haynes et al. (2005, p. 5) claim that there is no search strategy that will perform perfectly. Databases use different indexing schemes resulting in inconsistent indexing (Haynes et al., 2005 & Goss et al., 2007) besides, the indexing terms and methods evolve continuously while the evolution of search strategies might be slower than changes in terminology and appearance of new technical terms. Moreover, indexers only use a small number of terms to index a single record (although MEDLINE is more generous than many other databases) and many of these terms have similar meanings—for example, 'randomized controlled trials' and 'clinical trials' as MeSH and 'randomized controlled trial' and 'clinical trial' as publication types. Apparently, NLM is trying to emphasise that the term randomised control trial is used in two different contexts, one as a publication type and the other as a descriptor for documents that discuss general aspects of the randomised control trial.

Haynes et al. (2005) pointed out that search filters were developed based on clinical records of high quality only, but in reality databases in general and MEDLINE specifically will have clinical and non clinical studies as well as studies of lower quality which will affect the actual search performance. The Wong, Wilczynski and Haynes (2006) study objective was to design a search strategy to retrieve sound therapy studies and review articles in CINAHL, using year 2000 published studies as the gold standard. Index terms and text-words from clinical studies were selected and discussed by librarians and clinicians in the US and Canada, to decide on the final list of search terms to test on the CINAHL collection.

This study achieved best sensitivity for identifying treatment studies in CINAHL (99.4%) using a certain combination of search terms with a compromised specificity. Best specificity was also achieved using another combination of terms (98.5%) to identify the required studies but with a compromised sensitivity. Furthermore, to identify systematic reviews a combination of search terms achieved the highest rate for sensitivity, specificity and for optimised search strategy (sensitivity and specificity) of (91.3%, 99.6%, (76.4% and 76.7%))[♠] respectively, (Wong, Wilczynski &Haynes 2006, p. 197).

In line with the previous methodological filters, Wilczynski, McKibbon and Haynes (2011) decided to test if search precision can be improved by NOTing out some terms. They tested their approach on all filters categories on MEDLINE, EMBASE, CINAHL and PsycINFO. They found that precision has been improved for all filters on all databases. However, it is unavoidable that NOTing out content may result in the exclusion of relevant systematic reviews.

[♠] These are (sensitivity, specificity, (optimised sensitivity, specificity))

3.5.5 Qualitative Research Retrieval

Qualitative evidence can be difficult to search for and identify due to several reasons. Qualitative research is multidisciplinary (encompassing social psychology, sociology, health economics, nursing and allied health), and it uses multiple research methods which might cause confusion and inconsistency in naming of its concepts. It will be of interest to many different parties, making it a target for multiple databases with different coverage and indexing characteristics and accordingly the choice of which database to search will be more difficult and in some or most cases searching multiple databases is necessary. The reporting of qualitative research is another issue, as most qualitative studies do not have a structured abstract. Some qualitative research is reported in books or reports, and therefore the search will mainly depend on the title which often reflects the findings of the research rather than mentioning research method, and this aggravates the retrieval process (McKibbon, Wilczynski & Haynes, 2006).

Several attempts have been made to compare search strategies across the databases (MEDLINE, CINAHL and EMBASE) (Evans, 2002; Wilczynski, Marks and Haynes, 2007). From there several points emerge: 1) it appeared that a combination of index term and text words can provide good retrieval on CINAHL, while "text words only" appeared to be more effective on other databases: 2) the terms *interview*, *qualitative*, *themes*, and *experience* (italics in original) when combined appeared to be effective in terms of sensitivity for almost all databases, but these terms varied with CINAHL; 3) *study design, attitude* and *interview* (my italics) were the terms to be used on CINAHL (Wilczynski, Marks, Haynes, 2007). This reflects the variations of terms that might be used to address qualitative studies as well and how that might complicate the retrieval on different databases (indexing inconsistency). In general, text words seem to be more effective in retrieving qualitative research rather than index terms (used by indexers in the database) which emphasise the issues discussed earlier.

3.5.6 Clinical Trial Number and Clinical Registries

As mentioned earlier, (Chapter 2 and Section 3.5.1) it is clear that each sibling study has different issues that might influence the retrieval process, and this may be an added difficulty when trying to retrieve them all together. Consequently, it would be useful to have a specific or unique feature that can clearly differentiate one clinical trial from another and even better, if that unique feature can link each clinical trial to any other trial or study that is directly associated with it. The Clinical trial registry concept is partially serving this purpose, but up to now we are not that well served. There are several agencies and parties that emphasise the importance of clinical trial unique identifier for different reasons, but it would be ideal to expand the aims so that clinical trial and any other study based around that trial have that same identifier. Simply it will act like ISBN (International Standard Book Number) to uniquely identify trials and track all publications resulting from each trial.

CONSORT (2012) produces guidelines called the 'CONSORT statement' to ensure the quality of reporting randomised controlled trials (RCTs). A report should enable users to understand the trial, and this can only be achieved through complete transparency from authors when reporting an RCT. The main focus of the guidelines is on reporting the trial design, conduct, analyses and interpretation in a way that enhance users capability to assess the validity of the results. The CONSORT statement is a 25-item checklist with recommended flow diagram. One of the items that comprise the statement is item number 23 which recommends that the trial has a unique number to identify the trial and to register the trial under this number.

Furthermore, the WHO ICTRP (2012) is a network of international clinical trials registers that aims to establish a public platform which links clinical trials registers to ensure a single access point, unambiguous identification of trials and to facilitate the public availability and accessibility of trials information. Searchers (in the UK) may also use Current Controlled Trials (http://www.nres.nhs.uk/applications/). Trials registration is considered to be a scientific, ethical and moral responsibility in the clinical field. It is needed to ensure the decision makers are well informed about all of the available evidence, it is important to avoid publication bias, selective reporting of trial results and to maintain transparency of reporting, and moreover, to avoid unnecessary replication of trials. Such enhanced awareness of trial design and procedures employed in trials should make participant recruitment, and collaboration among researchers more efficient, furthermore, such awareness improves the ability to identify the gaps in clinical trials research as well as identifying potential problems early in the research process. As mentioned before, the WHO system is not a clinical trial registry, it is only a platform to link all registries into a single access point and therefore a trial should be registered on one of the registries on WHO network or through compliance with ICMJE.

Up to now, there has been no obligation to register trials and therefore, to make the process more obligatory, ICMJE decided that to accept a clinical trial for publication it should be registered before or at the beginning of a patient's enrolment, as a first step to make trial registration a norm when conducting and reporting an RCT (ICMJE, 2012). ICMJE has a set of guidelines for a clinical trials registry: it should be accessible, searchable and should contain information about the trial such as unique trial identifier, the intervention objectives, hypothesis, primary and secondary outcomes, eligibility criteria, key trial dates, target number of subjects, funding source, and contact information for the principal investigator.

Having a unique trial identifier for each trial can make the searching for RCTs more efficient, where this number can be as an identifier for associated publications that are related directly or indirectly related to a specific RCT. Therefore any economics study, qualitative study or process evaluation study that centred around that RCT should have the same unique identifier, making the search process to locate these relevant studies easier and more efficient. Unfortunately at present there seems to be considerable debate about the standard of compliance with data entry requirements for clinical trial registers and adequacy of registration information in reports of RCTs (Huić, Marušić & Marušić, 2011 & Milette, Roseman & Thombs, 2011). Expecting authors to add the trial number to sibling publications seems over hopeful at this stage.

3.6 Retrieval Effectiveness Measurements

An increased number of information retrieval systems providing their services to users required comparative evaluation of these systems' performance, as well as indicating possible improvements. Information retrieval systems need to be evaluated in terms of effectiveness and efficiency; where effectiveness means to what extent the system can provide the relevant information while suppressing the irrelevant, while efficiency means how the system performs economically (Kowalski & Maybury, 2000; Chowdhury, 2004). System availability, reliability and relevancy judgments are other evaluation metrics that should be considered as well. But still, according to Kowalski and Maybury (2000) information system evaluation is controlled and conditioned by human subjectivity.

Generally, the goal of all information retrieval systems is to retrieve only relevant information; therefore the term relevance will be the major factor to determine the system performance, Recall and precision are the most common measurements used in information retrieval systems evaluation for decades (Raghavan, Bollmann & Jung, 1989). And these of course are based primarily on relevancy judgement.

3.6.1 Recall and Precision

In summary, search strategies can be evaluated in terms of their comprehensiveness in identifying relevant studies. Recall and precision are the measurements to evaluate most search or retrieval process effectiveness, (Shaw et al., 2004). Raghvan, Jung & Bollmann (1989, p. 206) stated that from the users' point of view retrieval systems should behave as follows; *"Retrieve as many relevant items as possible 'Recall', and as few non relevant items as possible in response to a request 'Precision'"*.

However, there is a trade-off between sensitivity (Recall) and precision which means that if the search desires a high recall then the precision will be degraded and vice versa, and due to this, the searcher's needs shall determine the priority of each one when conducting the search, (Wong et al., 2003). As a result much research has been done in order to investigate optimal retrieval systems performance, beside the works that aimed to explore and explain the nature of recall and precision, since most studies point to the existence of an inverse relationship between recall and precision.

Buckland and Gey (1994) investigated the relationship between recall (sensitivity) and precision as they are recognised as the key performance measurements to assess retrieval effectiveness. The authors discussed the theoretical framework through which they explained their perspectives of how to achieve as high recall as possible and as high precision as possible, and whether one of the measurements should be sacrificed on account of the other. Table 1 represent the categories in which each document in the collection fall into:

- Recall = (a / a + b)
- Precision = (a / a + c)
- Specificity = (d / b + d)

Table 1: Two by Contingency Table for Retrieval Performance Calculation

	Retrieved	*Not retrieved*	**Total**
Relevant	A	B	a + b
Not relevant	C	D	c + d
Total	a + c	b + d	a + b + c + d

There were four theoretical cases of recall performance in actual practice. The initial assumption is that in a collection of 1000 records there are 100 records which are relevant. The *perfect retrieval* occurs when all the relevant items are retrieved before the first non relevant item is retrieved, regardless of retrieving non-relevant items as long as all the relevant items are retrieved. Another case when all non relevant items are retrieved first leaving the recall rate at zero, then when the system has left no choice but to retrieve the relevant items the recall rises rapidly to 1.0. *Random retrieval* occurs when for every 10 items there will be 1 relevant item to be retrieved making a straight line from the origin to the final point. However, in reality none of these could be the case. It is expected that recall performance will lie between perfect and random retrieval and that the performance will the best in the beginning and to start to deteriorate as the search is expanded, (Buckland & Gey 1994, pp. 13-14).

Buckland and Gey (1994, p. 15) also discussed the theoretical foundation of the precision also, in *random retrieval* the probability of retrieving a relevant item (randomly, no specific retrieval sequence is allocated) is a reflection of total proportion of the relevant items in the collection in relation to non-relevant items. Thus if 100 relevant items (a) out of 1000 (a+c) were retrieved this means that precision will be only 10%. *Perfect retrieval* of all relevant items will be retrieved first and then the non-relevant will being retrieved, after this point while the recall remain 1.0 the precision will start to degrade in line with the non-relevant items retrieved. As with recall, the realistic precision lies between perfect and random retrieval.

Consequently, the inverse relationship between recall and precision is observed. Therefore, Buckland and Gey (1994, p. 16) suggested a two–stage search process, where at the first stage a high recall ratio is achieved, followed by the second stage where the first stage result is searched again to improve the precision within this subset, and according to their analysis and findings both high recall and precision can be achieved using good search techniques and using a broader arrangement of clues that become available.

However, search strategy and database effectiveness is a matter of the degree of retrieved documents' relevance to the initial enquiry, so relevance judgement criteria are needed to measure relevance and from that to evaluate retrieval process effectiveness; (Borlund, 2000; Janes, 1994).

3.6.2 Relevancy

Relevance judgement is tricky as it is a human judgement which depends on the cognitive situation and the way a user perceives and processes the information. Saracevic (2007a, pp. 1916:1918-1919) believes that, relevance is *elusive, intangible, tacit and can be understood intuitively.* The author described relevance as a relation within two interacting worlds, those of systems and humans, and he stresses that relevance is a human notion and from this comes its strength as well as its weakness. Relevance cannot be explained, it is generally understood, it involves a relation between two types of object, tangible (documents or machines…) and intangible (ideas, concepts or information…) along with some explicit or implicit properties (topicality or utility…). It could be a measurement of relatedness and effectiveness of the relation with a given degree of the strength of relevance relation. Relevance can be created or inferred by systems and users, but not given explicitly.

According to Saracevic (2007a, pp. 1920-1921) relevance judgements have logical, philosophical and communicational bases. Relevance implies an inference that depends heavily on logic; it needs some philosophical skills for interpretation as well. Relevance is also based on communication and cognitive theory where the communication means to draw someone's attention by implying that the information is relevant, it is also proposing that relevance is contextual and comparative, thus the theory provides a number of explanations and operational, predictive principles about cognition and communication in terms of relevance. However, it does not come without weaknesses. Firstly, it uses proof by example to support its arguments, and secondly, the theory has not been tested empirically or experimentally up to now. On one hand, there is still a continuous search for a theory applicable in information science, on the other hand there is a debate that whether the relevancy theory is testable or not, either way, though, it is still untested although it does provide great insights and explanations about practical aspects of relevancy.

Saracevic (2007a pp. 1928-1931) classified relevance into several categories (*manifestations of relevance*) under which relevance may fall:

- *System or algorithmic relevance*: the retrieval algorithm interprets the relevancy as a relation between a request and information objects (documents) in the system which will either return a relevant object or non-relevant object or not retrieve the relevant object at all.

- *Topical or subject relevance*: the relation between the query subject and subject as described and represented by the system, under the assumption that both query and object are related by topic, in other words "Aboutness is the criterion by which topicality is inferred" (p. 1931).

- *Cognitive relevance or pertinence*: "Relation between the cognitive state of knowledge of a user, and information objects retrieved or in the systems file." (p. 1931).

- *Situational relevance or utility*: Relation between the situation, task, or problem at hand, and information objects retrieved or in the systems file. Usefulness in decision making and appropriateness of information in resolution of a problem, are the criteria by which situational relevance is inferred.

- *Affective relevance*: Relation between the intents, goals, emotions, and motivations of a user, and information object retrieved or in the systems file. Here satisfaction and accomplishment are the criteria for inferring motivational relevance.

Saracevic's categorisation of relevancy according to Borlund (2003, p. 915) can be categorised into a broader notion depending on the judge of relevancy. Consequently Saracevic's first category belongs to system relevance, while the four latter categories refer to subject or user relevance. Schamber et al. (1990) cited in (Borlund 2003), concludes there are three characteristics about the nature of relevance. They contend that relevance is a (1) cognitive concept that depends on users' perceptions and their situational information needs, which (2) dynamically evolves as the search process proceeds and relevance judgments are reached at certain point of time, yet (3) relevance is complex to infer. However, users' knowledge and situational and cognitive state are the only measures for relevancy to lessen the complexity, and this can be understood from Saracevic (2007b) as well.

Saracevic argued that relevance is *dynamic* and *situational.* It changes as the users' cognitive state or the associated knowledge, and cognitive changes evolve. And accordingly, users change criteria for relevance inference at each retrieval stage as the cognitive state of users changes as well as the cognitive knowledge state, thus the user may become more focused which strengthens the relevance inference and judgment, (Borlund 2003, p. 920; Saracevic 2007b, p. 2130). (They might also get very tired and miss items towards the end of scanning a large retrieval output!).

The context, situation and circumstances under which the information or information object was retrieved play a major role in this property of relevance (Saracevic 2007a). In addition to being a set of interdependent and interacting layers in which the relevance is the foundation, information retrieval is composed of several layers that interacted with each other during the retrieval process where the relevancy is the base for the retrieval. This involves both system and human factors which reflects a multi-faced nature for relevance.

Comparing the proposed Buckland and Gey (1994) system and what Borlund (2003) and Saracevic (2007) suggested above, Buckland and Gey needed to increase number of retrieved records to achieve a better performance, so they suggested the two stage search where in the second stage they narrow down the search using broader arrangement of clues that became available to enhance system precision. This reflects the nature of relevancy judgment which Borland and Saracevic discussed, the dynamic dimension of relevancy, so as the search advanced over time, more relevancy clues become available which will help with a more focused search and relevancy judgement.

Other factors that play a major role in relevancy are the presence or absence of relevance judges and relevance judgment. Relevance judges are often considered as domain experts and therefore bring in related factors such as experience and domain knowledge. The implication is that even though experiences cause differences in relevance inferences by a group of judges, the overall higher expertise among the group results in higher agreement, fewer differences and stronger inferences - individual peculiarities are ironed out. There are a lot of assumptions being made here – as some studies have indicated (e.g. Individuals make different relevance inferences to the group; individual cognitive, affective, situational, and other related variables are the main source of such variability) (Jane, 1994 Hripcsak 2002 & Saracevic 2007b). Expert and non-expert judges may make the same decision, the only difference is the time needed to make the decision (Nunn, 2008).

The main factors to consider when thinking about relevance judgment as the Saracevic (2007b, p. 2137) suggests are; first, topicality, where it considered as the primary factor that influences inferences of relevance of information or information objects. Second, measures and measuring of relevance inferences; using binary measurement as relevant and non relevant seem to be suitable but users might subdivide relevance judgments into further levels; middle, and high relevance assessments, where middle category objects are to be expected to be the prevalent category among other categorisations. Third, the degree of independence; where the objects are to be judged dependently or independently of each other. However, order and presentation of previous judgments have an effect on the successor judgment, and consequently, the more information is added the more the relevance judgments change, for example more information will be gained as the users advance from titles, to abstracts, to additional representations and then full texts. Furthermore, the consistency of relevance judgments depends on judges' experience and number of judges, thus "*higher expertise results in higher consistency and stringency while lower expertise results in lower consistency and more inclusion* (p. 2137)".

Xu and Chen (2005) believes that the relevance concept encapsulates topical relevance (aboutness) as the first or basic condition of relevance, followed by the processing and analysing of the document involving cognitive relevance, leading in the end to situational relevance (usefulness and appropriateness which involves both social and cultural factors). Xu and Chen (2005) applied Grice's (1975, 1989) theory –communication theory – and identified five core relevance criteria to be used as basics in relevance judgement: *scope, novelty, topicality, reliability and understandability*.

However, before getting to the relevance judgements, it is important to understand how users may approach searching, particularly of databases. Users construct search query(s) based on their previous knowledge and experience of the topic and accordingly they choose the appropriate database and search terms. Deciding on database and search terms might be recognised as an early relevance judgment. It is expected that users' search practices tend to be different. For example, PubMed daily query logs showed that a quarter of PubMed queries were navigational - using author information or citation information or both – while three quarters of queries were informational – using textwords or implicit MeSH terms. The search queries varied in length and topics in accordance to users' needs. However, it appears that users prefer short queries (median of terms per query was three terms). Users equally used narrow and broad search queries with a very rare usage of MeSH terms (Herskovic et al., 2007).

3.6.3 Database Performance

Database performance is usually measured using recall and precision as the key measurements. However, Stokes et al. (2009, p. 2) think that other performance measurements can be utilised to look deeper into their performance, such as databases' effectiveness, uniqueness, coverage, novelty and accessibility in addition to recall and precision. The comparative performance of the BNI, CINAHL, MEDLINE and EMBASE databases was investigated in terms of providing the most useful information for nursing and midwifery students using recall and precision, and odds estimator measurements. The odds calculations in this study used the pooled recall as the basis of the calculation – an odds estimator would be a better description (rather than the odds ratio used in the article) to avoid confusion with the use of true odds ratios in so much health services research.

Nine students who were registered for nursing degree (n=2), midwifery degree (n=2) and those who were following a Continuing Professional Development degree (n=5) participated. All participants provided a title or a specific subject area of their dissertation proposal which was the root of their search strings. The search method used was title search only using specific keywords in order to standardise the search process to reduce the bias and overcome indexing differences (due to inherent differences in databases' representation and processing standards). Search results were screened by participants for the relevancy judgments but the possibility that the participant may lose interest after reviewing a certain number of articles should be preserved in mind, (Stokes et al., 2009, p. 3). The search method used here was based on the results of a pilot study, examining whether year limits were necessary for title (only) searches.

According to Stokes et al. (2009, pp. 4-5) databases can be evaluated using the following six criteria: Recall, Precision, Novelty, Originality, Availability and Retrievability, the main measurements for each database to be analysed are:

- Effectiveness (relevant articles), where effectiveness *is a combination of Precision and Recall based on Relevancy*; relative recall was used here.

- Efficiency (unique articles), where it is *a combination of Novelty and Originality based on Uniqueness*. 'Novelty' refers to the proportion of relevant items retrieved and unique for the search strategy. 'Originality' refers to the proportion of relevant items retrieved and unique per database.

- Accessibility (obtained articles) *and it is a combination of Availability and Retrievability and is based on Obtainability*. Availability is the proportion of obtainable and relevant items retrieved per search. Retrievability is the proportion of relevant items retrieved that are obtainable (from database) to the total number of obtainable relevant items.

Stokes et al. (2009, p. 6) found that databases perform differently based on the topic being searched. For example EMBASE and MEDLINE performed better than BNI and CINAHL when the search did not contain nursing search terms, however, when the search became more focus on nursing topics EMBASE performance degraded while MEDLINE still performed well.

Precision, Originality and Availability were statistically significant at $P < 0.05$ level. Recall and Novelty were not statistically significant but indicate considerable differences. Retrievability was not significant at all. According to that, the findings support the argument about the existence of differences in precision, originality and availability between the databases; while it rejects the other arguments about the existence of differences in recall, novelty and retrievability, which do not differ significantly among databases, Stokes et al. (2009, p. 7).

In the odds estimator analysis for each database, in term of database effectiveness (relevant hits odds) BNI achieved the highest odds estimator retrieving few non-relevant hits, although CINAHL retrieves the higher proportion of relevant hits but it scores the second best odds estimator indicating a higher portion of non-relevant items compared to BNI. The other databases lose effectiveness since they retrieve a higher proportion of non-relevant items than nursing databases. In terms of databases' efficiency (the odds of retrieving unique and relevant items), CINAHL scored the highest odds estimator indicating the higher likelihood of retrieving relevant and unique compared to the other databases, at the same time MEDLINE and EMBASE retrieved the highest number of 'unique/relevant' items, but at the same time they retrieved a high rate of non-'unique/relevant' articles causing their efficiency to drop down. For Accessibility (the odds of retrieving an obtainable and relevant item), CINAHL and BNI have a high accessibility rate compared to MEDLINE and EMBASE, Stokes et al. (2009, p. 8) (but accessibility will vary according to institutional setting and subscriptions).

In conclusion, it was not an easy task to obtain a firm result to tell which databases are considered to be the most useful, but the approach used here helped to assess which database performs well in general without ignoring the fact that all the databases are likely to miss some relevant articles (from the manner of indexing or the way the search strategy is constructed, or simply faults in data entry). Moreover, database performance depends on the subject being searched for and therefore, the performance patterns detected here might not be the same with other subjects and this limits the generalisability. This research approach can be helpful in evaluating the comparative performance of databases when reviewing subscriptions or in other collection management decisions.

3.7 Text Mining in Health Research

Tracking the literature reveals the huge amount of available textual knowledge about every single research topic, more than anyone can possibly read or digest. One cannot deny or ignore the great role all these documents play in scientific discovery and evolution as they can be considered as a huge information repository that directs the wheel of science development and implementation. Electronic repository and records representation make the search for needed records and retrieval process more systematic and as reliable as possible, but one search query may retrieve hundreds to thousands of textual records on a single topic which make reading or reviewing these records exhausting and time consuming, (Sainani, 2008)

Text mining uses computers and technology to identify, extract, manage, integrate and exploit knowledge. Text mining involves analysing documents to discover relationships or patterns that are buried in the document collection and which would otherwise be extremely difficult, if not impossible, to discover. Text mining (TM) is based on natural language processing, information retrieval, information extraction, and data mining techniques to help in collecting, maintaining, interpreting, and discovery of the knowledge needed for research development and efficiency, (Ananiadou, Kell & Tsujii, 2006).

Ananiadou et al. (2006) noticed the increased attention on text mining techniques which are able to retrieve hidden knowledge and discover possible associations and patterns in texts. Systems biology is considered one of the areas in which text mining is effective due to its nature as an interdisciplinary subject, which involves iterative interplay between computational modelling, high-throughput and high content experimentation, and technology development, besides collating knowledge from wide areas of biology.

Text mining targets unstructured data to extract knowledge using three major steps, (Ananiadou et al., 2006, pp. 572-573:575):

1. Information retrieval (IR), in which relevant text and documents are needed to be defined and retrieved using ordinary search techniques.

2. Information extraction (IE), to identify and organise the extracted data or facts into tables without having the user read the entire document.

3. Data mining (DM), used with structured data to discover indistinct associations between the known facts extracted by IE. At this step text mining (dealing with unstructured data) and data mining integrate together.

Systematic review researchers are confronted with a huge amount of records that need to be read and reviewed to produce strong and reliable SRs (Ananiadou et al., 2007). With this data deluge, text mining techniques appear to be helpful in the creation of systematic reviews to ease the reviewing process. Ananiadou et al. (2007) at the National Centre for Text Mining (NaCTeM) described The Automatic Summarisation for Systematic Reviews using Text Mining (ASSERT)[♠] project and how text mining techniques can help in systematic reviews production.

Systematic reviews have their established standards and steps in doing the review, subsequently text mining can help in searching, screening and synthesising in systematic review (Ananiadou et al., 2009, p. 3).

♠ For more information visit: (http://www.nactem.ac.uk/assert/)

According to Ananiadou et al., (2009, p. 3), "Searching can be improved by using query expansion techniques based on the most important concepts (terms) similarities among terms but also ontologies and thesauri".

The ASSERT project works on extracting the most significant terms from a collection of documents using NaCTeM's TerMine service which, in turn, extracts and automatically ranks technical terms by assigning weights and produces a ranked list of documents similar to the original document. Clustering finds the set of representative terms with their association, and assigns each cluster a distinctive label based on contents. While document classification "identifies the underlying patterns and distinguishing features within documents that make them part of a defined grouping or class and uses this information to assign each new document to known classes." (Ananiadou et al., 2009, p. 6).

Query expansion calculates the similarity between documents (documents in collection and documents in user query) and then adds all important concepts and related documents (keywords) to the original query. And therefore;

> Screening can be improved by using *document clustering* which groups documents into topics (...) *Document classification* automatically assigns documents into existing categories, generating subsets of documents focused on a specific topic, allowing for more efficient and accurate analysis during subsequent stages of information filtering (…) Multi-topic classification is useful for systematic reviewing as single documents may be relevant to multiple review topics. (Ananiadou et al., 2007, p. 3).

This aims to limit the collection of documents to only the relevant and high quality documents and then clustering documents into clusters corresponding to a single topic that is shared by all the documents in the collection to improve the process. Moreover, visualisation enhances the associations between documents and topics. The key procedures which play a major role here are those for document classification.

For synthesis, this still needs the human factor as the logic and information processing is handled differently by humans and therefore the process can not be fully automated. However, text mining can facilitate the process so that it can be done quickly, as in the screening process the system already returns the most relevant records. Clustering and classification have highlighted the distinguishable features, patterns and concepts and help to assign documents to classes based on those, and to visualise the classes in a user friendly interface, and the user can use more classification criteria and feed those into the system so the system can do another search, with more filtering. The search process is iterative and can be refined according to users' requests and definition of more refined topics (Ananiadou et al., 2009). It produces a summary of documents based on the most significant terms of retrieved documents and then chooses the most informed sentences.

Synthesizing works on correlating and summarising evidences from several resources. 'Multi-document summarisation' is the technique proposed to improve the process. Sentences from each document are selected based on the significance of the term it contain, then classification techniques use these sentences to discover the most relevant passages within the important sections of a document i.e. *introduction, background, methodology, results, conclusions*.

Finally, it has been concluded that systematic reviews with the aid from text mining could proceed quickly and systematically, as TM tools improves searching, screening and synthesising. The process is now semi-automated and with further investigations it should facilitate the creation of systematic reviews by providing a "robust, scalable, efficient and rapidly responsive services for very large collections and the need to consult large-scale resources." Text mining tools could and should be expanded to go through the whole document, not just the abstract, as the information it contains is less than half of the actual information.

Unfortunately, text mining is currently limited to MEDLINE only. This limits search options when conducting systematic reviews, although MEDLINE is one of the leading biomedical databases. It is well established that searching MEDLINE only cannot be considered sufficient for comprehensive retrieval. Therefore, extending text mining to other databases is required; especially as the health services field has a social element that is covered by other databases that have more coverage in social sciences such as SCOPUS. Additionally, biomedical literature is multidisciplinary and accordingly text mining tools need to be flexible to handle different formats of information that are available.

3.8 Summary

Information retrieval systems retrieval performance based mainly on subject indexing and representation of documents. User information needs are interpreted into search query and databases return a set of documents that match the query as presented. Obviously, health professionals are likely to expect that the most relevant results to be presented first, as in a Google search output and therefore, relevance ranking techniques are needed. Meta-search is a practical tool in information retrieval systems, as it searches different information resources with different interfaces and contents representation, maintaining interoperability between them and returns the search result to a single virtual interface without users' awareness of the mechanisms involved.

Retrieval of RCTs has been a priority and progress has been made on the development of search filters to reach an optimal recall and precision. Many search filters are available. The Hedges team has developed filters for different types of study, designed specifically to work with specified database interfaces, i.e. MEDLINE. For economic, qualitative and process evaluations, there are different factors to consider when planning a search.

Database performance is another field of debate. There is no standardized approach to measure specific database performance and productivity, and different interfaces for the same database can produce different results for recall. Statistical approaches based on recall and precision calculations are the basic performance measurements for retrieval performance. Relevance criteria are generally reported by recall and precision. Yet, the actual relevance judgement is a tricky human judgement which depends on the cognitive situation and the way a user perceives and processes the information. Primarily, the concept of relevancy encapsulates topical relevance, although the perception of topicality is dynamic which seems to grow and evolve as more information become available.

Clinical trial number registers seem a practical solution when searching for related studies such as siblings, but these are not yet used to a useful extent. Text mining tools are a new trend in health information retrieval field. These aim to save time and effort when conducting SRs. However text mining still has limitations especially when it comes to social science aspects of systematic reviews i.e. qualitative studies.

Chapter Four
Methodology

4.1 Introduction

This chapter introduces the methods used in this research, the selection of the seed studies, the search strategies and databases used and investigated, and how data was prepared for screening, with consideration of the metrics for analysis. Research hypotheses and procedures for analysis are also presented. The methods used in this research reflect some of the debates discussed earlier in the thesis (Chapters two and three). For example:

- Information retrieval evaluation still tends to focus on recall and precision as the main measurements (although there are other measures, e.g, the indicator based on the odds estimator (Section 3.6.1)). Many of the IR studies in health, and associated with systematic reviews use recall and precision and for continuity, this research uses recall and precision metrics.

- Recall and precision are very dependent on the definition of relevance and even the "topical aboutness" (Sections 3.6.2) is probably not as objective a measure of relevance as some researchers assume. The information behaviour research reminds us about the dynamic nature of relevance, the importance of the user perspective(s) on relevance. Chapter 2 demonstrated that there are very different priorities for those interested in economic evaluation (Section 2.3), or qualitative evaluations (Section 2.4), and the term "complex interventions" (Section 2.5) is a good description of the problems of trying to integrate these different perspectives to provide informed decision making. This research study is primarily a technical investigation, but it is also important to investigate measures of relevance that may be more important to policymakers and practitioners than to well resourced systematic review teams.

- As the TREC series and the development of digital libraries, all demonstrate, the development of probabilistic retrieval systems, with the most relevant items appearing first (as in a Google search) deserve more research (despite the problems of not knowing how the underlying algorithm might work, and the fact that such algorithms, such as “related articles” searching in PubMed MEDLINE, are continually evolving). For that reason, emphasis is placed in this research on the use of a probabilistic ranking criterion (odds estimator) as well as exploration of federated searching.

- Text mining has been mooted as a way of helping systematic reviewers deal with large search outputs. A contact with the Nactem team in Manchester was made in July 2009 and their work on text mining was investigated (Section 3.7). Text mining might have been able to provide some help to systematic review teams, however, it is still limited to searching MEDLINE only and most of the performance analysis is based on its collection. This limits any further exploration of text mining as the nature of this research required searching databases other than MEDLINE/PubMed. Moreover, the way text mining carries out the analysis and categorisation is not compatible with this research categorisation, as the main focus of categorisation is based on study design and methods.

- Search filters have been developed in order to improve retrieval of study with specific research design. The Hedges team (Section 3.5.4) have contributed a lot to designing search filters for databases such as MEDLINE. This research will investigate the role of search filters in retrieving sibling studies.

- Indexing and reporting of studies has a big influence of search strategies, search filters and databases retrieval performance. Chapter 3 demonstrates issues with subject representation and difference among different databases (Section 3.2.3) and searching for evidenced based practice (Section 3.5.1).

As outlined earlier, the aim of this research was to investigate and compare the performance of different search strategies as well as the performance of different databases in retrieving direct and indirect siblings of RCT's. The research was conducted over two phases; the first phase was a pilot study that aimed to investigate some proposed search strategies and databases. In the second phase search strategies and databases were investigated more thoroughly in three sub-phases. The results of each sub-phase were screened using various inclusion criteria to identify relevant studies, direct and indirect siblings. The performance of search strategies and databases were determined using recall and precision, the odds estimator (Section 4.5) was used as a rank indicator to rank each search strategy or database according to its retrieval performance.

4.1.1 Resources to Locate Existing Literature

Procedures used for the literature review were based around use of the Google scholar search engine, as from that there were links to the databases and e-journals collections to retrieve the required articles. The link resolvers put in place by the University library service made this the easiest way of finding and retrieving full text material for general topics, to ensure that searching for this type of literature was as broad as possible. In addition, PubMed, and CENTRAL were searched for some clinical articles using the broad subject terms i.e. search strateg(y)(ies) AND qualitative research; qualitative search AND systematic review. As already noted, MEDLINE is not always easy to search for non-clinical topics. Once a good article was retrieved, snowballing of the reference list helped to find other articles that might be useful to retrieve. Other documents were recommended through personal contacts (e.g. my supervisor) and the reference list of another thesis on systematic reviews was examined. The Aberystwyth University network was used in order to gain access to most of databases and required literature i.e. LISTA and LISA for the library and information science literature. In other words, the main search strategies used to locate literature for this research were subject search using Boolean operators as seen appropriate, PubMed related articles and reference lists of some studies were reviewed. Background material on information retrieval was obtained from key textbooks (e.g. Chowdhury 2004) to complement, and contextualise the more recent material that could be obtained from journal articles. For the background literature on information retrieval and systematic reviewing, considerable emphasis was placed on snowballing of reference lists.

Journals regularly searched included *Health Information and Libraries Journal,* and *Journal of the Medical Library Association.* In addition, regular checks were made for publications by prominent members of the Cochrane IR methods group. The literature was regularly scanned for any new new studies to keep the literature up-to-date.

The identified studies and associated evidence were sought in order to be able to cover all aspects of the subjects discussed in the thesis. A Word file was created to store bibliographic details of the references obtained. The literature review work was ongoing, with periodic updating of the literature. The main literature review work ended for publications dated 2011, however a supplementary search on some topics for 2012, mainly search strategies and search filters, was done in order to maintain an up to date focus on these topics.

4.2 Seed Study Selection

Early pre-pilot investigations, and opinions sought from experts, indicated that qualitative and process evaluations were far more likely to occur alongside larger randomised controlled trials, at the Phase III stage. The outcome of discussions with members of the Cochrane IR Methods group was that we all believed initially that it would be possible to distinguish easily between direct siblings, that had a close relationship with a particular RCT, with the study conducted by one of the team or in close collaboration with the RCT team, and indirect siblings, that provided relevant evidence but were conducted independently, with no apparent association with the RCT. The group supposed that some direct siblings might share authorship with the RCT, but then some instances were identified where it was difficult to confirm a direct relationship, particularly when the sibling study had no authors in common with the RCT (Section 2.4). This influenced the choice of studies to use for the first stage of the research.

In collaboration with the Cochrane IR Methods Group, a range of studies with different characteristics, topic area, date and which had some known siblings as well were chosen:

- Telemedicine and diabetes (a known RCT, with many known direct siblings, that could be used to validate and checking the search strategies for their sensitivity), Shea et al. (2006).

- IVF (group collaboration for the RCT – multi-centre trial, not known if direct siblings existed), European and Israeli Study Group on Highly Purified Menotropin versus Recombinant Follicle-Stimulating Hormone (2002).

- Dexamethasone for chronic lung disease prevention in infants (a collaborative clinical network was involved, of interest for retrieval by author or group name, with economic and process improvement aspects of importance), The Vermont Oxford Network Steroid Study Group (2001).

- Tamoxifen for breast cancer prevention (known qualitative sibling, two RCTs involved), Fallowfield et al. (2001).

- Breast cancer therapy using body-mind-spirit group therapy (RCT, but of qualitative research, with Chinese researchers, of interest for retrieval as many Chinese researchers share the same family name in English, thus complicating author name searching), Liu et al. (2008).

As the aim of this research is to investigate and compare different search strategies' performance as well as the productivity of different databases, it was necessary to start with some trials known to have direct siblings. This provided a type of standard for comparison of search strategies.

4.3 Search Strategies and Databases

The following types of search strategies were used:

- Related article search in MEDLINE on PubMed (for the seed article).

- Cited reference search (with the seed article as the reference) on ISI, Web of Science, SCOPUS and CINAHL.

- Simple Author-Subject search (using a very simple combination of subject term, text terms, and the seed study authors) on PubMed, SCOPUS, CINAHL, E-Library♠ (with combination of British library integrated catalogue, Centre of research libraries, Directory of open access journals, Scirus (Elsevier), Science direct (Elsevier), Index of theses, Theses & Dissertations catalogue, Intute and Intute: Medicine) and web of knowledge (WoK). In all cases the search results were refined by author names, but BMS – body mind spirit group therapy - was an exception as the initial subject search retrieved a very small number of studies, which meant the refinement will contain the seed study only and therefore only subject search was employed in this case with no use of any author names.

It should be affirmed that WoK was used in general author-subject searching while WoS was used for citation searching only.

Author name variation is a serious issue which has implication on retrieval performance, especially in this research where authors are the main link between siblings. This issue will be addressed in this research by BMS seed study where Chinese authors names appears (many authors share the same last name (family name)).

- Simple subject search on an e-library♣ (with a combination of ISI (WoS), OCLC WorldCat, OCLC Articles First, EBSCO Business Complete, and EBSCO International Bibliography) using a simple subject term combination only. We chose 300 as the stopping point as after 200 the number of relevant studies retrieved started to diminish rapidly with the Meta-lib search engine used.

♠ This refers to the second selection of databases to be searched as a MetaLib search.

♣ This search was conducted in the pilot study. The results suggested a different choice of databases as indicated in the author-subject search on E-Library above.

The simple search terms were a selection of key terms which describe the main subject area taken from either the title or abstract. The main questions asked, in order to derive the key terms, were: what is the intervention or interventions, and what is the condition being treated? Subsidiary questions considered whether the patient group was important, and what the role of the intervention might be. The mean length of search string was 3 terms, extracted from the seed study title and/or abstract. The mean length of search string with search filters was, seven, eight and nine terms for costs, qualitative and economics filters respectively. Search terms were then combined with names of authors of the seed study and limiting the search to retrieve studies published between 1992 - 2010. This time to publication range was chosen as the seed studies were published in 2001, 2002 and 2006, so a window of 10 years before and after was chosen for all seed studies based on the earliest published seed study.

The rationale behind choosing to use a short search string in this research was that the siblings have different research designs and methodology, different objectives and different reporting style and therefore it would be better to use broad key terms that are expected to be common to all the siblings.

According to Cheng et al. (1998), Hopewell et al. (2007) and Takeda et al. (2008) the mean time to publication for the full report of RCT is between 4-5 years for positive trials and 6-8 years for negative trials. Therefore all records published before 1992 were excluded from the retrieval list as the seed studies were published between 2001 and 2006. Taking approximately 9-10 years time span either side of the seed study should be sufficient to pick up sibling studies. This was the base of the choice of publication date limit that was used in this research, though for the 2006 seed study the stop date 2010 was close to the limit. It should be emphasised that publication date may not be the same as the chronological date of actual publication. Many studies ostensibly dated 2010, for example, may in fact appear earlier, and the appearance of online early publications complicates matters further.

4.4 Phases of the Study - Overview

This study is based on 4 different retrieval phases where each phase focuses on retrieval of specific type of studies, for example phase one is conducted as a pilot study to assess search strategies and databases retrieval performance (section 4.5). The following phases were designed and refined based on the pilot study retrieval performance. Phase two (section 4.7) aimed to retrieve relevant studies with a possibility of being sibling studies (direct or indirect). Phases three and four (sections 4.8 & 4.9) used the pooled relevant results from phase two to identify indirect and direct siblings respectively.

In the second phases (with the three sub-phases) of the research Hedges search filters (section 3.5.4) were combined with either subject search or author-subject search on PubMed, WoK, CINAHL, SCOPUS and E-library (the second selection of databases) as seen suitable as following:

- Qualitative search filter to retrieve qualitative siblings.
- Economics search filter to retrieve economics siblings.
- Costs search filter to retrieve economics siblings.

With some seed studies there is no point in refining the search results by author names as this will restrict the retrieval results to no useful effect (Appendix one). Hedges filters are initially designed for MEDLINE/PubMed, EMBASE and PsycINFO. However the filters were used on other databases as other databases have a link to MeSH terms one way or another. For example, CINAHL uses the CINAHL subject heading which is based on MeSH but with added domain specific terms while SCOPUS uses both MeSH and Emtree terms for indexing. Table 2 shows the dates when each database was searched:

Table 2: Databases Search Dates

Database	Date searched
PubMed	30 November 09
PubMed with Hedges filter	27 October 2010
PubMed Sophisticated	14 November 2011
Citation(Web of Science)	24 December 2009
e-lib-Author – subject	24 December 2009
SCOPUS	20 March 2010
SCOPUS with Hedges filter	06 April 2012
CINAHL	25 March 2010
CINAHL with Hedges filter	04 April 2012
WoK & E-lib	25 May 2010
WoK & E-lib with Hedges filters	18 November 2010

Despite the different dates when each database was searched, the stop date to include sibling studies in the study was December 2010 in all databases searched. It is worth mentioning here that much of the literature is now published ahead of time. For example, something published in 2010 may actually appear in 2009. Moreover, many articles are published online ahead of their nominal publication date.

It is important to compare a simple search strategy with the more usual type of sophisticated search strategy used for retrieving RCTs for systematic reviews. Therefore sophisticated search strings were prepared for each seed study, and were conducted on the PubMed database to compare a simple search performance with sophisticated search performance. For each seed study the PubMed, MeSH database was searched for the most appropriate MeSH terms for each seed study and then these were combined with the appropriate text terms and specific search filters. The sophisticated search strategies formats were:

- Author-subject search for all seed studies.
- Author-subject search with process evaluation/qualitative MeSH terms.
- Author-subject search with Hedges costs filter.
- Author-subject search with Hedges economics filter.

In fact the simple search for PubMed is not "simple" in the sense that few terms are actually used. PubMed matches search terms input against the MeSH translation table, a Journal translation table, the Full author translation table, Author index, the Full investigator translation table and an investigator translation table using automatic mapping feature. It adds the appropriate MeSH terms and expands the search accordingly (See Section 3.5.3).

The rationale for testing these approaches was that some of these strategies exploit the strength of Boolean searching (with AND/OR/NOT) to refine the search. In some databases, author names were used to limit the retrieval, but this assumes that the direct sibling studies will at least share one author with the seed study, and that cannot be guaranteed, of course, but it seems the most logical approach to refining a large search output for subject searching. Therefore when the retrieval was relatively small the author names were not used in the search, and only used when the retrieval output was large (over a thousand). This assumes, of course, that any direct siblings that do not share author names, will be picked up from scanning references of direct siblings or seed studies that do share author names (See Section 7.3.4).

Related articles searching uses a different (and evolving) methodology. The PubMed algorithm uses term frequencies to determine the topicality relevance which they refer to as *pmra*, the topic-based content similarity model (Lin & Wilbur, 2007). The information about SCOPUS and CINAHL related article searching is much less clear. The rationale for using related article searching was that this approach might find siblings if the shared subject content was similar. Related article searching (PubMed) and subject searching (with the clinical query filter) are considered to be a useful combination for searches used in updating existing systematic reviews (Sampson, 2009).

Citation searching assumes that siblings (direct or indirect) might cite the specific seed article of a high quality study (RCT) or be cited by other siblings and accordingly it will be identified and retrieved. As aforementioned, several sources use citing reference; SCOPUS, CINAHL and WoS, and some studies reported that SCOPUS yields more citing references than ISI in the health sector (Sampson, 2009).

Given the popularity and growth in availability of meta-search engines for databases, a MetaLib search was conducted as well using the Aberystwyth University interface and a selection of the available databases, (Section 4.3). It is to be assumed that meta-search might result in good retrieval searching several carefully selected sources with a good utilisation of time (Coiera et al., 2008).

Additionally, the reference list of each seed study was examined. This might identify sibling studies either direct or indirect that other search strategies might fail to pick up.

Authors of the seed studies were contacted by email to see if they provide any additional information that may help to identify sibling studies. After the first phase of searching, an information sheet was prepared and sent to a contact author of each seed study along with a list of relevant studies (all relevant either siblings or only relevant), to ask for their opinion of the list provided, whether the set of siblings provided are really siblings and if they can recommend any other sibling study(s) to be added into the sibling lists (Cheng, 1998; Hopewell, 2007; Takeda et al. 2008). One author (for IDEATeL, telemedicine and diabetes RCT) provided a complete list for comparison. The list was investigated in detail in order to derive a theme that might define the sibling relationship. Shared authors with the IDEATeL seed study was the only obvious grouping factor and therefore was used as an indicator of the sibling relationship for the other four seed studies (Section 4.9). After I created the siblings lists for each seed study, authors of each seed study were contacted (if contact information was obtainable) and provided with the list of direct siblings created and an information sheet to explain the research problem and objectives (See Appendix Nine) to confirm whether the siblings list I identified using relative recall matched their opinions.

The entire search results were uploaded and stored into the Web End Note reference management program, for organizing and screening proposes. Search results were organised in different folders according to the search strategy and databases being used to retrieve that set of records and then all results sets were reviewed and duplicates and any record published before 1992 were excluded from that dataset. In addition, all of the retrieved records (from all search strategies and databases for x_i seed study) for each seed study were pooled creating a pooled retrieval list. Again, duplicates were processed and removed from all retrieval datasets, and finally the datasets were ready for screening.

For each seed study, a record (study) was eligible if it describes the main subject of the study, addressing any aspects that are related to the seed study main interests, with the emphasis on “topical aboutness”. For example, any RCT that is done within the same scope as the seed study, a qualitative study discussing issues associated with the seed subject topic, a study investigating effectiveness and cost-effectiveness, or process evaluation. This is the general frame that was used in the first screening to narrow down the retrieval results into a dataset of relevant (about the topic) studies that might be direct or indirect siblings for consideration in the second screening process. After that another list of both direct and indirect siblings were pooled. In the end a complete list of direct siblings was pooled from all search strategies for each seed study (Figure 4).

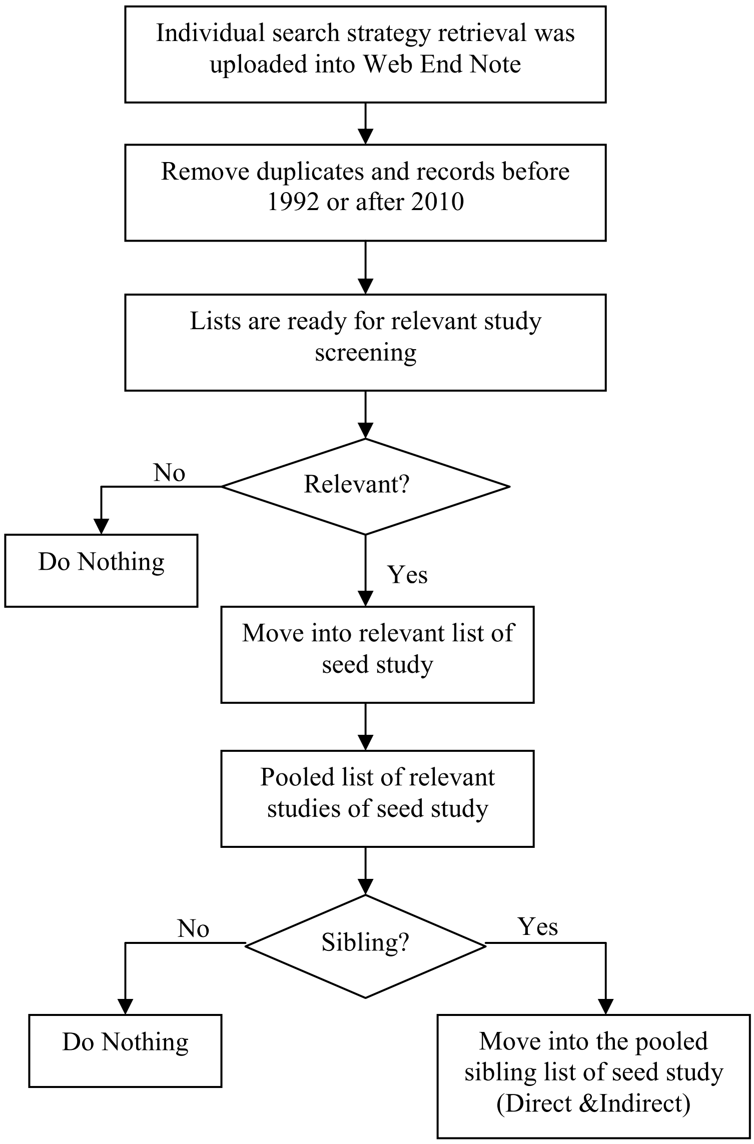


Figure 4: Screening Process for Sibling Studies Identification

The pooled retrieval lists for the sibling studies acted as the gold standards, for performance metrics. This approach differs from the conventional method for gold standard creation in some systematic review database evaluations, but was necessary in this case mainly due to the inclusion of federated searching. The approach used is referred to as relative recall: it is comparing a proportion of articles retrieved from one database to a pooled list of relevant articles retrieved from all the databases under investigation (Sampson et al., 2006c; Hoogendam et al., 2009) and this approach is now being used as an alternative to a conventionally created gold standard. According to Sampson et al. (2006c) several authors have advised the use of relative recall, though the term relative recall is not always used to describe the method. Hersh et al. (1994) used this method to asses the performance of SAPHIRE algorithm based on MEDLINE, as it was difficult to know the exact number for the relevant collection of documents in the MEDLINE, and a pooled collection of relevant documents were created from three search engines (Hersh et al., 1994). Some authors have suggested that this method, the relative recall, might overestimate recall when some studies were missed by search(s). However they claim that there was the benefit of a more generalisable performance to gain, as this is a more realistic reflection of the real world situation where the possibility of a miss is unavoidable. Sampson et al. (2006c) stated that relative recall appeared to be a good alternative to the conventional gold standard.

All the titles of retrieved records (and abstracts where available) were scanned to identify the relevant (on topic) studies which could be considered as siblings in the final stage of siblings identification. Ideally, the screening process is done by two or more investigators to reach consensus on whether a record is eligible for inclusion. In this research, the relevance judgements made by two researchers [myself, and later by my supervisor] checking against each seed study's subject, aims and objectives. This relevance judgement (paired) was done for each phase, the initial on-topic relevance screening to identify the relevant pool of possible siblings, and the later phases (See Sections 5.2; 5.3; 5.4; 6.2).

The limitation of the screening process to be addressed here is that the abstract of some of retrieved records was not available, making it difficult to judge the record as relevant or non-relevant. Accordingly those records with no abstract available were excluded if the title did not explicitly imply its relevancy.

4.5 Metrics for Analysis

Recall, precision and a modified odds estimator (indicator) were chosen as the performance metrics to be used in this study. Recall represents the number of relevant and retrieved, divided by the number of all relevant studies (pooled relevant studies). Precision represents the number of relevant and retrieved studies divided by the number of items retrieved by the search strategy, where the total retrieval varied among search strategies and databases (Stokes, Foster, & Urquhart, 2009). These two metrics, as indicated earlier, are the standard measures used in assessment of search strategies for systematic reviews, as conducted by members of the Cochrane IR Methods Group (See Section 3.6.1, Table 1).

The odds are a way of representing probabilities. It means the ratio of the probability that the event in interest occurs to the probability that it does not. Odds ratio (OR) is a measure of association between an exposure and an outcome. In health research odds ratios are mostly common in case-control studies, but that does not mean that OR is not used in cross-sectional and cohort designs however in these two cases some modifications for the assumptions are needed. The example below helps to explain the OR calculation.

Example:

In the study, 200 health professionals were assessed for the use of Boolean search strategy on MEDLINE after receiving formal training (n=100) or informal training with a help sheet (n=100). Six months later, of the trained group, 25 health professionals used a Boolean search strategy. The remaining 75 health professionals did not use a Boolean search strategy. The other 100 health professionals received informal training from a help sheet only. Of this group, only 10 health professionals used a Boolean search strategy after six months.

In this example the exposure is training in use of Medline and the outcome is use of Boolean search strategy after 6 months (of formal training or help sheets only). And accordingly what are the odds (ratio) of using a Boolean search strategy after receiving formal training compared to informal training. OR will be calculated using the following formula:

$$OR = \frac{A*D}{B*C}$$

Where:

A: Trained and using Boolean search strategy after 6 months (on retest) = 25 (++).

B: Trained and not using Boolean search strategy after 6 months (on retest) = 75 (+-).

C: Not trained and using Boolean search strategy after 6 months = 10 (-+).

D: Not trained and not using Boolean search strategy after 6 months = 90 (--).

Odds ratio is derived as the following:

1. The probability of health professionals receiving formal training and using Boolean search strategy is 25/100 (0.25) and the odds is 25:75 (1:3) = 25/75.
2. The probability of informally trained health professionals using Boolean search strategy is 10/100 (0.1) and the odds is 10:90 (1:9) = 10/90.
3. The odds (ratio) of using a Boolean search strategy after receiving formal training compared to informal training is the division of the above two odds:

$$\frac{\text{Odds of health professional trained using Boolean search strategy}}{\text{Odds of health professional trained informally using Boolean search strategy}} = \frac{A/B}{C/D} = \frac{A*D}{B*C}$$

Then:

$$\text{Odds Ratio (OR)} = \frac{25/75}{10/90} = \frac{25*90}{10*75} = 3$$

Thus, the odds of using Boolean search strategy is 3 times higher given formal training compared to no formal training.

In this research it was desirable to have a type of odds estimator, to represent the chance of retrieving sibling studies rather than non-sibling studies, for a particular database or search strategy, in comparison to the other databases and search strategies considered. Odds ratio is a common way to represent the results of statistical analysis for combining the results of several studies that are used within systematic reviews. Usually odds ratio deals with a definite number of participants in order to calculate the odds of an event happening. The intention within systematic reviews is slightly different as the question there is – overall, given the pooled results, some of which may show intervention A is better than intervention B, and others that show the reverse, what are the chances that A really is better than intervention B.

The question here is different, as the comparison is among the different strategies and databases. Recently, Stokes et al. (2009) used the odds ratio to compare database retrieval. This indicator proved to give a good performance indicator by giving a performance rank for each individual database based on the other participant (databases) retrieval performance. The advantage of the metric was that it included A, B, C, D, whereas recall and precision do not (see Section 3.6.1 and Table 3 in this section).

Stokes used the term odds ratio, but I have used the term odds estimator in this research to prevent the confusion between the odds ratio that is used in intervention effectiveness calculations (odds ratio for intervention outcomes that have a definite number of participants and therefore the D value can be easily configured). The odds estimator follow the same formula as odds ratio, however that value of D is different as explained below.

The odds estimator is more novel, but is used to compare the odds for two groups in the same way as the odds ratio, and it can be used to indicate the probability of a search strategy being able to retrieve relevant studies rather than non-relevant studies (Land & Altman, 2000; Ounger & Boddy, 2008). In this research, we were interested in comparing the search strategies' performance for five seed studies. Comparisons of the modified odds estimators can provide a way of taking recall and precision into account. From the practitioner perspective, the chances of finding siblings from a particular search strategy or database, rather than finding non-relevant material, is likely to be important. The idea behind using the odds estimator is that it gives an overall ranking score for the search strategy/database, which makes comparative performance easier to assess. The ranking is more important than the gross figures for the odds estimator (which makes use of this metric a little different from recall and precision).

Usually the number of non-relevant and non-retrieved items (D) would be calculated for each search strategy/database. However, the D value would be extremely high, and this would weight the calculated metric so much that the indicator would be useless. For example, a search on PubMed on publication date 1992-2010 yields 10891607 records – clearly, even if a thousand items were retrieved as relevant, the non-relevant number is going to swamp the conventional odds ratio calculation. For example the odds ratio for the related articles search on PubMed to retrieve sibling studies for the IDEATeL seed study will be as the following (See Table 3):

$$OR = (20*10891441)/(50*166)$$
$$OR = 26244.44$$

Another example is the odds ratio for the simple author-subject search on WoK to retrieve sibling studies for the IDEATeL seed study. The approximate number of records in WoK between 1992-2010 is 60606706 records and therefore the OR will be:

$$OR = (49* 60,606,693)/(21*13)$$
$$OR = 10878120.62$$

Obviously, both numbers are very large and are independent retrieval indicators that only provide the retrieval odds of the two databases as individual performances regardless of the other database retrieval performance values. Therefore the ratio is independent from the other databases' performance and does not reflect the ranking aspect that is required.

This, and the impossibility of coming to a reliable estimate of the non-relevant and non-retrieved for federated searching, made the use of another odds estimator preferable. The procedure followed was essentially identical to that used by Stokes et al. (2009).

Just as one can calculate the pooled relevant and retrieved items from all searches, the pooled non-relevant and not retrieved items by specific search strategy/database can be counted. For example, for the related article search for the IDEATeL seed study, the number of non-relevant and not retrieved is 1111 studies which is obtained by taking out the number of non-siblings/non-relevant records and retrieved by the related articles search from the pooled number of non-relevant records from all other search strategies. In other words, the non-siblings and not retrieved score (D) was the pooled non-siblings and retrieved by the remaining databases. Tables 3, 4 and 5 clearly demonstrate the calculation procedure.

Table 3: Two by Two Table Showing Modified Odds Estimator Calculations

	Relevant	Non-Relevant
Retrieved	A	B
Not Retrieved	C	D♠

♠ Pooled non-relevant and retrieved studies from other searches and databases.

- Odds Estimator = $\frac{(A*(D^{\spadesuit}-B))}{(B*C)}$
- Recall = $\frac{A}{(A+C)}$
- Precision = $\frac{A}{(A+B)}$

Example 1: Odds Estimator for PubMed related articles search strategy to retrieve direct sibling studies for the IDEATeL seed study.

Table 4: Siblings Retrieval from PubMed Related Articles Search of the IDEATeL Seed Study

	Relevant	Non-Relevant
Retrieved	A=20	B = 166
Not Retrieved	C= 50	D♠ = 1277

Odds Estimator (OE) = $\frac{(A*(D^{\spadesuit}-B))}{(B*C)}$

OE = $\frac{20*(1277-166)}{166*50}$

D♠ (in Table 4) is the pooled figure for all the databases/search strategies. D value that is used to calculate OE for each database is calculated by taking the non-relevant but retrieved figure from the pooled figure (D♠). For example, in Table 4, D♠ is 1277 and therefore D value will be 1111 (1277 after taking out the 166 non relevant retrieved items).

$$OE = \frac{20*1111}{166*50} = 2.68$$

Similarly, the total number of siblings is the pooled number of siblings that are retrieved from all search strategies and databases which is referred to as relative recall. Accordingly, recall value depends on the pooled number of siblings retrieved where each search strategy and database has an independent value of A and C (See Table 3). For example, the pooled number for the IDEATeL seed study is 70 siblings from all search strategies and databases. The PubMed related articles search strategy retrieved 20 siblings (Table 4) and accordingly the recall value is:

$$\text{Recall} = \frac{A}{(A+C)} = \frac{20}{70} = 0.29 = 29\%$$

While precision is:

$$\text{Precision} = \frac{A}{(A+B)} = \frac{20}{186} = 0.11 = 11\%$$

Furthermore, when a division by zero problem arose, the zero was substituted by 1 as a neutral value to overcome such mathematical problems (Breslow, 1981), (As illustrated in Table 5).

Example 2: Odds estimator for CINAHL author-subject search strategy to retrieve direct sibling studies for the IDEATeL seed study (division by zero problem) (Section 6.3.1).

Table 5: Siblings Retrieval from CINAHL Author-Subject Search of IDEATeL Seed Study

	Relevant	Non-Relevant
Retrieved	A=19	B = 0
Not Retrieved	C= 51	D = 1277

$$\text{Recall} = \frac{A}{(A+C)}$$

$$\text{Recall} = \frac{19}{(19+51)} = \frac{19}{70} = 0.27 = 27\%$$

$$\text{Precision} = \frac{A}{(A+B)}$$

$$\text{Precision} = \frac{19}{19} = 1 = 100\%$$

$$\text{OE} = \frac{19*(1277-0)}{(0*51)} = \text{division by zero}$$

Substitute B by 1, then

$$\text{OE} = \frac{19*1276}{1*51} = 475.37$$

Precision and recall are inversely related – high precision (A/A+B) is normally associated with low recall (A/A+C). If either B or C is large compared to A then the value of the ratio is low. High values of the ratio require B and C to be low in comparison to A.

In the end, the odds estimator appeared to provide a good indicator of retrieval performance for both search strategies and databases and that is the desired outcome for using the odds estimator measure. What is important in this research again is to provide a performance rank that reflects the retrieval performance of each database in relation to the other databases performances and not independent ranks based on each database only. Table 6 provides the ranks of search strategies and databases based on OR and OE measures, using available OR estimates as explained above.

As expected OR ranks the databases differently from OE as each one used a different value for D, however OE seems more realistic in this regard as it takes other databases into consideration when calculating the rank of each search strategy or database. In other words, X database has Y odds to retrieve sibling studies compared to the other databases.

Table 6: Search Strategies and Databases Ranks (OR and OE Ranks)

Search Strategy	OE	Search Strategy	OE	OR
WoK-Economics	0.92	PubMed-Economics	1.42	12925.07
WoK-Costs	1.19	PubMed –Costs	1.92	17288.15
PubMed-Economics	1.42	Sophisticated2 - Subject	1.53	20623.6
Sophisticated2 - Subject	1.53	Sophisticated	2.52	23231.09
PubMed -Costs	1.92	PubMed-Related articles	2.68	26244.44
Sophisticated	2.52	PubMed-Author-subject	3.88	36977.22
PubMed-Related articles	2.68	PubMed -Qualitative	4.25	38350.48
PubMed-Author-subject	3.88	WoK-Economics	0.92	45995.43
PubMed -Qualitative	4.25	WoK-Costs	1.19	61095.38
WoK-Qualitative	4.55	Sophisticated-Author-Subject-Qualitative	6.68	191080.6
Sophisticated-Author-Subject-Qualitative	6.68	Sophisticated2-Author-Subject-Costs	25.74	220032.4
Sophisticated2-Author-Subject-Costs	25.74	WoK-Qualitative	4.55	230224.6
Sophisticated2 – Author-Subject	29.03	Sophisticated2 – Author-Subject	29.03	254773.6
Sophisticated-Subject-Economics	98.15	Sophisticated-Subject- Economics	98.15	837815.8
WoK-Author-subject	263.54	WoK-Author-subject	263.54	16835194

4.6 Phase I: Exploring the Proposed Search Strategies Performance - Pilot Study

This phase was conducted as a pilot study to explore simple author-subject search, related articles search and citation search on different databases, for their potential in retrieving sibling studies for the five selected seed studies. The pilot phase results (appendix five) did not provide high recall values, but provided good precision and indicators of the good likelihood of retrieving relevant studies with a good proportion of siblings. Based on these results, it was decided to carry the research further and investigate search filters performance (phase two) in retrieving sibling studies. Moreover, based on federated (MetaLib) search performance it was decided to use different selection of databases in order to see how a different selection might change performance results of a MetaLib search. At the end of this phase a decision on performance measurements was made and the use of recall, precision, and the odds estimator (modified odds ratio) was confirmed.

4.7 Phase II: Exploring Combined Search Strategies/Search Filters and Databases Performance.

The second phase of this research was conducted based on the performance of search strategies and databases from the first phase. In this phase, some of search strategies were combined with specific search filters designed particularly to improve the retrieval of studies with specific study type or research design (Hedges, 2007a). Qualitative, economics and costs filters were chosen to explore their potential in retrieving sibling studies in combination with some search strategies and databases, based on the performance of the search strategies and databases explored in the first phase of this research. The following are the expanded search strategies/databases used in the second phase:

- Subject search on PubMed, WoK, SCOPUS and CINAHL combined with qualitative, economics and costs search filters.

- Author-Subject search on E-library (with combination of British library integrated catalogue, Centre of research libraries, Directory of open access journals, Scirus (Elsevier), Science direct (Elsevier), Index of theses, Theses & Dissertations catalogue, Intute and Intute: Medicine) combined with qualitative, economics and costs search filters.

The rationale for this is that this research aims to retrieve economic, qualitative and process evaluation siblings and these studies have specific research designs associated with them It was imperative to investigate the performance of existing specialised filters in retrieving sibling studies.

For the telemedicine and diabetes seed study two sophisticated search strings were formulated and used; the first one was advised by two researchers (myself and my supervisor) and the second was a revised search string for the first search string after advice from the PRESS forum. Together with my supervisor, I formed a sophisticated search string to search for the IDEATeL siblings. The search string was reviewed and validated by the PRESS forum and was used as well, with the initial one that was originally submitted to the forum as following:

- Sophisticated search on PubMed, using subject terms and MeSH terms. This search to be carried out based on retrieval results of the IDEATeL sophisticated search strategy performance.

The rationale for using a sophisticated search strategy was to examine how the sophisticated search strategy compared to simple search strategy in terms of performance (recall, precision) and the odds of each approach for retrieving sibling studies.

4.8 Phase III: Direct and Indirect Siblings Identification

After general topical relevance judgements were made and the relevant records were identified, stricter inclusion criteria were necessary in order to identify relevant studies likely to be direct siblings. Indirect siblings are the studies that are based on or emerge from the seed studies and aim to investigate other aspects that may interfere, affect or explain the intervention output using either the same or a different research design and meet at least one of the following inclusion criteria:

1. Any relevant study that appeared in the reference list of the seed study.
2. Relevant studies that came up with a citation search strategy.

Any relevant study that met at least one of these inclusion criteria was identified indirect siblings. The above inclusion criteria are used to screen the retrieved studies, and any study that meets one of these criteria is indirect, until it meets stricter inclusion criteria in the next and final filtering phase, at which point it will be promoted to the rank of a direct sibling, or remain an indirect sibling.

4.9 Phase III: Direct Sibling Identification

This phase was the last phase of this doctoral research, where in this phase only relevant studies with direct sibling rank were considered. Direct siblings in this context refer to the studies that are based on or emerge from the seed studies and aim to investigate other aspects that may interfere, affect or explain the intervention output using either the same or a different research design. Moreover, direct sibling studies must share at least one author with the seed study. The main inclusion criteria applied in this phase - authors' names of the seed study - was applied to direct and indirect siblings list identified in the previous phase. There may be other direct siblings but at least we can be surer about identification of direct siblings with some clear association on author names. This inclusion condition is based on the theme derived from the IDEATeL siblings list that was supplied to me by one of the IDEATeL seed study authors. All the studies in the list shared at least one author of the seed study.

4.10 Research Hypotheses and Hypotheses Testing

As aforementioned, this doctoral research aims to investigate the performance efficiency and effectiveness of different search strategies and databases in retrieving sibling studies. Accordingly, a set of hypotheses has been generated and tested. The set of hypotheses proposed the existence of relationships in order to derive inferences and themes about the nature and behaviour of findings. Two SPSS non-parametric test were performed with the direct sibling sets.

Set 1: Search strategies and databases performance efficiency relationship

H_0: There is no difference between databases in terms of precision.

H_1: There is a difference between databases in terms of precision.

In other words, this hypothesis is testing whether using database A rather than database B (and so on) will significantly affect the precision values for the sibling studies, across the five seed studies.

H_0: There is no difference between the recall from the databases used.

H_1: There is a difference between the recall from the databases used.

In other words, this hypothesis is testing whether the recall from database A, as compared to database B (and so on) will be significantly different, when examining all the recall figures obtained for the five seed studies.

H_0: There is no difference between search strategies' precision.

H_1: There is a difference between search strategies' precision.

In other words, this hypothesis is testing whether using one search strategy rather than another, will make a significant difference to the precision values obtained across the five seed studies.

H_0: There is no difference between search strategies recall.

H_1: There is a difference between search strategies recall.

In other words, this hypothesis is testing whether using one search strategy rather than another, will make a significant difference to the recall values obtained across the five seed studies.

Set 2: Search strategies, databases and siblings retrieval relationship

H_0: Database X is not more likely to retrieve sibling studies rather than non-siblings.

H_1: Database X is more likely to retrieve sibling studies rather than non-siblings.

In other words, this hypothesis is testing whether using one database rather than another database will be more productive in retrieving sibling studies rather than non-sibling studies.

H_0: Search strategy X is not more likely to retrieve sibling studies rather than non-siblings.

H_1: Search strategy X is more likely to retrieve sibling studies rather than non-siblings.

In other words, this hypothesis is testing whether using one search strategy rather than another search strategy will be more productive in retrieving sibling studies rather than non-sibling studies.

Set 3: Search strategy, databases and clinical area dependency relationship

H_0: Search strategy performance is independent of the clinical topic.

H_1: Search strategy performance is dependent on the clinical topic.

H_0: Database performance is independent of the clinical topic.

H_1: Database performance is dependent of the clinical topic.

Both of the hypothesis sets in Set 3 are testing whether the search strategy or database performance depend on the clinical topic – as a preliminary step in working out whether (if there is a relationship) what that relationship might be.

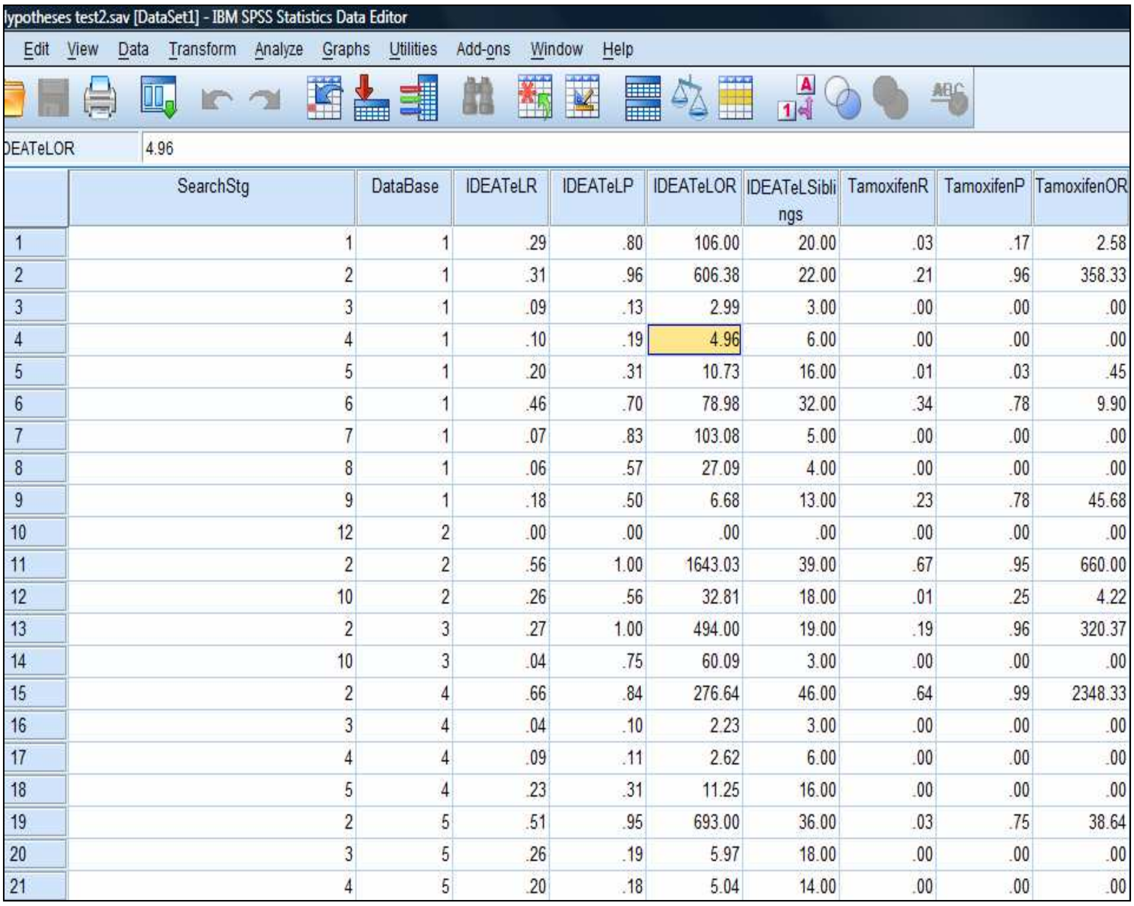
The hypotheses were analysed quantitatively using SPSS version 19. The analyses variables were search strategy, database, and clinical area as independent variables and recall, precision, odds estimator (performance measurements) and siblings studies retrieved number as dependent variables. The analyses targeted the existence of relationship between the independent and dependent variables, therefore the analyses were inferential rather than descriptive.

Both Friedman's and the Kruskal-Wallis tests were conducted to test the performance of search strategies and databases and to test the effect the clinical area might have over the outputs from using the search strategies and databases. These two tests were chosen because the data here are non-parametric (they do not follow a normal distribution - see Appendix Two to show the tests done to determine whether the data followed a normal distribution or not) (Green & Salkind, 2005; Field, 2009). Moreover, both of the tests are useful to test the significance of relationship when there are related although independent groups of data. Therefore, these tests are useful to test any significant differences in search strategy performance across different clinical areas (in this case the clinical areas are independent, although related through the retrieval measurements).

By handling the data and viewing it in two different ways, it was possible to benefit from two non-parametric tests. The Friedman test examines differences between several related groups, when there are more than two conditions and the same participants have been used in all conditions. According to those assumptions, for Friedman there are five different seeds, which are basically considered to be independent from each other but at the same instant are related to each other by the measurement type being used. For example, there are five seed studies related to each other by recall and again the same seed studies grouped together under precision and so on. So, in other words the measurements are the grouping factors for the Freidman test (Field, 2009).

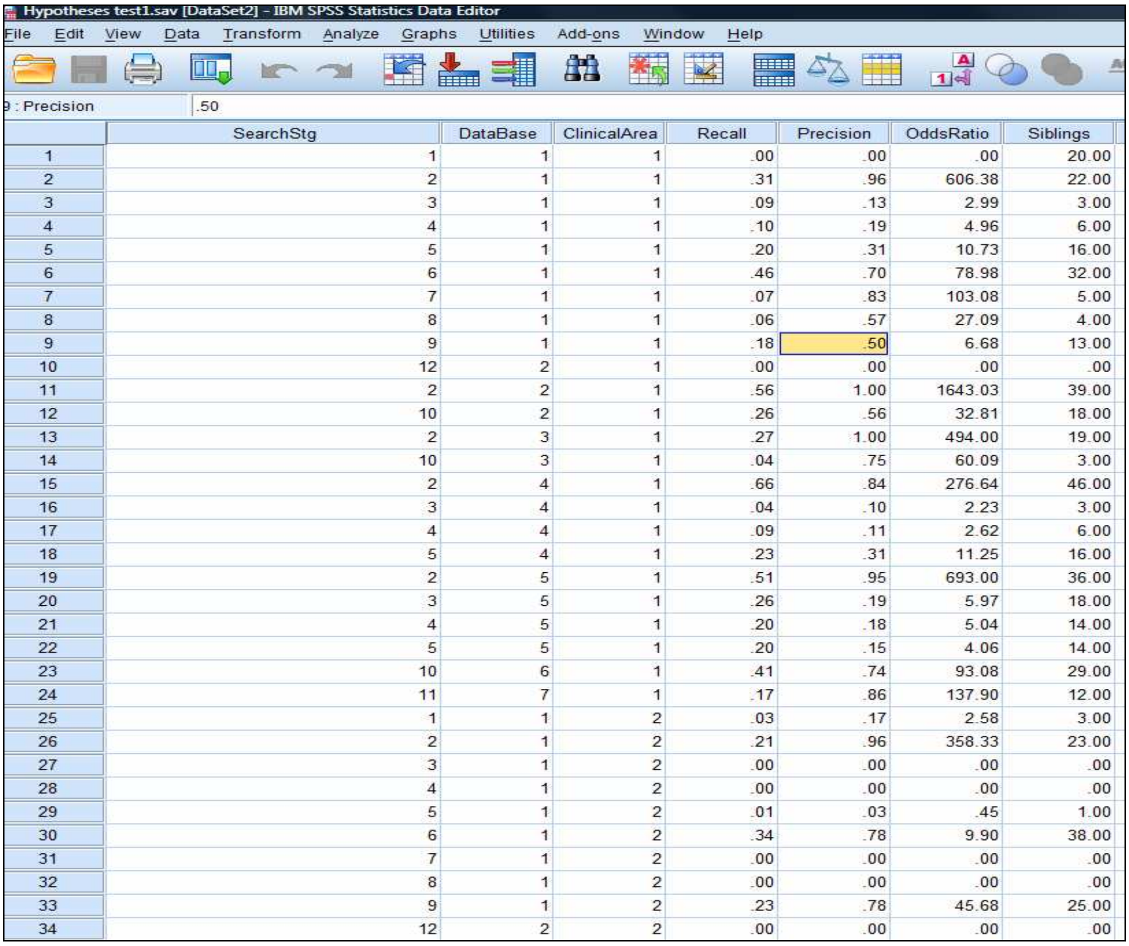
Figure 5 present a snapshot of the data view obtained from SPSS. Thus the last letter or letters stands for the measurement. R represents recall, P represents Precision, OE represents odds estimator and siblings represents the number of direct siblings retrieved for certain seed study according to its name, for example IDEATeLR and IDEATeLP stands for recall and precision values for the IDEATeL seed study obtained from different search strategies respectively. Finally, the Friedman test focuses on evaluating the differences between the median of different groups so it can tell us something about the general behaviour of the grouping factor used for each group under investigation. In this case the median refers to the median recall, precision, odds estimator or number of direct siblings retrieved for each group (the five seed studies).

The Kruskal-Wallis test is a non-parametric test to test for the difference between several independent groups (Field, 2009). In a way, the Kruskal-Wallis looks at the data "the other way round". I conducted the Kruskal-Wallis test on two different datasets – the same data with different representation- the first independent groups in the first dataset are the recall, precision, odds estimator and number of direct siblings retrieved. In this dataset I blinded seed study clinical area (by pooling all the values) and tested the four measurement variables against search strategies and databases. Again the test variables are independent from each other since each one represents different measurements as illustrated in Figure 6. In the second dataset I made use of the fact that the variables are independent from each other in terms of seed study clinical area and conducted the Kruskal-Wallis test. This test extended what was possible with the Friedman test in testing, for example, whether recall was significantly affected by choice of search strategy. The significance level of $p< 0.05$ was used (Field, 2009).



	SearchStg	DataBase	IDEATeLR	IDEATeLP	IDEATeLOR	IDEATeLSiblings	TamoxifenR	TamoxifenP	TamoxifenOR
1	1	1	.29	.80	106.00	20.00	.03	.17	2.58
2	2	1	.31	.96	606.38	22.00	.21	.96	358.33
3	3	1	.09	.13	2.99	3.00	.00	.00	.00
4	4	1	.10	.19	4.96	6.00	.00	.00	.00
5	5	1	.20	.31	10.73	16.00	.01	.03	.45
6	6	1	.46	.70	78.98	32.00	.34	.78	9.90
7	7	1	.07	.83	103.08	5.00	.00	.00	.00
8	8	1	.06	.57	27.09	4.00	.00	.00	.00
9	9	1	.18	.50	6.68	13.00	.23	.78	45.68
10	12	2	.00	.00	.00	.00	.00	.00	.00
11	2	2	.56	1.00	1643.03	39.00	.67	.95	660.00
12	10	2	.26	.56	32.81	18.00	.01	.25	4.22
13	2	3	.27	1.00	494.00	19.00	.19	.96	320.37
14	10	3	.04	.75	60.09	3.00	.00	.00	.00
15	2	4	.66	.84	276.64	46.00	.64	.99	2348.33
16	3	4	.04	.10	2.23	3.00	.00	.00	.00
17	4	4	.09	.11	2.62	6.00	.00	.00	.00
18	5	4	.23	.31	11.25	16.00	.00	.00	.00
19	2	5	.51	.95	693.00	36.00	.03	.75	38.64
20	3	5	.26	.19	5.97	18.00	.00	.00	.00
21	4	5	.20	.18	5.04	14.00	.00	.00	.00

Figure 5: SPSS File for Data Set 1



Hypotheses test1.sav [DataSet2] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

9 : Precision .50

	SearchStg	DataBase	ClinicalArea	Recall	Precision	OddsRatio	Siblings
1	1	1	1	.00	.00	.00	20.00
2	2	1	1	.31	.96	606.38	22.00
3	3	1	1	.09	.13	2.99	3.00
4	4	1	1	.10	.19	4.96	6.00
5	5	1	1	.20	.31	10.73	16.00
6	6	1	1	.46	.70	78.98	32.00
7	7	1	1	.07	.83	103.08	5.00
8	8	1	1	.06	.57	27.09	4.00
9	9	1	1	.18	.50	6.68	13.00
10	12	2	1	.00	.00	.00	.00
11	2	2	1	.56	1.00	1643.03	39.00
12	10	2	1	.26	.56	32.81	18.00
13	2	3	1	.27	1.00	494.00	19.00
14	10	3	1	.04	.75	60.09	3.00
15	2	4	1	.66	.84	276.64	46.00
16	3	4	1	.04	.10	2.23	3.00
17	4	4	1	.09	.11	2.62	6.00
18	5	4	1	.23	.31	11.25	16.00
19	2	5	1	.51	.95	693.00	36.00
20	3	5	1	.26	.19	5.97	18.00
21	4	5	1	.20	.18	5.04	14.00
22	5	5	1	.20	.15	4.06	14.00
23	10	6	1	.41	.74	93.08	29.00
24	11	7	1	.17	.86	137.90	12.00
25	1	1	2	.03	.17	2.58	3.00
26	2	1	2	.21	.96	358.33	23.00
27	3	1	2	.00	.00	.00	.00
28	4	1	2	.00	.00	.00	.00
29	5	1	2	.01	.03	.45	1.00
30	6	1	2	.34	.78	9.90	38.00
31	7	1	2	.00	.00	.00	.00
32	8	1	2	.00	.00	.00	.00
33	9	1	2	.23	.78	45.68	25.00
34	12	2	2	.00	.00	.00	.00

Figure 6: SPSS File for Data Set 2

4.11 Summary: Reflection on Methodology

The methodology of much information retrieval research has been quantitative, assessing whether search strategies or particular databases perform better in terms of recall and precision, or what the optimal trade-off between recall and precision is. One perennial difficulty is that these metrics of recall and precision depend on human judgements of relevance, and there is considerable debate about the difficulty of assessing relevance – it is inevitably subjective. Another difficulty is that these metrics are difficult or impossible to apply to Web searching, as the size of the test collection is almost infinite, and the same – though to a lesser extent, perhaps, applies to federated searching in digital libraries.

It was important to decide on limits that should be used with each search strategy in a way that it will not affect the retrieval performance in one hand and will not risk the lose of some sibling studies in the other. For example, with the federated search used in this research a decision about stopping was needed. Based on the retrieved ranked list it was noted that no relevant study after the first 250 to 300 record, so stopping point of first 300 was chosen. This emphasises the importance of the underlying ranking algorithm. In some instances when the retrieval size was low (under 100 records) author names were not used.

The research data are ordinal data that does not have normal distribution (Appendix Two). Accordingly, non-parametric tests (Kruskal-Wallis and Friedman) were used to analyse the association between research variables.

Moreover, as the link between an RCT and its sibling study is still unclear, it was imperative to decide on what might be considered a link that associate the siblings to the seed study. The IDEATeL seed study was the only seed study that had a complete list of siblings provided by one of its authors at the start of this research, and therefore it was used as a case study to deduce and decide on what the relation or link could be. In the end, the only clear option was to use authors as the base of determining the direct sibling relationship.

In conclusion, the multidisciplinary nature of this research directed the choice of search strategies as well as databases. It was very important to search databases other than PubMed as the social research element for some types of siblings (e.g. qualitative studies) cannot be neglected, moreover indexing variations will influence the retrieval performance and therefore, sibling retrieval can be optimised taking into account those two factors. As aforementioned retrieval performance, was assessed using recall, precision and an odds estimator (modified odds ratio).

Chapter Five
Relevant Retrieval: An Explanatory Phases for Siblings Retrieval

5.1 Introduction

This section introduces the results demonstrating the performance of each search strategy and database proposed in this research, grouped under each seed study individually. The findings set out the results for the overall performance of each search strategy/database in retrieving relevant studies, and also how each performs in order to provide a guide for sibling identification.

The results will be presented at the study level, indicating the performance of each search strategy and database based on that particular seed study, followed by performance assessment comparison for all the seeds. The reason for that is the need to explore whether there is a particular pattern for retrieving sibling studies regardless of the seed study's clinical topic, or whether the performance is strongly associated with the seed study's main clinical topic or type of intervention.

The results are presented chronologically, from four different search phases; the first three phases are presented in this chapter, at each phase specific inclusion criteria were used to narrow down the results first to relevant studies in order to identify the siblings in a later stage of this research and which will be detailed later in chapter six. The first phase (Section 5.2) outline the pilot study which aimed to explore possible search strategies and databases and decide if more or fewer search strategies might be more effective in retrieving sibling studies. The second phase results are presented in section 5.3. In section 5.4 phase three (direct and indirect siblings retrieval) results are presented.

The main performance metrics that are recall, precision and odds estimator for comparison (section 4.5).

5.2 Pilot Study

The first phase was carried out testing three main search strategies, Author-Subject search, Related articles search and Citation search, on six different databases (or sets of databases), PubMed, Web of Science (WoS), SCOPUS, CINAHL, Web of Knowledge (WoK) and a Metalib search using Aberystwyth university e-library interface on a combination of selected databases♠. This phase was initially planned as a pilot study starting with the IDEATeL seed study to investigate the possibility of finding direct siblings for it, and based on the performance obtained from the IDEATeL and search strategies applied on it the research was expanded to cover four different clinical topics to explore the possible differences and performance fluctuations that may rise.

These initial results showed that WoK, SCOPUS and CINAHL databases contributed to retrieving relevant on-topic studies more than other search databases (for the author-subject searches). Related article search strategy provided good retrieval for two seed studies – IVF and CLD- suggesting that this search strategy might be efficient in retrieving sibling studies. Subject search on e-library (federated search) provided a good retrieval results for some seed studies, suggesting the potential this type of search might have. Based on these results it was decided to carry this research into the next level. Further search strategies were added by incorporating search filters and advising another federated search using a different selection of databases (See Section 4.3). Table 7 present the top three search strategies in terms of recall, precision and odds estimator for the pilot study. Appendix Five provides the full tables for each seed study search strategies performance.

♠ This is the first selection of databases (see section 4.3)

Table 7: Three Search Strategies and Databases (Pilot Retrieval of Relevant on Topic)

IDEATeL		***Tamoxifen***		***BMS***		***CLD***		**IVF**	
				Recall					
WoK author-subject	45%	SCOPUS author-subject	39%	WoK author-subject	22%	PubMed related articles	50%	PubMed related articles	49 %
subject search(e-library)	38%	WoK author-subject	38%	subject search(e-library) PubMed related articles	17%	subject search(e-library)	36%	WoS citation	38 %
SCOPUS author-subject	32%	subject search(e-library)	31%	subject search(e-library)	9%	SCOPUS citation	16%	SCOPUS author-subject (Author(s))	32 %
				Precision					
CINAHL author-subject	94%	WoK author-subject	48%	SCOPUS Author-subject (Author(s)) CINAHL Author-subject CINAHL citation	100%	CINAHL author-subject WoK author-subject	100%	WoK author-subject	36 %
WoK author-subject	89%	CINAHL author-subject	46%	WoK author-subject	71%	SCOPUS author-subject (Author(s))	75%	WoS citation	25 %
SCOPUS author-subject	75%	SCOPUS author-subject	32%	WoS citation SCOPUS Citation	67%	subject search(e-library)	62%	SCOPUS author-subject (Author(s))	24 %
				Odds ratio					
CINAHL author-subject		WoK author-subject		WoK author-subject		CINAHL author-subject		WoK author-subject	
WoK author-subject		CINAHL author-subject		WoS citation SCOPUS citation CINAHL author-subject CINAHL citation		WoK author-subject		WoS citation	
SCOPUS author-subject		SCOPUS author-subject		subject search(e-library)		SCOPUS author-subject (Author(s))		SCOPUS author-subject (Author(s))	

5.3 Second Stage Search Strategies

Based on the results from the pilot study, the research sought out to incorporate specific search filters with some selected search strategies and databases, according to their performance from the previous run of the search(s) (See Sections 4.4 and 5.2). Incorporating search filters with search strategies yielded additional studies which were then added to search results from the previous run of the search, creating a pooled set of relevant studies from all refined searches/databases, changing the odds estimators, and other indicators accordingly. The results here represent a cumulative retrieval results from both runs of search strategies.

5.3.1 Diabetes Telemedicine (IDEATeL) Seed Study

The search for the telemedicine and diabetes siblings from all search strategies yielded 1347 studies in total, and among these 350 studies were categorised as relevant with a strong possibility of being among the direct siblings for the seed study.

5.3.1.1 Recall

According to Table 8, E-Library author-subject with Hedges qualitative and economics filters retrieved the highest number of relevant studies with a value of 96 and 95 respectively, scoring a recall value of 27%. The author-subject search on E-library along with the Hedges costs search filter, scored the second highest recall value of 23%, retrieving 80 relevant studies. The third best performance resulted from author-subject search on author-subject search on WoK scoring the third highest number of relevant studies being retrieved (55 relevant studies) with a recall of value 16%.

5.3.1.2 Precision

In terms of precision, Hedges search filters appeared to provide the highest precision. For example, the CINAHL Author-subject, CINAHL subject search with Hedges qualitative, economics and costs filters and SCOPUS with the Hedges qualitative filter were the best search strategies scoring a precision value of 100%, followed by WoK author-subject search with a precision value of 89%. Finally, the search strategies that performed the third best were SCOPUS Author-subject (Author) and SCOPUS subject search with the Hedges costs filter with a precision value of 75% (see Table 8).

5.3.1.3 Odds Estimator

Furthermore, the odds estimator calculations showed that CINAHL author-subject search scored the highest odds estimator with score of 57.17, indicating the best chances of retrieving relevant studies rather than non-relevant studies, SCOPUS subject search strategy with the Hedges qualitative filter odds estimator was the second with score of 41.5 indicating the second best likelihood of obtaining the relevant studies rather than non-relevant studies and CINAHL subject search with the Hedges qualitative and economics filters was the third most likely to retrieve relevant studies rather than non-relevant studies with odds estimator of 17.37. The search with the smallest odds of retrieving the relevant, rather than non-relevant studies was the subject search on e-library with odds estimator of 0.36, as illustrated in Table 8.

Table 8: The IDEATeL Search and Databases Retrieval

Search Strategy	Relevant/R (A)	Relevant N/R©	Not relevant (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
Diabetes - Telemedicine								
PubMed-Related articles	25	369	161	0.07	7%	0.13	13%	0.40
PubMed-Author-subject	23	371	134	0.07	7%	0.15	15%	0.45
Citation(Web of Science)	14	380	43	0.04	4%	0.25	25%	0.92
Subject search (e-library)	39	355	257	0.11	11%	0.14	14%	0.36
SCOPUS Author-subject (Author(s))	39	355	13	0.11	11%	0.75	75%	9.49
SCOPUS citation	32	362	32	0.09	9%	0.50	50%	3.03
CINAHL Author-subject	19	375	0	0.05	5%	1.00	100%	57.17
CINAHL citation	4	390	3	0.01	1%	0.57	57%	3.83
WoK-Author-subject	55	339	7	0.16	16%	0.89	89%	0.89
E-lib – Author-subject	38	356	17	0.11	11%	0.69	69%	0.69
PubMed-Economics	48	346	37	0.14	14%	0.56	56%	4.12
PubMed –Costs	37	357	40	0.11	11%	0.48	48%	2.83
PubMed -Qualitative	45	349	36	0.13	13%	0.56	56%	3.94
E-lib-Economics	95	299	134	0.27	27%	0.41	41%	1.58
E-lib-Costs	80	314	134	0.23	23%	0.37	37%	1.91
E-lib-Qualitative	96	298	172	0.27	27%	0.36	36%	1.81
WoK-Economics	30	364	32	0.09	9%	0.48	48%	2.83
WoK-Costs	54	340	45	0.15	15%	0.55	55%	3.86
WoK-Qualitative	51	343	43	0.15	15%	0.54	54%	3.78
SCOPUS- Qualitative	14	380	1	0.04	4%	1.00	100%	41.50
SCOPUS-Costs	3	391	1	0.01	1%	0.75	75%	8.61
SCOPUS-Economics	2	392	2	0.01	1%	0.50	50%	2.86
CINAHL- Qualitative	6	388	1	0.02	2%	1.00	100%	17.37
CINAHL-Costs	4	390	1	0.01	1%	1.00	100%	11.51
CINAHL-Economics	6	388	1	0.02	2%	1.00	100%	17.37

Total relevant R without duplicates	**350**
Total retrieved without duplicate	**1347**
Total non-relevant R without duplicates (**D**)	**997**
Odds Estimator = A * (D – B) / B * C	

5.3.1.4 Sibling Studies Retrieval by Study Type

Each retrieved item was categorised into one of four sibling types (RCT, Qualitative, Economics or Process evaluation) in order to investigate search strategies and databases performance against specific study type. According to Table 9, the CINAHL author-subject search odds estimator indicate the highest likelihood of retrieving RCT studies rather than non-RCTs, with odds estimator values of 101.63. WoK author-subject performed the second best in retrieving RCTs rather than non-RCTs with odds estimator of 32.14, while subject search with Hedges qualitative filter on SCOPUS scored the third best possibility of retrieving RCT studies rather than non-RCTs with odds estimator value of 18.79. That seems to be odd for a qualitative search filter to retrieve RCTs, however this might imply that there was a randomised aspect of the study design.

According to Table 9, the subject search with Hedges qualitative filter on SCOPUS suggested the best possibility of retrieving qualitative studies rather than non-qualitative with odds estimator value of 84.09. The searches that performed second and third best in retrieving the qualitative siblings were first; CINAHL author-subject search and CINAHL subject search with Hedges qualitative filter with odds estimator of 40.65 followed by the WoK author-subject search with odds estimator of 38.9 suggesting the second and third best possibility of retrieving qualitative rather than non-qualitative studies. SCOPUS subject search with both Costs and Economics Hedges search filters and CINAHL subject search with Costs Hedges search filter did not retrieve any RCTs or qualitative studies indicating the least possibility of retrieving either RCTs or qualitative studies rather than non-RCTs or non-qualitative.

Furthermore, economics studies were best retrieved by CINAHL subject search with both the Hedges economics and costs filters scoring the highest odds estimator with value of 71.14 indicating the best possibility of retrieving the economics rather than non-economics, followed by subject search on SCOPUS with the Hedges economics filter as the second best search expected to retrieve economics studies rather than non-economics with odds estimator of 23.14. The CINAHL author-subject with odds estimator value of 22.64 with the third best possibility of retrieving economics studies rather than non-economic studies. CINAHL citation, subject search with Hedges qualitative filter on PubMed, WoK, SCOPUS and CINAHL were the searches with the least possibility of retrieving economics studies with odds estimator value of zero, as illustrated in Table 9.

And finally, for the process evaluation siblings, the search with highest odds of retrieving process evaluation studies rather than non-process evaluation was the CINAHL author-subject search with odds estimator value of 29.51 while the second best odds of retrieving process evaluation rather than non-process evaluation studies was scored by subject search with Hedges qualitative filter on SCOPUS with odds estimator value of 29.39 and the third best likelihood of retrieving process evaluation rather than non-process evaluation studies was scored by author-subject search on WoK with odds estimator value of 10.96. Citation search did not retrieve any process evaluation siblings yielding odds estimator value of zero, (See Table 9).

Table 9: IDEATeL Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	9	99	0.47	9	93	0.5	3	42	0.37	4	135	0.15
PubMed-Author-subject	8	100	0.52	9	93	0.62	2	43	0.3	4	135	0.19
Citation(Web of Science)	4	104	0.85	8	94	1.89	2	43	1.03	0	139	0
Subject search (e-library)	13	95	0.39	11	91	0.35	3	42	0.21	5	134	0.11
SCOPUS Author-subject (Author(s))	17	91	14.14	12	90	10.09	2	43	3.52	8	131	4.62
SCOPUS citation	19	89	6.44	8	94	2.57	2	43	1.4	3	136	0.66
CINAHL Author-subject	10	98	101.6	4	98	40.65	1	44	22.64	4	137	29.51
CINAHL citation	1	107	3.1	3	99	10.04	0	45	0	0	139	0
WoK-Author-subject	20	88	32.14	22	80	38.89	3	42	10.1	10	129	10.96
E-lib – Author-subject	18	90	4.19	13	89	3.06	1	44	0.48	6	133	0.95
PubMed-Economics	13	95	3.55	5	97	1.34	16	29	14.32	14	125	2.9
PubMed -Costs	11	97	2.71	2	100	0.48	15	30	11.96	9	130	1.66
PubMed -Qualitative	7	101	1.85	20	82	6.51	0	45	0	18	121	3.97
E-lib-Economics	23	85	1.74	32	70	2.94	6	39	0.99	34	105	2.09
E-lib-Costs	19	89	1.93	16	86	1.68	25	20	11.32	20	119	1.52
E-lib-Qualitative	23	85	1.3	23	79	1.4	21	24	4.2	29	110	1.26
WoK-Economics	12	96	3.77	0	102	0	9	36	7.54	9	130	2.09
WoK-Costs	19	89	4.52	4	98	0.86	13	32	8.59	18	121	3.15
WoK-Qualitative	7	101	1.54	24	78	6.83	0	45	0	20	119	3.73
SCOPUS- Qualitative	2	106	18.79	8	94	84.09	0	45	0	4	135	29.39
SCOPUS-Costs	0	108	0	0	102	0	2	43	18.19	1	138	2.83
SCOPUS-Economics	0	108	0	0	102	0	2	43	23.14	0	139	0
CINAHL- Qualitative	1	107	9.31	4	98	40.65	0	45	0	1	138	7.22
CINAHL-Costs	0	108	0	0	102	0	3	42	71.14	1	138	7.22
CINAHL-Economics	1	107	9.31	1	101	9.86	3	42	71.14	1	138	7.22

5.3.2 Breast Cancer (Tamoxifen) Seed Study

The search for the breast cancer prevention using Tamoxifen siblings from all search strategies yielded 1844 studies, where 387 studies were categorised as relevant with a strong possibility of being one of the direct siblings for the seed study (See Table 10).

5.3.2.1 Recall

According to Table 10, SCOPUS author-subject search retrieved 78 relevant studies scoring the highest recall value among all search strategies and databases with value of 20%. WoK author-subject scored the second best performance retrieving 72 of studies that has been categorised as relative relevant with a recall value of 19%. The third search that followed was the e-library subject search retrieving 59 relevant studies scoring a recall value of 15%.

5.3.2.2 Precision

Table 10 shows that in terms of precision, the CINAHL author-subject search performed the best among other search strategies and databases scoring precision value of 54%, followed by WoK author-subject search as the second best performing search scoring a precision value of 49%. And finally the third best performing search was SCOPUS subject with the Hedges economics filter with precision value of 45%.

5.3.2.3 Odds Estimator

Odds estimator calculations show that the search with the highest likelihood of retrieving relevant studies rather than non-relevant was CINAHL author-subject search with odds estimator value of 4.56. WoK author-subject was the second best search of retrieving related studies scoring odds estimator with value of 4.21 indicating the second best chances of retrieving relevant studies rather than non-relevant studies. SCOPUS subject with the Hedges economics filter odds estimator with value of 3.18 indicating the third best chances of retrieving relevant studies rather than non-relevant studies. The search with the least possibility of retrieving relevant studies was author-subject search on PubMed with odds estimator value of 0.16, as illustrated in Table 10.

Table 10: **Breast Cancer (Tamoxifen): Search and Databases Retrieval Performance**

Search Strategy	Relevant/R (A)	Relevant N/R©	Not relevant (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
				Breast Cancer (Tamoxifen)				
PubMed-Related articles	18	365	182	0.05	5%	0.09	9%	0.34
PubMed-Author-subject	24	359	427	0.06	6%	0.05	5%	0.16
Citation(Web of Science)	7	376	46	0.02	2%	0.13	13%	0.57
Subject search (e-library)	59	324	229	0.15	15%	0.20	20%	0.96
SCOPUS Author-subject (Author(s))	78	305	151	0.20	20%	0.34	34%	2.18
SCOPUS citation	4	379	55	0.01	1%	0.07	7%	0.27
CINAHL Author-subject	22	361	19	0.06	6%	0.54	54%	4.56
CINAHL citation	1	382	3	0.003	0.3%	0.25	25%	1.26
WoK-Author-subject	72	311	75	0.19	19%	0.49	49%	4.21
E-lib – Author-subject	4	379	39	0.01	1%	0.09	9%	0.38
PubMed-Economics	21	362	52	0.05	5%	0.29	29%	1.55
PubMed -Costs	16	367	50	0.04	4%	0.24	24%	1.21
PubMed -Qualitative	29	354	104	0.07	7%	0.22	22%	1.05
E-lib-Economics	52	331	161	0.13	13%	0.24	24%	0.99
E-lib-Costs	38	345	120	0.10	10%	0.24	24%	1.21
E-lib-Qualitative	50	333	165	0.13	13%	0.23	23%	1.16
WoK-Economics	43	340	96	0.11	11%	0.31	31%	1.77
WoK-Costs	31	352	68	0.08	8%	0.31	31%	1.78
WoK-Qualitative	55	328	151	0.14	14%	0.27	27%	1.43
SCOPUS- Qualitative	46	341	88	0.12	12%	0.34	34%	2.10
SCOPUS-Costs	26	361	63	0.07	7%	0.29	29%	1.59
SCOPUS-Economics	18	369	22	0.05	5%	0.45	45%	3.18
CINAHL- Qualitative	7	380	15	0.02	2%	0.32	31%	1.77
CINAHL-Costs	3	384	8	0.01	1%	0.27	27%	1.42
CINAHL-Economics	7	380	23	0.02	2%	0.23	23%	1.15

Total relevant R without duplicates	**387**
Total retrieved without duplicate	**1844**
Total non-relevant R without duplicates (**D**)	**1457**
Odds Estimator = A * (D – B) / B * C	

5.3.2.4 Sibling Studies Retrieval by Study Type

According to Table 11, the odds estimator shows that there was neither search strategy nor database that was superior in performance to any other search or database in retrieving all four types of siblings. Accordingly, the RCT studies were more likely to be retrieved rather than non-RCTs by WoK author-subject as indicated by its odds estimator with value of 9.79, followed by CINAHL author-subject with odds estimator value of 7.63. SCOPUS author-subject scored the third best odds estimator for retrieving RCTs studies rather than non-RCTs with value of 4.023. Finally, citation search on WoS, CINAHL citation, subject search with Hedges economics filter on PubMed and SCOPUS, E-Library author-subject and CINAHL subject search with Hedges costs filter did not retrieve any RCT resulting in odds estimator value of zero.

For the qualitative studies, CINAHL citation search scored the highest odds estimator value of 4.81 indicating the best possibilities of retrieving the qualitative studies rather than non-qualitative studies. The second best chances of retrieving the qualitative studies rather than non-qualitative studies was achieved by subject Hedges qualitative filter on SCOPUS as indicated by odds estimator value of 3.35, and the third best likelihood of retrieving the qualitative studies rather than non-qualitative studies was scored by CINAHL author-subject with odds estimator value of 3.08. The search with the least likelihood of retrieving the qualitative studies rather than non-qualitative studies was CINAHL subject search with Hedges costs filter with odds estimator value of zero, (Table 11).

Economic evaluation siblings were best retrieved by subject with Hedges economics filter on SCOPUS with odds estimator value of 39.86 indicating the highest likelihood of retrieving economics studies rather than non-economics studies. WoK subject search with costs Hedges filter was next (odds estimator value of 25.14), while author-subject with Hedges costs filter on PubMed with odds estimator value of 24.84 was third. However, several search strategies and databases did not retrieve any economics siblings at all, i.e. Author-Subject on PubMed, Citation search on Web of Science, e-library subject search, SCOPUS author-subject, SCOPUS citation, CINAHL author-subject, CINAHL citation, WoK author-subject and finally E-library - author-subject as illustrated in Table 11.

Finally, odds estimator calculations indicate that CINAHL subject search with Hedges qualitative filter yielded the highest odds estimator with value of 4.58 pointing out the highest possibility of retrieving process evaluation studies rather than non-process evaluation ones. The CINAHL author-subject search scored the second best (odds estimator value of 3.6.) and the third highest likelihood of retrieving process evaluation siblings was achieved by WoK subject search with Hedges economics filter (odds estimator value of 2.91) (Table 11).

According to the Table 11, SCOPUS citation search, CINAHL citation author-subject on E-library and CINAHL subject search with Hedges costs filter did not retrieve any process evaluation studies with odds estimator value of zero.

Table 11: Tamoxifen Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	8	134	0.42	8	120	0.47	1	28	0.25	1	87	0.08
PubMed-Author-subject	10	132	0.75	10	118	0.2	0	29	0	4	84	0.12
Citation(Web of Science)	0	142	0	5	123	1.25	0	29	0	2	86	0.71
Subject search (e-library)	27	115	1.26	18	106	0.88	0	29	0	14	74	1.02
SCOPUS Author-subject (Author(s))	47	95	4.02	19	109	1.42	0	29	0	12	76	1.28
SCOPUS citation	1	141	0.18	3	125	0.61	0	29	0	0	88	0
CINAHL Author-subject	13	129	7.63	5	123	3.08	0	29	0	4	84	3.6
CINAHL citation	0	142	0	1	127	3.81	0	29	0	0	88	0
WoK-Author-subject	51	91	9.79	12	116	1.81	0	29	0	9	79	1.99
E-lib – Author-subject	0	142	0	4	124	1.17	0	29	0	0	88	0
PubMed-Economics	0	142	0	2	126	0.42	14	15	24.84	5	83	1.6
PubMed -Costs	1	141	0.2	1	127	0.22	12	17	19.64	2	86	0.65
PubMed -Qualitative	4	138	0.37	18	110	2.08	4	25	2.04	3	85	0.45
E-lib-Economics	15	127	0.91	8	120	0.52	10	19	4.07	19	69	2.13
E-lib-Costs	3	139	0.24	6	122	0.55	13	16	9.05	16	72	2.48
E-lib-Qualitative	3	139	1.21	6	118	1.37	3	26	0.9	9	79	0.89
WoK-Economics	6	136	0.63	7	121	0.82	15	14	15.19	15	73	2.91
WoK-Costs	3	139	0.44	2	126	0.32	16	13	25.14	10	78	2.62
WoK-Qualitative	18	124	0.92	23	105	1.39	1	28	0.23	13	75	1.1
SCOPUS- Qualitative	10	132	1.18	23	105	3.35	3	26	1.79	10	78	1.98
SCOPUS-Costs	1	141	0.04	9	119	0.43	14	15	5.35	2	86	0.13
SCOPUS-Economics	0	142	0	5	123	2.65	11	18	39.86	2	86	1.52
CINAHL- Qualitative	1	141	0.68	2	126	1.53	0	29	0	4	84	4.58
CINAHL-Costs	0	142	0	0	128	0	3	26	20.9	0	88	0
CINAHL-Economics	1	141	0.44	1	127	0.49	3	26	7.19	2	86	1.45

5.3.3 Breast Cancer (Body Mind Spirit) Seed Study

Search strategies and databases searching for studies related to breast cancer study using body mind spirit therapy yielded 614 studies, of these 71 studies were categorised to be relevant with a strong possibility of being a sibling.

5.3.3.1 Recall

As shown in Table 12, author-subject search on WoK retrieved 15 relevant studies (highest recall value of 21%). Related search on PubMed and subject on e-library each yielded the same number of 12 relevant studies (second highest recall of 17%). WoK subject search with Hedges qualitative filter retrieved, E-library author-subject and subject search with qualitative Hedges filter on both SCOPUS and CINAHL search retrieved 6 relevant studies (third highest recall with value of 8%).

5.3.3.2 Precision

Moreover, comparing precision values demonstrate that SCOPUS author-subject (Author(s)), SCOPUS subject search with Hedges qualitative filter, CINAHL author-subject and CINAHL citation search scored the highest precision value, with value of 100%, as they all retrieved a small number of studies where all retrieved studies were relevant with no non relevant studies. Subject search on WoK with Hedges qualitative filter scored the second best precision with value of 75%. The third best searches were WoK author-subject search and subject search on PubMed with Hedges qualitative filter where they scored a precision with value of 71%, as illustrated in Table 12.

5.3.3.3 Odds Estimator

Finally, SCOPUS subject search with Hedges qualitative filter is more likely to retrieve relevant studies rather than non-relevant studies (highest odds estimator with score of 50), WoK subject search with Hedges qualitative filter performed the second best (odds estimator value of 24.97) and author-subject on WoK search strategy performed the third best (odds estimator value of 23.97). Finally, Author-subject search on PubMed is considered to be the search with the least chances of retrieving relevant studies with odds estimator value of 0.38, It should be mentioned that not all search strategies retrieved studies related to this specific seed study (see Table 12).

Table 12: The BMS Search and Databases Retrieval

Search Strategy	Relevant/R (A)	Relevant N/R©	Not relevant (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
Breast Cancer (BMS)								
PubMed-Related articles	12	59	96	0.17	17%	0.11	11%	0.95
PubMed-Author-subject	3	68	64	0.04	4%	0.04	4%	0.38
Citation(Web of Science)	2	69	1	0.03	3%	0.67	67%	15.71
Subject search(e-library)	12	59	11	0.17	17%	0.52	52%	9.84
SCOPUS Author-subject (Group)	5	66	21	0.07	7%	0.19	19%	1.88
SCOPUS Author-subject (Author(s))	1	70	0	0.01	1%	1.00	100%	7.74
SCOPUS citation	2	69	1	0.03	3%	0.67	67%	15.71
CINAHL Author-subject	2	69	0	0.03	3%	1.00	100%	15.71
CINAHL citation	2	69	0	0.03	3%	0.67	100%	15.71
WoK-Author-subject	15	56	6	0.21	21%	0.71	71%	23.97
E-lib – Author-subject	6	65	77	0.08	8%	0.07	7%	0.56
PubMed-Economics	--	--	--	0.00	--	0.00	--	0.00
PubMed -Costs	--	--	--	0.00	--	0.00	--	0.00
PubMed -Qualitative	5	66	2	0.07	7%	0.71	71%	20.49
E-lib-Economics	--	--	--	0.00	--	0.00	--	0.00
E-lib-Costs	--	--	--	0.00	--	0.00	--	0.00
E-lib-Qualitative	--	--	--	0.00	--	0.00	--	0.00
WoK-Economics	--	--	--	0.00	--	0.00	--	0.00
WoK-Costs	--	--	--	0.00	--	0.00	--	0.00
WoK-Qualitative	6	65	2	0.08	8%	0.75	75%	24.97
SCOPUS- Qualitative	6	65	0	0.08	8%	1.00	100%	50.03
SCOPUS-Costs	0	71	0	0.00	0%	0.00	0%	0.00
SCOPUS-Economics	0	71	0	0.00	0%	0.00	0%	0.00
CINAHL- Qualitative	6	65	8	0.08	8%	0.43	43%	6.17
CINAHL-Costs	0	71	3	0.00	0%	0.00	0%	0.00
CINAHL-Economics	5	66	7	0.07	7%	0.42	42%	5.80

Total relevant R without duplicates	**71**
Total retrieved without duplicate	**614**
Total non-relevant R without duplicates (**D**)	**543**
Odds Estimator = A * (D – B) / B * C	

5.3.3.4 Sibling Studies Retrieval by Study Type

Further analysis was done by categorising each retrieved study into one of the four sibling types (RCT, qualitative, process evaluation and economical evaluation). As Table 13 demonstrates, there is no search strategy among all these which performed the best for all four siblings category; consequently, CINAHL author-subject performed the best for RCTs with odds estimator of 67.5 followed by SCOPUS author-subject (Author(s)) search in retrieving the RCTs with odds estimator of 31.82, and SCOPUS subject search with

Hedges qualitative filter was third with odds estimator value of 31.58. WoS citation and SCOPUS citation among other search strategies and databases can be considered to be the strategies with the least possibility of retrieving RCTs siblings as indicated by their odds estimator with value of zero (Table 13).

For the qualitative siblings, SCOPUS subject search with Hedges qualitative filter scored highest with odds estimator value of 66.5, followed by WoK subject search with Hedges qualitative filter, and subject search PubMed with Hedges qualitative filter with odds estimator value of 41.15 for both, and WoK author-subject scored the next best (with odds estimator value of 22). WoS citation, SCOPUS author-subject, SCOPUS citation, CINAHL author-subject, CINAHL citation searches as indicated by their odds estimator, subject search on SCOPUS with both costs and economics Hedges filter and CINAHL subject search with Hedges costs filter would be the least likely searches to retrieve qualitative studies with odds estimator of zero (Table 13).

There was no economics sibling retrieved by any of search strategies on different databases. And finally for process evaluation siblings, subject search on e-library scored the highest odds estimator with a value of 334.1 indicating the highest likelihood of retrieving process evaluation studies rather than non-process evaluation ones. The second was jointly scored by both SCOPUS citation and CINAHL citation search with odds estimator of 180. Citation search on WoS was the search with the third with odds estimator value of 77.29. There was several search strategies and databases that did not retrieve any process evaluation studies i.e. SCOPUS Author-subject (Author(s)) and CINAHL Author-subject scoring odds estimator of zero and indicating the least possibility of retrieving process evolution studies rather than non-process evaluation, as illustrated in Table 13.

Table 13: BMS Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	5	13	1.79	6	39	0.71	0	0	0	1	7	0.66
PubMed-Author-subject	1	17	0.44	0	45	0	0	0	0	2	6	2.48
Citation(Web of Science)	0	18	0	1	44	12.3	0	0	0	1	7	77.27
Subject search(e-library)	7	11	30.37	3	42	3.44	0	0	0	7	1	334.09
SCOPUS Author-subject (Group)	2	16	3.1	1	44	0.56	0	0	0	2	6	8.25
SCOPUS Author-subject (Author(s))	1	17	31.82	0	45	0	0	0	0	0	8	0
SCOPUS citation	0	18	0	0	45	0	0	0	0	2	6	180
CINAHL Author-subject	2	16	67.5	0	45	0	0	0	0	0	8	0
CINAHL citation	0	18	0	0	45	0	0	0	0	2	6	180
WoK-Author-subject	4	14	25.38	9	36	22	0	0	0	2	6	29.72
E-lib – Author-subject	5	13	2.3	1	44	0.14	0	0	0	0	8	0
PubMed-Economics	--	--	--	--	--	--	--	--	--	--	--	--
PubMed -Costs	--	--	--	--	--	--	--	--	--	--	--	--
PubMed -Qualitative	0	18	0	5	40	41.15	0	0	0	0	8	0
E-lib-Economics	--	--	--	--	--	--	--	--	--	--	--	--
E-lib-Costs	--	--	--	--	--	--	--	--	--	--	--	--
E-lib-Qualitative	--	--	--	--	--	--	--	--	--	--	--	--
WoK-Economics	--	--	--	--	--	--	--	--	--	--	--	--
WoK-Costs	--	--	--	--	--	--	--	--	--	--	--	--
WoK-Qualitative	0	18	0	6	39	41.15	0	0	0	0	8	0
SCOPUS- Qualitative	1	17	31.59	5	40	66.5	0	0	0	0	8	0
SCOPUS-Costs	--	--	--	--	--	--	--	--	--	--	--	--
SCOPUS-Economics	--	--	--	--	--	--	--	--	--	--	--	--
CINAHL- Qualitative	2	16	8.23	4	41	6.4	0	0	0	0	8	0
CINAHL-Costs	0	18	0	0	45	0	0	0	0	0	8	0
CINAHL-Economics	0	18	0	5	40	9.39	0	0	0	0	8	0

5.3.4 In Vitro Fertilisation (hMP vs rFSH) Seed Study

When the search based on in vitro fertilization was carried out, the search strategies and databases yielded 1650 studies, 72 studies were categorised as relevant with a strong possibility of being one of the direct siblings as shown in Table 14.

5.3.4.1 Recall

As shown in Table 14, author-subject search on E-library with Hedges economics filter retrieved 21 related studies scoring the highest recall value of 29%. Related search strategy on PubMed retrieved the second highest number of relevant studies (18 relevant studies, recall score of 25%), and the third search strategy was WoS citation search (14 relevant studies, recall value of 19%).

5.3.4.2 Precision

In addition, precision calculations demonstrate that subject search on SCOPUS with Hedges economics filter scored the highest precision with score of 86%. WoK subject search with Hedges economics, and WoK subject search with Hedges costs filters were both the second best search (precision value of 63%). The third best performing searches were subject search on PubMed with Hedges economics, subject search on PubMed search with Hedges costs filter, and subject search on SCOPUS with Hedges costs filter (precision value of 56%) as indicated in Table 14.

5.3.4.3 Odds Estimator

The odds estimator calculations showed that subject search on SCOPUS with Hedges economics filter is most likely to retrieve relevant studies rather than non-relevant ones with the highest odds estimator value of 143.36, followed WoK subject search with Hedges economics, and WoK subject search with Hedges costs filters, as the second best (both with odds estimator score of 39.18). Next were subject search on PubMed with Hedges economics, subject search on PubMed search with Hedges costs filters, and subject search on SCOPUS with Hedges costs filter (odds estimator value of 29.37). Author-subject on PubMed was the search strategy with the least chances of retrieving sibling studies with odds estimator value of 0.17, while there was some search strategies which did not retrieve any relevant studies for the IVF seed studies indicating zero possibility of retrieving any relevant studies rather than non-relevant studies, as illustrated in Table 14.

Table 14: The IVF Search Strategies and Databases Retrieval

IVF								
Search Strategy	**Relevant/R (A)**	**Relevant N/R©**	**Not relevant (B)**	**Recall**		**Precision**		**Odds Estimator**
				Value	**%**	**Value**	**%**	
PubMed-Related articles	18	54	88	0.25	25%	0.17	17%	5.64
PubMed-Author-subject	13	59	901	0.18	18%	0.01	1%	0.17
Citation(Web of Science)	14	58	42	0.19	19%	0.25	25%	8.83
Subject search(e-library)	12	60	279	0.01	1%	0.13	13%	2.21
SCOPUS Author-subject (Author(s))	12	60	38	0.17	17%	0.24	24%	8.11
SCOPUS citation	--	72	--	--	--	--	--	--
CINAHL Author-subject	--	72	--	--	--	--	--	--
CINAHL citation	--	72	--	--	--	--	--	--
WoK-Author-subject	5	67	9	0.07	7%	0.36	36%	13.01
E-lib – Author-subject	9	63	39	0.13	13%	0.19	19%	5.64
PubMed-Economics	5	67	4	0.07	7%	0.56	56%	29.37
PubMed -Costs	5	67	4	0.07	7%	0.56	56%	29.37
PubMed -Qualitative	5	67	23	0.07	7%	0.18	18%	5.05
E-lib-Economics	21	51	221	0.29	29%	0.09	9%	2.14
E-lib-Costs	10	62	158	0.14	14%	0.06	6%	1.45
E-lib-Qualitative	12	60	166	0.17	17%	0.07	7%	1.70
WoK-Economics	5	67	3	0.07	7%	0.63	63%	39.18
WoK-Costs	5	67	3	0.07	7%	0.63	63%	39.18
WoK-Qualitative	4	68	27	0.06	6%	0.13	13%	3.38
SCOPUS- Qualitative	2	70	2	0.03	3%	0.50	50%	22.51
SCOPUS-Costs	5	67	4	0.07	7%	0.56	56%	29.37
SCOPUS-Economics	6	66	1	0.08	8%	0.86	86%	143.36
CINAHL- Qualitative	--	72	--	--	--	--	--	--
CINAHL-Costs	--	72	--	--	--	--	--	--
CINAHL-Economics	--	72	--	--	--	--	--	--

Total relevant R without duplicates	**72**
Total retrieved without duplicate	**1650**
Total non-relevant R without duplicates (**D**)	**1578**
Odds Estimator = A * (D – B) / B * C	

5.3.4.4 Sibling Studies Retrieval by Study Type

Furthermore, each relevant study being retrieved was categorised into one of the four sibling types (RCT, qualitative, process evaluation and economical evaluation). Table 15 shows that search strategy and database performed differently for the four siblings types. Consequently, for the RCTs siblings, subject search with Hedges qualitative filter on SCOPUS performed the best with odds estimator of 36.65 indicating the best likelihood of retrieving the RCTs rather than non-RCTs, followed by WoK author-subject search (odds estimator of 12.45). Subject search with Hedges economics filter and Hedges costs filter on

WoK were third (both with odds estimator value of 11.93). The search with the least likelihood of retrieving RCTs siblings was author-subject search on PubMed with odds estimator of 0.43, regardless of the searches that did not even retrieve any studies.

There were only six qualitative studies retrieved for the IVF seed study. Three studies were retrieved by WoK subject search with Hedges qualitative filter, scoring odds estimator value of 57.44 pointing to the best possibility of obtaining the qualitative studies rather than non-qualitative ones. Subject search on PubMed with qualitative Hedges filter was the second, retrieving 2 qualitative studies with odds estimator value of 33.08, and finally, SCOPUS subject-search was third (odds estimator value of 8.11) as indicated in Table 15. Moreover, most of search strategies did not retrieve any qualitative studies scoring odds estimator of zero.

Furthermore, the best likelihood of retrieving the economics siblings was achieved by SCOPUS subject search with Hedges economics filter with odds estimator value of 985.63 indicating the best possibility of retrieving economic studies rather than non-economic studies, while subject search on SCOPUS with Hedges economics filter was second (odds estimator value of 245.94). WoK subject search with Hedges economics filter and WoK subject search with Hedges costs filter scored third (both with odds estimator value of 233.33). There was several search strategies did not any economics studies, i.e. WoK subject search with Hedges qualitative filter and subject search with Hedges qualitative filter on PubMed indicating the least likelihood of retrieving the economic studies, as illustrated in Table 15.

And finally for process evaluation siblings, WoS citation search scored the highest odds estimator with a value of 36.57 indicating the highest probability of retrieving process evaluation studies rather than non-process evaluation studies, and subject search on e-library (using hMP search term) was second (with odds estimator of 22.4). Author-subject on SCOPUS was third (odds estimator value of 5.79) (Table 15). Retrieving process evaluation studies seems to be difficult as well as the qualitative, as mentioned earlier, with this type of studies there was several search strategies which did not retrieve any process evaluation studies, i.e. WoK author-subject.

For the IVF seed study, many search strategies did not retrieve any studies, which might suggests that this particular seed study (effect of clinical topic or focus?) might have an influence on search performance for both search strategies and databases (See Section 7.3).

Table 15: IVF Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	**R/N**	**OE**	**RR**	**R/N**	**OE**	**RR**	**R/N**	**OE**	**RR**	**R/N**	**OE**
Related Search(PubMed)	15	30	8.47	0	6	0	2	11	3.08	1	7	2.42
Author-Subject(PubMed)	9	36	0.19	0	6	0	3	10	0.23	1	7	0.11
Citation(Web of Science)	8	37	7.91	0	6	0	2	11	6.65	4	4	36.57
Subject search(e-library, hMP)	2	43	7.29	0	6	0	1	12	13.07	1	7	22.4
Subject search(e-library, rFSH)	8	37	1	0	6	0	2	11	0.85	1	7	0.67
SCOPUS Author-subject	8	37	8.76	1	5	8.11	2	11	7.37	1	7	5.79
SCOPUS citation	--	--	--	--	--	--	--	--	--	--	--	--
CINAHL Author-subject	--	--	--	--	--	--	--	--	--	--	--	--
CINAHL citation	--	--	--	--	--	--	--	--	--	--	--	--
WoK-Author-subject	3	42	12.45	0	6	0	2	11	31.7	0	8	0
E-lib - Author-subject	6	39	6.07	0	6	0	3	10	11.84	0	8	0
Economics-Hedges filter	1	44	8.94	0	6	0	4	9	174.89	0	8	0
Costs-Hedges filter	1	44	8.94	0	6	0	4	9	174.89	0	8	0
Qualitative-Hedges filter	3	42	4.83	2	4	33.8	0	13	0	0	8	0
E-lib-Economics	11	34	1.99	0	6	0	9	4	13.82	1	7	0.88
E-lib-Costs	6	39	1.38	0	6	0	4	9	3.99	0	8	0
E-lib-Qualitative	7	38	1.57	1	5	1.7	1	12	0.71	3	5	5.1
WoK-Economics	1	44	11.93	0	6	0	4	9	233.33	0	8	0
WoK-Costs	1	44	11.93	0	6	0	4	9	233.33	0	8	0
WoK-Qualitative	1	44	1.31	3	3	57.44	0	13	0	0	8	0
SCOPUS- Qualitative	2	43	36.65	0	6	0	0	13	0	0	8	0
SCOPUS-Costs	0	45	0	0	6	0	5	8	245.94	0	8	0
SCOPUS-Economics	0	45	0	0	6	0	5	8	985.63	0	8	0
CINAHL- Qualitative	--	--	--	--	--	--	--	--	--	--	--	--
CINAHL-Costs	--	--	--	--	--	--	--	--	--	--	--	--
CINAHL-Economics	--	--	--	--	--	--	--	--	--	--	--	--

5.3.5 Chronic Lung Disease (Dexamethasone) Seed Study

The search for chronic lung disease sibling studies returns 1011 studies. 153 studies were classed as relevant studies with a very strong possibility of being one of the direct siblings for the seed study, from all the different search strategies and databases.

5.3.5.1 Recall

As demonstrated in Table 16, related articles search on PubMed retrieved the highest number of related studies, 52 relevant studies, scoring the highest recall value of 34%. Author-subject search on E-library with Hedges qualitative filter retrieved 46 studies (recall value of 32%). The third search that retrieved the third highest number of relevant studies was subject search on e-library yielding 33 relevant studies (recall value of 22%).

5.3.5.2 Precision

Precision calculations revealed that CINAHL author-subject, WoK author-subject, as well as subject on WoK with Hedges economics filter performed best, all scoring the highest precision with score of 100% each, however, as might be expected, the number of relevant studies these strategies retrieved is lower. The second best precision was achieved by SCOPUS author-subject (Author(s)) and CINAHL subject search with either qualitative or economics Hedges filter with value of 75% and again a low number of relevant studies being retrieved (only 3 relevant studies), followed by subject search on e-library (third highest precision with value of 62%) as shown in Table 16.

5.3.5.3 Odds Estimator

As shown in Table 16, WoK author-subject scored the highest odds estimator value of 34.98 indicating the best likelihood of retrieving relevant rather than non-relevant studies, followed by CINAHL author-subject with odds estimator value of 28.95. The third best odds estimator was scored by SCOPUS author-subject and CINAHL subject search with either qualitative or economics Hedges filter with odds estimator value of 17.14. The search with the least likelihood of retrieving relevant rather than non-relevant studies was the CINAHL subject search with the Hedges costs filter subject search with odds estimator value of zero.

Table 16: The CLD Search Strategies and Databases Retrieval

CLD Infant Infection (Dexamethasone)								
Search Strategy	**Relevant/R (A)**	**Relevant N/R©**	**Not relevant (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Estimator**
PubMed-Related articles	52	101	675	0.34	34%	0.07	7%	0.14
PubMed-Author-subject	3	150	36	0.02	2%	0.08	8%	0.46
Citation(Web of Science)	5	148	30	0.03	3%	0.14	14%	0.93
Subject search(e-library)	33	120	20	0.22	22%	0.62	62%	11.52
SCOPUS Author-subject (Group)	9	144	29	0.06	6%	0.24	24%	1.79
SCOPUS Author-subject (Author(s))	3	150	1	0.02	2%	0.75	75%	17.14
SCOPUS citation	20	133	24	0.13	13%	0.45	45%	5.23
CINAHL Author-subject	5	148	0	0.03	3%	1.00	100%	28.95
CINAHL citation	1	152	1	0.01	1%	0.50	50%	5.64
WoK-Author-subject	6	147	0	0.04	4%	1.00	100%	34.98
E-lib – Author-subject	2	151	36	0.01	1%	0.05	5%	0.30
PubMed-Economics	3	150	4	0.02	2%	0.43	43%	4.27
PubMed -Costs	2	151	3	0.01	1%	0.40	40%	3.77
PubMed -Qualitative	6	147	5	0.04	4%	0.55	55%	6.96
E-lib-Economics	26	127	31	0.17	17%	0.46	46%	2.84
E-lib-Costs	11	142	9	0.07	7%	0.55	55%	7.31
E-lib-Qualitative	49	104	41	0.32	32%	0.54	54%	9.39
WoK-Economics	1	152	0	0.01	1%	1.00	100%	5.64
WoK-Costs	1	152	3	0.01	1%	0.25	25%	1.88
WoK-Qualitative	6	147	12	0.04	4%	0.33	33%	2.88
SCOPUS- Qualitative	6	147	30	0.04	4%	0.17	17%	1.13
SCOPUS-Costs	3	150	22	0.02	2%	0.12	12%	0.76
SCOPUS-Economics	3	150	10	0.02	2%	0.23	23%	1.70
CINAHL- Qualitative	3	150	1	0.02	2%	0.75	75%	17.14
CINAHL-Costs	0	153	1	0.00	0%	0.00	0%	0.00
CINAHL-Economics	3	150	1	0.02	2%	0.75	75%	17.14

Total relevant R without duplicates	**153**
Total retrieved without duplicate	**1011**
Total non-relevant R without duplicates (**D**)	**858**
Odds Estimator = A * (D – B) / B * C	

5.2.5.4 Sibling Studies Retrieval by Study Type

Each relevant study was categorised into one of the four sibling type (RCT, quantitative, economics and process evaluation). Table 17 demonstrates the performance of each search strategy and database in retrieving a specific type of siblings, with specific pattern to be noticed. Therefore, The RCTs siblings were best retrieved by WoK author-subject search, odds estimator value of 47.61, followed by author-subject search on SCOPUS with odds estimator of 23.16 as the second best likelihood of retrieving the RCTs rather than non-

RCTs. Subject search on CINAHL with Hedges economics filter was third (odds estimator with value of 15.3). The least likelihood of retrieving RCTs studies were author-subject on PubMed, CINAHL citation, subject search on PubMed with Hedges costs filter and subject search on CINAHL with Hedges costs filter as indicated by odds estimator value of zero.

For the qualitative studies, CINAHL subject search with Hedges qualitative filter was the most likely search to retrieve qualitative studies rather than non-qualitative as indicated by odds estimator with value of 142.83, followed by subject search on PubMed with Hedges qualitative filter (odds estimator of 28.43) and third was author-subject search on E-library with Hedges costs filter (odds estimator value of 18.56). And finally, there was more than one search strategy and databases with odds estimator value of zero indicating least possibility of retrieving qualitative studies rather then non-qualitative i.e. author-subject on PubMed and SCOPUS author-subject (Author(s)) search as indicated in Table 17.

There was three economics studies retrieved for CLD seed study. Subject search on PubMed with Hedges costs filter retrieved 2 of these siblings scoring the highest odds estimator value of 570, followed by subject search on PubMed with Hedges economics filter (odds estimator with value of 427). Author-subject search on E-library with Hedges economics filter scored third (odds estimator with value of 53.56) as shown in Table 17.

Process evaluation siblings were best retrieved by SCOPUS author-subject (Group name) search with odds estimator value of 223.57 indicating the highest possibility of retrieving process evaluation studies rather than non-process evaluation, followed by SCOPUS citation (odds estimator with value of 137.12), and third was CINAHL citation and subject search on CINAHL with either Hedges qualitative or economics filter with odds estimator value of 30.61. Many search strategies had an odds estimator value of zero for process evaluation siblings as illustrated in Table 17.

Table 17: CLD Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	41	73	0.15	4	3	0.36	0	3	0	7	22	0.09
PubMed-Author-subject	0	114	0	0	7	0	1	2	11.42	2	27	1.69
Citation(Web of Science)	3	111	0.75	0	7	0	0	3	0	2	27	2.04
Subject search(e-library)	30	84	14.96	0	7	0	0	3	0	3	26	4.84
SCOPUS Author-subject (Group)	3	111	23.16	0	7	0	0	3	0	6	23	223.57
SCOPUS Author-subject (Author(s))	2	112	0.51	1	6	4.76	0	3	0	0	29	0
SCOPUS citation	18	96	7.58	0	7	0	0	3	0	2	27	137.12
CINAHL Author-subject	1	113	6.52	0	7	0	0	3	0	4	25	2.57
CINAHL citation	0	114	0	0	7	0	0	3	0	1	28	30.61
WoK-Author-subject	6	108	47.61	0	7	0	0	3	0	0	29	0
E-lib – Author-subject	2	112	0.41	0	7	0	0	3	0	0	29	0
PubMed-Economics	1	113	1.89	0	7	0	2	1	427	0	29	0
PubMed -Costs	0	114	0	0	7	0	2	1	570	0	29	0
PubMed -Qualitative	2	112	3.05	1	6	28.43	0	3	0	3	26	19.69
E-lib-Economics	20	94	5.68	0	7	0	2	1	53.36	4	25	4.27
E-lib-Costs	7	107	7.28	1	6	18.55	0	0	0	3	26	12.85
E-lib-Qualitative	30	84	7.12	3	4	14.95	1	2	9.96	12	17	14.06
WoK-Economics	1	113	7.58	0	7	0	0	3	0	0	29	0
WoK-Costs	1	113	2.52	0	7	0	0	3	0	0	29	0
WoK-Qualitative	3	111	1.91	1	6	11.75	0	3	0	2	27	5.22
SCOPUS- Qualitative	2	112	0.49	1	6	4.59	0	3	0	3	26	3.17
SCOPUS-Costs	2	112	0.68	0	7	0	1	2	19	0	29	0
SCOPUS-Economics	2	112	1.51	0	7	0	1	2	42.4	0	29	0
CINAHL- Qualitative	1	113	7.58	1	6	142.83	0	3	0	1	28	30.61
CINAHL-Costs	0	114	0	0	7	0	0	3	0	0	29	0
CINAHL-Economics	2	112	15.3	0	7	0	0	3	0	1	28	30.61

5.3.6 Summary: Reflection on Retrieval Performance of Search Strategies and Databases

Table 18 summarises the performance retrieval results detailed in previous sections. It presents the top three performing search strategies/databases for each seed study. The measurements matrices used in this research were the standard information retrieval performance measurements of recall and precision for search strategies/database performance measures, a third measure of odds estimator focuses on ranking search strategies and/or databases based on their efficiency of retrieving relevant and/or siblings rather than non-relevant and/or non-siblings. According to the summary table below, the search filters appeared to have enhanced the precision rather than recall, as reflected in both the precision and odds estimator figures. This at least demonstrates that the filters are working, but the filters also seem to boost recall for some of the seed studies. However the recall is still considered to be low (maximum of 34%). This called for combining the top search strategies to make the search more comprehensive and increase the recall (See Section 7.5). Table 18 shows no stable retrieval pattern. The performance results varied among the seeds studies, which indicate the effect of clinical area on the retrieval. Moreover this might have to do with the database coverage of the seed study clinical area and the type of siblings, i.e. CINAHL is expected to index more qualitative siblings rather than the other types of siblings (See Section 7.3).

Table 18: Top Three Search Strategies and Databases in Retrieving Relevant (on-topic) Studies

DEATeL		*Tamoxifen*		*BMS*		*CLD*		IVF	
Recall									
E- library author-subject - (Economics & Qualitative)	27%	SCOPUS author-subject	20%	WoK author-subject	21%	PubMed related articles	34%	E-library author-subject- Economics	29%
E-library author-subject -Costs	23%	WoK author-subject	19%	subject search(e-library) PubMed related articles	17%	E-library Author-subject- Qualitative	32%	PubMed related articles	25%
Author-subject on WoK	16%	subject search(e-library)	15%	E-library Author-subject, WoK subject-Qualitative, CINAHL- Subject-Qualitative & SCOPUS subject-Qualitative	8%	subject search(e-library)	22%	WoS citation	19%
Precision									
CINAHL author-subject, SCOPUS- Qualitative & CINAHL (Qualitative, Costs & Economics)	100%	WoK author-subject	54%	SCOPUS Author-subject (Author(s)) CINAHL Author-subject, CINAHL citation & SCOPUS-Qualitative	100 %	CINAHL author-subject, WoK author-subject WoK subject (economics filter)	100%	SCOPUS-Economics	86%
WoK author-subject	89%	CINAHL author-subject	50%	WoK Subject-Qualitative	75%	SCOPUS author-subject (Author(s)), CINAHL (Qualitative & Economics)	75%	WoK subject search (economics & costs filters)	62%
SCOPUS author-subject & SCOPUS-Costs	75%	SCOPUS -Economics	45%	WoK author-subject search PubMed subject-Qualitative	71%	subject search(e-library)	62%	PubMed subject (economics & costs) & SCOPUS-Costs	56%
Odds Estimator									
CINAHL author-subject		CINAHL author-subject		SCOPUS-Qualitative		WoK author-subject		SCOPUS-Economics	
SCOPUS-Qualitative		WoK author-subject		WoK subject-Qualitative		CINAHL author-subject		WoK subject search (economics & costs filters)	
CINAHL-(Qualitative & Economics)		SCOPUS-Economics		WoK Author-subject		SCOPUS author-subject (Author(s)) & CINAHL(Qualitative & Economics)		SCOPUS-Costs & PubMed (economics and costs)	

5.4 Indirect Siblings

To prepare for the final stage of this project several identification criteria for siblings were created and applied to the related sets in order to extract studies that qualify as siblings according to those inclusion criteria (Section 4.8):

- Relevant study that appeared in the reference list of the seed study;
- Relevant studies that came up with a citation search strategy.

At this stage the inclusion criteria characteristics were intended to narrow down the possibilities, to the likelihood that the sibling is indirect sibling and not simply a relevant item. Sibling results are presented using recall, precision and finally odds estimator as this study's main measurements metrics. And finally, the term siblings in this section refers to studies that met this stage's inclusion criteria, meaning that each study is a possible candidate to be direct sibling (same principles as the seed study with at least one shared author (Section 4.9)) and if not it would be classified as a indirect siblings (same principles as the seed study with no shared author (Section 4.8)); a term I chose for studies that are very closely related to specific seed study.

5.4.1 Diabetes Telemedicine (IDEATeL) Seed Study

IDEATeL seed study search strategies and databases yielded 394 studies that were recognised as relevant to the seed study from the entire retrieval pool, but not all these relevant items were siblings. After applying the inclusion criteria 93 studies were considered as siblings, either direct or indirect.

5.4.1.1 Recall

As shown in Table 19, WoK author-subject search identified 46 (out of 55 relevant and retrieved, see Table 6) sibling studies scoring the best recall with value of 49%. SCOPUS author-subject search identified 39 (out of 39 relevant and retrieved, see Table 8) sibling s scoring the second highest best recall value of 42%. E-library author-subject retrieved 36 (of 38 relevant and retrieved, see Table 8) siblings scoring the third best recall value of 39%.

5.4.1.2 Precision

CINAHL Author-subject, CINAHL subject search with Hedges qualitative, costs or economics filters and SCOPUS subject search with Hedges qualitative filter scored the best precision value of 100%, followed by author-subject on WoK with precision value of 79%, SCOPUS author-subject and SCOPUS subject with Hedges costs filter search each scored the third best precision value of 75% (Table 19).

5.4.1.3 Odds Estimator

CINAHL Author subject search did not retrieve any non-sibling studies causing some difficulties in odds estimator calculation due to the division by zero mathematical error, however this problem was solved (See Section 4.5). CINAHL Author subject search scored the highest odds estimator indicating the highest possibility of retrieving sibling studies rather than non-siblings with value of 321.2. SCOPUS subject search with Hedges qualitative filter scored the second best odds estimator with value of 221.7, and third was CINAHL subject with Hedges qualitative or economics filter with odds estimator value of 86.28 (See Table 19).

Table 19: The IDEATeL Search Strategies and Databases Retrieval

Diabetes Telemedicine (IDEATeL)								
Search Strategy	**Siblings (A)**	**Siblings N/R©**	**Non Siblings (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Estimator**
PubMed-Related articles	23	70	163	0.25	25%	0.12	12%	2.23
PubMed-Author-subject	22	71	135	0.24	23%	0.14	14%	2.56
Citation(Web of Science)	14	79	1	0.15	15%	0.25	25%	4.98
Subject search (e-library)	29	64	266	0.32	32%	0.1	10%	1.77
SCOPUS Author-subject (Author(s))	39	54	13	0.42	42%	0.75	75%	68.83
SCOPUS citation	29	64	35	0.31	31%	0.45	45%	15.76
CINAHL Author-subject	19	74	0	0.2	20%	1	100%	321.2
CINAHL citation	4	89	3	0.04	4%	0.57	57%	18.71
WoK-Author-subject	46	44	16	0.49	49%	0.79	79%	75.61
E-lib – Author-subject	36	57	19	0.39	39%	0.65	65%	40.99
PubMed-Economics	7	86	78	0.08	8%	0.08	8%	1.23
PubMed -Costs	8	85	69	0.09	9%	0.1	10%	1.61
PubMed -Qualitative	18	75	67	0.19	19%	0.21	21%	4.53
E-lib-Economics	24	69	205	0.26	26%	0.1	10%	1.78
E-lib-Costs	17	76	197	0.18	18%	0.08	8%	1.2
E-lib-Qualitative	18	75	250	0.19	19%	0.07	7%	0.96
WoK-Economics	4	89	58	0.04	4%	0.06	6%	0.93
WoK-Costs	10	83	89	0.11	11%	0.1	10%	1.57
WoK-Qualitative	20	73	74	0.22	22%	0.21	21%	4.36
Sophisticated	15	78	94	0.16	16%	0.14	14%	2.37
Sophisticated2 – Subject	33	60	471	0.35	35%	0.07	7%	3.15
Sophisticated2 – Author-Subject	32	61	36	0.34	34%	0.47	47%	46.39
SCOPUS- Qualitative	14	79	0	0.15	15%	1	100%	221.7
SCOPUS-Costs	3	90	1	0.03	3%	0.75	75%	41.7
SCOPUS-Economics	2	91	2	0.02	2%	0.50	50%	13.74
CINAHL- Qualitative	6	87	0	0.07	7%	1	100%	86.28
CINAHL-Costs	4	89	0	0.04	4%	1	100%	56.23
CINAHL-Economics	6	87	0	0.07	7%	1	100%	86.28

Total SIBLINGS R without duplicates	93
Total NON-SIBLINGS R without duplicate	301
Total Relevant R without duplicates	394
Total non-relevant R without duplicates (D)	**1252**

Odds Estimator = A * (D - A)/ B * C

5.4.1.4 Sibling Studies Retrieval by Study Type

Sibling studies were investigated more thoroughly after identification in order to classify each sibling into one of our four sibling categories (RCT, qualitative, process evaluation and economics). According to Table 20, RCT siblings are best retrieved by CINAHL author-subject search as it scored the best odds estimator with value of 431.38 with more likelihood of retrieving RCTs rather than non-RCT siblings. WoK author-subject search scored second with odds estimator value of 81.32. SCOPUS subject search with Hedges qualitative scored third with odds estimator value of 67.62. SCOPUS subject search with Hedges economics or costs filter and CINAHL subject search with Hedges costs filter search scored the lowest odds estimator of zero.

According to Table 20, the qualitative siblings were best retrieved by SCOPUS subject search with Hedges qualitative filter, with odds estimator value of 625.5. CINAHL author-subject and CINAHL subject with Hedges qualitative filter odds estimator were second with odd estimator value of 250.2, followed by SCOPUS author-subject search with odds estimator value of 112.64 as third. Subject search with Hedges economics and costs filter on SCOPUS and subject search on CINAHL with Hedges costs filter did not retrieve any qualitative siblings, and therefore the odds estimator value was zero.

The CINAHL subject search with Hedges costs or economics filter scored the highest odds estimator for retrieving economics siblings rather than the other sibling types with value of 3753. Subject search on SCOPUS with Hedges costs filter scored second, with odds estimator value of 1251, while subject search on SCOPUS with Hedges economics filter scored third with odds estimator value of 625. Several search strategies did not retrieve any economics siblings scoring odds estimator value of zero, as illustrated in Table 20.

For process evaluation siblings the CINAHL author-subject search and SCOPUS subject search with Hedges qualitative filter each scored the best odds estimator with value of 227.46 indicating the highest likelihood of retrieving process evaluation rather than non-process evaluation siblings. Subject search on SCOPUS with Hedges costs and CINAHL subject search with Hedges qualitative, costs or economics filter scored the second with odds estimator value of 50. SCOPUS author-subject search was the third best with odds estimator value of 44.36. Subject search with Hedges economics filter on either SCOPUS or WoK did not retrieve any qualitative siblings with odds estimator value of zero (Table 20).

Table 20: The IDEATeL Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
Related Search(PubMed)	8	31	1.72	6	19	2.23	1	3	2.23	8	18	2.97
Author-Subject(PubMed)	6	33	1.5	10	15	5.91	3	1	24.82	3	23	1.08
Citation(Web of Science)	4	35	3.21	6	19	9.37	2	2	28.12	2	24	2.34
Subject search(e-library)	12	27	1.65	10	15	2.65	2	2	3.71	6	20	1.11
SCOPUS Author-subject	16	23	66.3	13	12	112.64	2	2	95.31	8	18	42.36
SCOPUS citation	16	23	24.19	7	18	14.32	2	2	34.77	4	22	6.32
CINAHL Author-subject	10	29	431.38	4	21	250.2	1	3	417	4	22	227.46
CINAHL citation	2	37	22.51	1	24	18.1	0	4	0	1	25	16.65
WoK-Author-subject	20	19	81.32	15	10	108.15	3	1	231.75	9	17	40.9
E-lib - Author-subject	17	22	50.15	11	14	54.91	1	3	21.63	7	19	23.91
Economics-Hedges filter	2	37	0.81	2	23	1.37	1	3	5.02	2	24	1.25
Costs-Hedges filter	2	37	0.93	1	24	0.75	3	1	51.44	2	24	1.43
Qualitative-Hedges filter	4	35	2.16	6	19	6.29	0	4	0	8	18	8.39
E-lib-Economics	6	33	0.93	8	17	2.55	0	4	0	10	16	3.19
E-lib-Costs	5	34	0.79	5	20	1.41	2	2	5.36	5	21	1.28
E-lib-Qualitative	4	35	0.46	5	20	1.06	3	1	12.02	6	20	1.2
WoK-Economics	2	37	1.11	0	25	0	2	2	20.59	0	26	0
WoK-Costs	5	34	1.92	0	25	0	2	2	13.07	3	23	1.7
WoK-Qualitative	5	34	2.34	6	19	5.31	0	4	0	9	17	8.43
Sophisticated	8	31	3.18	2	23	1.12	2	2	12.32	3	23	1.61
Sophisticated2 - Subject	13	26	3.28	10	15	4.68	3	1	19.66	7	19	2.42
Sophisticated2 – Author-Subject	13	26	44.21	10	15	63.16	2	2	88.43	7	19	32.58
SCOPUS-Qualitative	2	37	67.62	8	16	625.5	0	4	0	4	22	227.46
SCOPUS-Costs	0	39	0	0	24	0	2	2	125	1	25	50.04
SCOPUS-Economics	0	39	0	0	24	0	2	2	625	0	26	0
CINAHL-Qualitative	1	38	32.92	4	20	250.2	0	4	0	1	25	50.04
CINAHL-Costs	0	39	0	0	24	0	3	1	3753	1	25	50.04
CINAHL-Economics	1	38	32.92	1	23	54.39	3	1	3753	1	25	50.04

5.4.2 Breast Cancer (Tamoxifen) Seed Study

The search for the Tamoxifen seed study resulted in 383 of relevant studies, and 132 studies from the relevant set met the inclusion criteria for siblings, therefore were marked/categorised as siblings of the Tamoxifen seed study.

5.4.2.1 Recall

As Table 21 shows, SCOPUS author-subject search retrieved the highest number of siblings with value of 77 siblings (from 78 relevant studies retrieved), scoring the highest recall value of 58%, followed by WoK author-subject retrieving 72 siblings(72 relevant items) with recall value of 55%. The third best recall was scored by author-subject on PubMed retrieving 23 siblings and scoring recall value of 17%.

5.4.2.2 Precision

CINAHL Author subject search scored the highest precision with value of 54%. WoK author-subject precision was the second highest with value of 49%. SCOPUS Author-subject scored the third best precision with value of 34%. The precision values for this seed study are low compared to the IDEATeL seed study (See Table 21).

5.4.2.3 Odds Estimator

WoK-Author – subject odds estimator was the highest value, indicating the highest likelihood of retrieving siblings rather than non-siblings with value of 26.16, followed by author-subject on CINAHL scoring second (odds estimator value of 17.8) and third was SCOPUS author-subject with odds estimator value of 14.4. Several search strategies did not retrieve any direct or indirect siblings and accordingly resulted in odds estimator of zero (See Table 21).

Table 21: Breast Cancer (Tamoxifen): Search and Databases Retrieval Performance

Search Strategy	Siblings (A)	Siblings N/R©	Non Siblings (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
PubMed-Related articles	9	123	191	0.07	7%	0.05	5%	0.61
PubMed-Author-subject	23	109	428	0.17	17%	0.05	5%	0.63
Citation(Web of Science)	7	125	46	0.05	5%	0.13	13%	2.03
Subject search(e-library)	6	126	282	0.05	5%	0.02	2%	0.24
SCOPUS Author-subject (Author(s))	77	55	152	0.58	58%	0.34	34%	14.4
SCOPUS citation	4	128	55	0.03	3%	0.07	7%	0.94
CINAHL Author-subject	22	110	19	0.17	17%	0.54	54%	17.8
CINAHL citation	1	131	3	0.01	1%	0.25	25%	4.34
WoK-Author-subject	72	60	75	0.55	55%	0.49	49%	26.16
E-lib – Author-subject	4	128	39	0.03	3%	0.09	9%	1.34
PubMed-Economics	1	131	72	0.01	1%	0.01	1%	0.17
PubMed –Costs	1	131	65	0.01	1%	0.02	2%	0.19
PubMed -Qualitative	4	128	130	0.03	3%	0.03	3%	0.38
E-lib-Economics	6	126	207	0.05	5%	0.03	3%	0.35
E-lib-Costs	2	130	156	0.02	2%	0.01	1%	0.15
E-lib-Qualitative	3	129	212	0.02	2%	0.01	1%	0.16
WoK-Economics	3	129	136	0.02	2%	0.02	2%	0.27
WoK-Costs	3	129	96	0.02	2%	0.03	3%	0.39
WoK-Qualitative	2	130	204	0.02	2%	0.01	1%	0.11
SCOPUS- Qualitative	6	126	128	0.04	4%	0.05	5%	0.59
SCOPUS-Costs	0	132	89	0	0%	0	0%	0
SCOPUS-Economics	0	132	40	0	0%	0	0%	0
CINAHL- Qualitative	1	131	21	0.01	1%	0.05	5%	0.61
CINAHL-Costs	0	132	11	0	0%	0	0%	0
CINAHL-Economics	1	131	29	0.01	1%	0.03	3%	0.44

Total SIBLINGS R without duplicates	132
Total NON-SIBLINGS R without duplicate	249
Total Relevant R without duplicates	383
Total non-relevant R without duplicates (D)	**1710**

Odds Estimator = A * (D – B)/ B * C

5.4.2.4 Sibling Studies Retrieval by Study Type

As mentioned earlier, sibling studies were investigated further after identification in order to classify each sibling into one of the four siblings categories (RCT, qualitative, process evaluation and economics). Therefore, as shown in Table 22, WoK author-subject search scored the highest odds estimator of 58.13 indicating the best odds of retrieving RCT rather than non-RCT siblings. The SCOPUS author-subject odds estimator value of 27.33 was second and third was the CINAHL author-subject with odds estimator value of 25.22. According to Table 22, there was many search strategies which had zero values for the odds estimator.

As illustrated in Table 22, qualitative siblings were best retrieved by CINAHL citation with odds estimator value of 15.81, with the highest likelihood of retrieving qualitative siblings rather than non-qualitative siblings, author-subject on WoK odds estimator was second with value of 6, followed by WoS citation search with odds estimator value of 5.65. There was many search strategies with zero odds estimator values.

Subject search with Hedges costs filter on PubMed scored the highest odds estimator for retrieving economics siblings rather than the non-economics with a value of 25.31, while subject search with Hedges economics filter on PubMed scored second value of 22.75 and WoK subject with Hedges costs filter was third with odds estimator value of 16.81, as shown in Table 22.

For process evaluation siblings, CINAHL author-subject search scored the best odds estimator with value of 17.8 indicating the best likelihood of retrieving process evaluation rather than non-process evaluation siblings. Author-subject search on WoK was second, with odds estimator value of 17.44, and SCOPUS author-subject was third with odds estimator value of 16.11 as shown in Table 21. There was more than one search strategy that did not retrieve any economics or process evaluation siblings and hence the odds estimator value was zero, as shown in Table 22.

Table 22: The Tamoxifen Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
Related Search(PubMed)	2	75	0.21	6	31	1.54	0	2	0	1	17	0.47
Author-Subject(PubMed)	15	62	0.72	3	34	0.26	0	2	0	5	13	1.15
Citation(Web of Science)	0	77	0	5	32	5.65	0	2	0	2	16	4.52
Subject search(e-library)	4	73	0.28	2	35	0.29	0	2	0	0	18	0
SCOPUS Author-subject	56	21	27.33	10	27	3.8	0	2	0	11	7	16.11
SCOPUS citation	1	76	0.4	3	34	2.66	0	2	0	0	18	0
CINAHL Author-subject	17	60	25.22	2	35	5.09	0	2	0	3	15	17.8
CINAHL citation	0	77	0	1	36	15.81	0	2	0	0	18	0
WoK-Author-subject	56	21	58.13	8	29	6.01	0	2	0	8	10	17.44
E-lib - Author-subject	0	77	0	4	33	5.19	0	2	0	0	18	0
Economics-Hedges filter	0	77	0	0	37	0	1	1	22.75	0	18	0
Costs-Hedges filter	0	77	0	0	37	0	1	1	25.31	0	18	0
Qualitative-Hedges filter	2	75	0.35	1	36	0.37	0	2	0	1	17	0.77
E-lib-Economics	3	74	0.29	1	36	0.2	1	1	7.26	1	17	0.43
E-lib-Costs	1	76	0.13	0	37	0	1	1	9.96	0	18	0
E-lib-Qualitative	2	75	0.19	0	37	0	0	2	0	1	17	0.42
WoK-Economics	1	76	0.15	0	37	0	1	1	11.57	1	17	0.68
WoK-Costs	1	76	0.22	0	37	0	1	1	16.81	1	17	0.99
WoK-Qualitative	0	77	0	0	37	0	0	2	0	2	16	0.92
SCOPUS-Qualitative	2	75	0.33	2	35	0.71	0	2	0	2	16	1.55
SCOPUS-Costs	0	77	0	0	37	0	0	2	0	0	18	0
SCOPUS-Economics	0	77	0	0	37	0	0	2	0	0	18	0
CINAHL-Qualitative	0	77	0	0	37	0	0	2	0	1	17	4.73
CINAHL-Costs	0	77	0	0	37	0	0	2	0	0	18	0
CINAHL-Economics	0	77	0	0	37	0	0	2	0	1	17	3.4

5.4.3 Breast Cancer (BMS) Seed Study

The search for the breast cancer (BMS) seed study yielded 71 relevant studies, but only 10 studies of those relevant studies met the inclusion criteria for sibling studies.

5.4.3.1 Recall

The recall values which are presented in Table 23 showing that the recall ratio is very low for this seed study siblings retrieval. The highest recall value of 40% was scored by both e-library subject search and WoK author-subject search (4 siblings out of 12 and 15 relevant and retrieved respectively), followed by author-subject on PubMed with recall value of 30% (3 siblings of 3 relevant and retrieved studies, see Table 10). The third best recall was scored by several search strategies, all scoring a recall value of 20%.

5.4.3.2 Precision

According to Table 23, citation search on WoS score was the highest in term of precision with a value of 67%. The second best precision was scored by CINAHL author-subject and CINAHL citation with precision value of 50%. And finally, SCOPUS citation and SCOPUS subject search with Hedges qualitative filter scored the third best precision with value of 33%.

5.4.3.3 Odds Estimator

The best likelihood of retrieving the siblings rather than non-siblings was achieved by citation search on WoS with odds estimator value of 142.5, and author-subject search and citation search on CINAHL were second with odds estimator value of 63.33, while SCOPUS subject search with qualitative filter was third with odds estimator value of 35.44, as shown in Table 23.

Table 23: The BMS Search Strategies and Databases Retrieval

Breast Cancer (BMS)								
Search Strategy	**Siblings (A)**	**Siblings N/R©**	**Non Siblings (B)**	**Recall**		**Precision**		**Odds Estimator**
				Value	**%**	**Value**	**%**	
PubMed-Related articles	2	8	106	0.2	20%	0.02	2%	1.2
PubMed-Author-subject	3	7	64	0.3	30%	0.04	4%	3.4
Citation(Web of Science)	2	8	1	0.2	20%	0.67	67%	142.5
Subject search(e-library)	4	6	20	0.4	40%	0.17	17%	18.37
SCOPUS Author-subject (Group)	2	2	24	0	0	0.08	8%	22.8
SCOPUS Author-subject (Author(s))	0	10	1	0	0	0	0%	0
SCOPUS citation	1	9	2	0.1	10%	0.33	33%	31.61
CINAHL Author-subject	1	9	1	0.1	10%	0.5	50%	63.33
CINAHL citation	1	9	1	0.1	10%	0.5	50%	63.33
WoK-Author-subject	4	6	17	0.4	40%	0.19	19%	21.73
E-lib – Author-subject	2	8	81	0.2	20%	0.02	2%	1.51
PubMed-Economics	0	10	0	0	0%	0	0%	0
PubMed -Costs	0	10	0	0	0%	0	0%	0
PubMed -Qualitative	1	9	6	0.1	10%	0.14	14%	10.46
E-lib-Economics	0	10	0	0	0%	0	0%	0
E-lib-Costs	0	10	0	0	0%	0	0%	0
E-lib-Qualitative	0	10	0	0	0%	0	0%	0
WoK-Economics	0	10	0	0	0%	0	0%	0
WoK-Costs	0	10	0	0	0%	0	0%	0
WoK-Qualitative	2	8	6	0.2	20%	0.25	25%	23.54
SCOPUS- Qualitative	2	8	4	0.2	20%	0.33	33%	35.44
SCOPUS-Costs	0	10	0	0	0%	0	0%	0
SCOPUS-Economics	0	10	0	0	0%	0	0%	0
CINAHL- Qualitative	2	8	12	0.2	20%	0.14	14%	11.65
CINAHL-Costs	0	10	3	0	0%	0	0%	0
CINAHL-Economics	1	9	11	0.1	10%	0.08	8%	5.65

Total SIBLINGS R without duplicates	10
Total NON-SIBLINGS R without duplicate	61
Total Relevant R without duplicates	71
Total non-relevant R without duplicates (D)	**571**

Odds Estimator = A * (D – B)/ B * C

5.4.3.4 Sibling Studies Retrieval by Study Type

After constructing the siblings set, each sibling was categorised into one of the four sibling types, and the odds estimator for each sibling category were calculated, although it must be acknowledged that the low number of siblings makes the calculations unreliable. Table 24 shows that for the RCTs siblings, odds estimator values indicated that CINAHL author-subject search scored the best odds estimator, with odds estimator value of 285. Subject search on WoK with Hedges qualitative filter was next, with odds estimator of value of 94.17, followed by subject search on SCOPUS with Hedges qualitative filter with odds estimator value of 70.88. A qualitative filter would not normally be expected to retrieve RCTs preferentially!

As illustrated in Table 24, qualitative siblings were best retrieved by citation search on WoS, scoring odds estimator value of 190 and subject search on SCOPUS with Hedges qualitative filter scored the second best odds estimator value of 47.25, with the third best being author-subject search on WoK with odds estimator value of 32.59.

There were no economics siblings for breast cancer with Body-Mind-Spirit therapy seed study retrieved by any search strategy or database.

For process evaluation siblings, CINAHL citation and WoS citation were the best possible searches to retrieve process evaluation siblings rather than non-process evaluation with odds estimator with value of 285. SCOPUS citation scored the second best odds estimator with value of 142.25 followed by author-subject on WoK with value of 16.294 as shown in Table 24. According to Table 24 there was more than one search which did not retrieve any RCT, qualitative, economics and process evaluation siblings, e.g. Author-subject on SCOPUS. This might indicate that for this seed study and the nature of clinical area and the procedures employed in this type of siblings of this particular clinical area either introduces some difficulties in retrieving siblings if there are any, or that there are a very few siblings available.

Table 24: BMS Search Strategies and Databases Odds Estimator per Sibling Type.

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	1	2	2.19	0	4	0	0	0	0	1	2	2.19
PubMed-Author-subject	1	2	3.96	0	4	0	0	0	0	2	1	15.84
Citation(Web of Science)	0	3	0	1	3	190	0	0	0	1	2	285
Subject search(e-library)	1	2	13.78	2	2	27.55	0	0	0	1	2	13.78
SCOPUS Author-subject (Group)	1	2	11.4	0	4	0	0	0	0	1	2	11.4
SCOPUS Author-subject (Author(s))	0	3	0	0	4	0	0	0	0	0	3	0
SCOPUS citation	0	3	0	0	4	0	0	0	0	1	2	142.25
CINAHL Author-subject	1	2	285	0	4	0	0	0	0	3	0	0
CINAHL citation	0	3	0	0	4	0	0	0	0	1	2	285
WoK-Author-subject	1	2	16.29	2	2	32.59	0	0	0	1	2	16.29
E-lib – Author-subject	1	2	3.03	1	3	2.02	0	0	0	0	3	0
PubMed-Economics	0	3	0	0	4	0	0	0	0	0	3	0
PubMed -Costs	0	3	0	0	4	0	0	0	0	0	3	0
PubMed -Qualitative	0	3	0	1	3	31.67	0	0	0	0	3	0
E-lib-Economics	0	3	0	0	4	0	0	0	0	0	3	0
E-lib-Costs	0	3	0	0	4	0	0	0	0	0	3	0
E-lib-Qualitative	0	3	0	0	4	0	0	0	0	0	3	0
WoK-Economics	0	3	0	0	4	0	0	0	0	0	3	0
WoK-Costs	0	3	0	0	4	0	0	0	0	0	3	0
WoK-Qualitative	2	1	94.17	0	4	0	0	0	0	0	2	0
SCOPUS- Qualitative	1	2	70.88	1	3	47.25	0	0	0	0	3	0
SCOPUS-Costs	0	3	0	0	4	0	0	0	0	0	3	0
SCOPUS-Economics	0	3	0	0	4	0	0	0	0	0	3	0
CINAHL- Qualitative	1	2	23.29	1	3	15.53	0	0	0	0	3	0
CINAHL-Costs	0	3	0	0	4	0	0	0	0	0	3	0
CINAHL-Economics	0	3	0	1	3	16.97	0	0	0	0	3	0

5.4.4 Chronic Lung Disease (Dexamethasone) Seed Study

The total number of studies that are relevant to CLD seed study was 153, after applying sibling studies identification and inclusion criteria 71 studies were identified and marked as siblings of this seed study.

5.4.4.1 Recall

The recall ratios for CLD siblings retrieval was slightly low as the highest recall value was 44% and was scored by related articles search on PubMed (with 31 siblings, out of 52 relevant and retrieved, see Table 14), followed by subject search on e-library with recall value of 35% (25 siblings out of 33 relevant and retrieved). SCOPUS citation search scored third with recall value of 28%, as illustrated in Table 25.

5.4.4.2 Precision

SCOPUS author-subject (first author), CINAHL author-subject and WoK-author-subject search all did not retrieve any non-siblings resulting in a complete set of relevant studies which are siblings as well, therefore their precision value was 100%. CINAHL citation, CINAHL subject search with Hedges qualitative or economics filters scored the second best precision with value of 50% while subject search on e-library scored third with a precision value of 47% as shown in Table 25.

5.4.4.3 Odds Estimator

According to Table 25, odds estimator calculations reveal that author-subject search on CINAHL is the most likely search to retrieve sibling studies rather than non-siblings with odds estimator value of 71.14. The second was scored by SCOPUS author-subject and WoK author-subject with odds estimator value of 56.06. Finally, subject search on e-library scored the third best odds estimator with value of 17.7.

Table 25: The CLD Search Strategies and Databases Retrieval

Chronic Lung Disease (CLD)								
Search Strategy	Siblings (A)	Siblings N/R©	Non Siblings (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
PubMed-Related articles	31	40	696	0.44	44%	0.04	4%	0.3
PubMed-Author-subject	3	68	36	0.04	4%	0.08	8%	1.11
Citation(Web of Science)	5	66	30	0.07	7%	0.14	14%	2.3
Subject search(e-library)	25	46	28	0.35	35%	0.47	47%	17.7
SCOPUS Author-subject (Group)	6	65	32	0.09	9%	0.16	16%	2.62
SCOPUS Author-subject (Author(s))	4	67	0	0.06	6%	1	100%	56.06
SCOPUS citation	20	51	24	0.28	28%	0.45	45%	14.97
CINAHL Author-subject	5	66	0	0.07	7%	1	100%	71.14
CINAHL citation	1	70	1	0.01	1%	0.5	50%	13.41
WoK-Author-subject	6	65	0	0.09	9%	1	100%	56.06
E-lib – Author-subject	2	69	36	0.03	3%	0.05	5%	0.73
PubMed-Economics	2	69	5	0.03	3%	0.29	29%	5.42
PubMed -Costs	1	70	4	0.01	1%	0.2	2%	3.34
PubMed -Qualitative	3	68	8	0.04	4%	0.27	27%	5.14
E-lib-Economics	4	67	53	0.06	6%	0.07	7%	1
E-lib-Costs	1	70	19	0.01	1%	0.05	7%	0.69
E-lib-Qualitative	9	62	81	0.13	13%	0.1	10%	1.54
WoK-Economics	0	71	1	0.00	0%	0	0%	0
WoK-Costs	0	71	4	0.00	0%	0	0%	0
WoK-Qualitative	3	68	15	0.04	4%	0.17	17%	2.72
SCOPUS- Qualitative	2	69	34	0.03	3%	0.06	6%	0.77
SCOPUS-Costs	2	69	23	0.03	3%	0.08	8%	1.16
SCOPUS-Economics	1	70	12	0.01	1%	0.08	8%	1.11
CINAHL- Qualitative	2	69	2	0.03	3%	0.5	50%	13.6
CINAHL-Costs	0	71	1	0	0%	0	0%	0
CINAHL-Economics	2	69	2	0.03	3%	0.5	50%	13.6

Total SIBLINGS R without duplicates	71
Total NON-SIBLINGS R without duplicate	82
Total Relevant R without duplicates	153
Total Non-Relevant R without duplicate (D)	**940**

Odds Estimator = A * (D – B)/ B * C

5.4.4.4 Sibling Studies Retrieval by Study Type

Table 26 shows that WoK author-subject was the search with the highest possibility of retrieving RCT siblings rather than non-RCT siblings as indicated by its odds estimator value of 106.3, followed by SCOPUS author-subject search with odds estimator value of 68.29, and the third best odds estimator value of 22.34 was scored by subject search on e-library.

There was only one qualitative sibling that was retrieved, and subject search with qualitative filter on CINAHL odds estimator performed best with odds estimator of 469, followed by subject search with Hedges qualitative filter on PubMed (odds estimator 117.13) and third best likelihood of retrieving qualitative rather than non-qualitative was with subject search with Hedges qualitative filter on WoK with odds estimator value of 61.67 on as illustrated in Table 26. According to the results it appears that search strategies with the Hedges qualitative filter were the only search strategies that retrieved the qualitative study, and this is investigated further (See Section 7.2.3.2).

There were only two direct and indirect economics siblings for the CLD seed study which were retrieved by three search strategies only, subject with Hedges costs filter on PubMed, subject with Hedges economics filter on PubMed and author-subject on PubMed with odds estimator value of 234, 187 and 25.11 respectively, see Table 26. It appears that adding economics search filter to simple subject with or without authors' names seems to be effective in retrieving economics sibling studies.

According to Table 26, CINAHL author-subject was the best search for retrieving process evaluation siblings scoring odds estimator value of 751.2, with citation search on CINAHL next with odds estimator value of 117.38, and the third was subject with Hedges qualitative filter on PubMed with odds estimator value of 14.641.

Table 26: CLD Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	27	32	0.3	1	0	0.35	0	2	0	3	6	0.18
PubMed-Author-subject	0	59	0	0	1	0	1	1	25.11	2	7	7.18
Citation(Web of Science)	3	56	1.63	0	1	0	0	2	0	2	7	8.67
Subject search(e-library)	24	35	22.34	0	1	0	0	2	0	1	8	4.07
SCOPUS Author-subject (Group)	4	55	2.06	0	1	0	0	2	0	2	7	8.11
SCOPUS Author-subject (Author(s))	4	55	68.29	0	1	0	0	2	0	0	9	0
SCOPUS citation	18	41	16.76	0	1	0	0	2	0	2	7	10.91
CINAHL Author-subject	1	58	16.2	0	1	0	0	2	0	4	5	751.2
CINAHL citation	0	59	0	0	1	0	0	2	0	1	8	117.38
WoK-Author-subject	6	53	106.3	0	1	0	0	2	0	0	9	0
E-lib – Author-subject	2	57	0.88	0	1	0	0	2	0	0	9	0
PubMed-Economics	1	58	3.22	0	1	0	1	0	187	0	9	0
PubMed -Costs	0	59	0	0	1	0	1	0	234	0	9	0
PubMed -Qualitative	1	58	2.02	1	0	117.13	0	2	0	1	8	14.64
E-lib-Economics	3	56	0.9	0	1	0	1	1	16.74	0	9	0
E-lib-Costs	1	58	0.84	0	1	0	0	2	0	0	9	0
E-lib-Qualitative	7	52	1.43	1	0	10.61	0	2	0	1	8	1.33
WoK-Economics	0	59	0	0	1	0	0	2	0	0	9	0
WoK-Costs	0	59	0	0	1	0	0	2	0	0	9	0
WoK-Qualitative	1	58	1.06	1	0	61.67	0	2	0	1	8	7.71
SCOPUS- Qualitative	1	58	0.46	1	1	26.65	0	2	0	0	9	0
SCOPUS-Costs	2	57	1.4	0	1	0	0	2	0	0	9	0
SCOPUS-Economics	1	58	1.33	0	1	0	0	2	0	0	9	0
CINAHL- Qualitative	1	58	8.09	1	1	469	0	2	0	0	9	0
CINAHL-Costs	0	59	0	0	1	0	0	2	0	0	9	0
CINAHL-Economics	2	57	16.46	0	1	0	0	2	0	0	9	0

5.4.5 In Vitro Fertilisation (IVF) Seed Study

The total number of relevant studies that were identified for the IVF seed study was 70 studies, 22 sibling studies were identified and extracted from the relevant set of studies.

5.4.5.1 Recall

Table 26 shows that related articles search on PubMed scored the best recall with value of 59% by retrieving 13 of the sibling studies (out of 18 relevant and retrieved, Table 14) followed by citation search on WoS with recall value of 55% (with 11 siblings out of 14, Table 12). Both author-subject search on PubMed and SCOPUS author-subject search scored the third best recall value by retrieving 11 siblings (of 13 and 12 relevant and retrieved respectively, Table 14) scoring a recall value of 50%.

5.4.5.2 Precision

Precision calculations indicated that subject search on e-library/MetaLib scored the best precision value of 82% followed by subject search with Hedges economics filter on SCOPUS with value of 43%, while author-subject on WoK scored the third with value of 25%, as shown in Table 27.

5.4.5.3 Odds Estimator

According to Table 27, citation search on WoS is the most likely search strategy to retrieve sibling studies rather than non-siblings with odds estimator of value 43.2. The second was SCOPUS author-subject search with odds estimator value of 40.74, followed by SCOPUS subject search with Hedges economics filter, with value of 36.56.

Table 27: The IVF Search Strategies and Databases Retrieval

In Vitro Fertilization (IVF)								
Search Strategy	**Siblings (A)**	**Siblings N/R©**	**Non Siblings (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Estimator**
PubMed-Related articles	13	9	93	0.59	59%	0.12	12%	25.28
PubMed-Author-subject	11	11	903	0.50	50%	0.01	1%	0.8
Citation(Web of Science)	12	10	44	0.55	55%	0.21	21%	43.2
Subject search(e-library)	9	13	2	0.41	41%	0.82	82%	3.3
SCOPUS Author-subject (Author(s))	11	11	39	0.50	50%	0.22	22%	40.74
SCOPUS citation	0	22	0	0	0%	0	0%	0
CINAHL Author-subject	0	22	0	0	0%	0	0%	0
CINAHL citation	0	22	0	0	0%	0	0%	0
WoK-Author-subject	4	18	10	0.18	18%	0.29	29%	35.96
E-lib – Author-subject	2	20	46	0.09	9%	0.04	4%	3.44
PubMed-Economics	2	20	7	0.09	9%	0.22	22%	23.16
PubMed -Costs	2	20	7	0.09	9%	0.22	22%	23.16
PubMed -Qualitative	0	22	28	0.00	0%	0	0%	0
E-lib-Economics	4	18	238	0.18	18%	0.02	2%	1.3
E-lib-Costs	3	19	165	0.14	14%	0.02	2%	1.4
E-lib-Qualitative	3	19	175	0.14	14%	0.02	2%	1.31
WoK-Economics	2	20	6	0.09	9%	0.25	25%	27.03
WoK-Costs	2	20	6	0.09	9%	0.25	25%	27.03
WoK-Qualitative	0	22	31	0	0%	0	0%	0
SCOPUS- Qualitative	0	22	4	0	0%	0	0%	0
SCOPUS-Costs	1	21	8	0.05	5%	0.11	11%	8.57
SCOPUS-Economics	3	19	4	0.14	14%	0.43	43%	36.56
CINAHL- Qualitative	0	22	0	0	0%	0	0%	0
CINAHL-Costs	0	22	0	0	0%	0	0%	0
CINAHL-Economics	0	22	0	0	0%	0	0%	0

Total SIBLINGS R without duplicates	22
Total NON-SIBLINGS R without duplicate	48
Total Relevant R without duplicates	70
Total non-relevant R without duplicates (D)	**1628**

Odds Estimator = A * (D – B)/ B * C

5.4.5.4 Sibling Studies Retrieval by Study Type

Again, each sibling study was categorised into one of the four sibling types, and Table 28 shows that related articles search on PubMed, with odds estimator value of 41.26 was most likely to retrieve the RCT siblings rather than non-RCTs siblings, followed by author-subject search on SCOPUS with odds estimator value of 40.74, and citation search was the third, with odds estimator with value of 27. There were no qualitative siblings retrieved for the IVF seed study.

For the economics siblings, subject search on subject search on SCOPUS with Hedges economics filter odds estimator was the best at retrieving economics siblings rather than non-economics scoring a value of 694.71, WoK author-subject with either Hedges economics or costs filters scored the second best odds estimator with value of 540.67 and third was subject search with either Hedges economics or costs filter on PubMed with odds estimator value of 463.14, as shown in Table 28.

According to Table 28, WoS citation search was the best search to retrieve process evaluation siblings with odds estimator value of 144 followed by SCOPUS author-subject search scoring the second best possibility of retrieving process evaluation siblings with odds estimator value of 27.16. The third best possibility of retrieving process evaluation rather than non process evaluation siblings was scored by related articles search on PubMed with odds estimator value of 4.13.

And finally, as Table 28 shows, there was more than one search strategy which did not retrieve any type sibling studies indicating (as for the CLD seed study) the possibility of either a low number of existing siblings or an interaction between search strategy or database and clinical area.

Table 28: IVF Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
Related Search(PubMed)	10	4	41.26	0	0	0	2	1	33.01	1	4	4.13
Author-Subject(PubMed)	6	8	0.6	0	0	0	0	3	0	2	3	0.54
Citation(Web of Science)	6	8	27	0	0	0	2	1	72	4	1	144
Subject search(e-library)	6	8	3.58	0	0	0	2	1	9.55	1	4	1.19
SCOPUS Author-subject	7	7	40.74	0	0	0	2	1	81.49	2	3	27.16
SCOPUS citation	0	14	0	0	0	0	0	3	0	0	5	0
CINAHL Author-subject	0	14	0	0	0	0	0	3	0	0	5	0
CINAHL citation	0	14	0	0	0	0	0	3	0	0	5	0
WoK-Author-subject	2	12	26.97	0	0	0	2	1	323.6	0	5	0
E-lib - Author-subject	1	13	2.65	0	0	0	1	2	17.2	0	5	0
Economics-Hedges filter	0	14	0	0	0	0	2	1	463.14	0	5	0
Costs-Hedges filter	0	14	0	0	0	0	2	1	463.14	0	5	0
Qualitative-Hedges filter	0	14	0	0	0	0	0	3	0	0	5	0
E-lib-Economics	2	12	0.97	0	0	0	2	1	11.68	0	5	0
E-lib-Costs	2	12	1.48	0	0	0	1	2	4.43	0	5	0
E-lib-Qualitative	1	13	0.64	0	0	0	1	2	4.15	1	4	2.08
WoK-Economics	0	14	0	0	0	0	2	1	540.67	0	5	0
WoK-Costs	0	14	0	0	0	0	2	1	540.67	0	5	0
WoK-Qualitative	0	14	0	0	0	0	0	3	0	0	5	0
SCOPUS-Qualitative	0	14	0	0	0	0	0	3	0	0	5	0
SCOPUS-Costs	0	14	0	0	0	0	1	2	89.94	0	5	0
SCOPUS-Economics	0	14	0	0	0	0	3	0	694.71	0	5	0
CINAHL-Qualitative	0	14	0	0	0	0	0	3	0	0	5	0
CINAHL-Costs	0	14	0	0	0	0	0	3	0	0	5	0
CINAHL-Economics	0	14	0	0	0	0	0	3	0	0	5	0

5.4.6 Summary: Reflection on Search Strategies and Databases Performance

In this section, indirect siblings were the target. Indirect siblings are studies that either appear in the seed study reference list, are cited by or cite the seed study (section 4.8). Standard information retrieval performance measures of recall and precision were used. Odds estimator was the third measure that focuses on ranking search strategies and/or databases based on their efficiency of retrieving direct and indirect siblings rather than non-direct and indirect siblings. Table 29 shows the top three search strategies and/or databases for each seed study independently. According to the results the author-subject type of search performed better than either related articles search or citing search for most of the seed studies, as there was at least one author-subject search for each seed study alongside the related articles search and citing reference search. According to the results, author-subject search strategy (as a type of strategy) can be considered a winner over the related articles search strategy and/or citing reference search strategy.

However the performance of the author-subject search strategy varied among different databases. WoK and SCOPUS databases provided a good recall with a readable retrieval size. Retrieval size of related articles search and author-subject search on PubMed was slightly bigger than Wok and SCOPUS. Although both searches did not entirely outperform either SCOPUS or WoK recall, they provided a good recall for some of the seed studies suggesting that related articles search in particular have a good chance of retrieving both direct and indirect sibling studies. The low recall suggested that combining the top performing strategies might boost up the recall value and provide a more comprehensive search and therefore better retrieval performance (See sections 6.9 & 7.5). Moreover, type of clinical study should be taken into consideration (See section 7.3). PubMed related articles search provided a good recall for some seed studies (in the pilot study the related search was among the best performing strategies as well, see Tables 7 & 29). This indicates that PubMed related article should be considered to complement other search strategies for a comprehensive retrieval. In terms of precision and odds estimator, author-subject search on CINAHL was the best. WoK and SCOPUS provided good chances of retrieving direct and indirect siblings rather than non-siblings, but less likelihood than the CINAHL database.

Table 29: Top Three Search Strategies and Databases Retrieval to Retrieve Direct and Indirect Siblings

IDEATeL		***Tamoxifen***		***BMS***		***CLD***		**IVF**	
Recall									
WoK Author-subject	49%	SCOPUS Author-subject	58%	e-library subject & WoK Author-subject	40%	PubMed related articles	44%	PubMed related articles	59%
SCOPUS Author-subject	42%	WoK Author-subject	54%	PubMed author-subject	30%	subject search(e-library)	35%	WoS citation	55%
E-library Author-subject	39%	PubMed Author-subject	17%	PubMed related articles, WoS citation E-libray Author-subject & WoK Author-subject (qualitative filter)	20%	SCOPUS citation	28%	PubMed Author-subject SCOPUS Author-subject	50%
Precision									
CINAHL Author-subject, SCOPUS -Qualitattive, CINAHL -Qualitative, CINAHL - Costs & CINAHL-Economics	100%	CINAHL Author-subject	54%	WoS citation	67%	SCOPUS author-subject (first author) CINAHL author-subject, WoK-author-subject	100%	subject search(e-library)	82%
WoK Author-subject	79%	WoK-Author – subject	49%	CINAHL Author-subject & CINAHL citation	50%	CINAHL citation, CINAHL-Economics & CINAHL-Qualitative	50%	SCOPUS-Economics	43%
SCOPUS Author-subject & SCOPUS-Costs	75%	SCOPUS Author-subject	34%	SCOPUS citation	33%	subject search(e-library)	47%	WoK Author-subject	29%
Odds Estimator									
CINAHL Author-subject		WoK Author-subject		WoS citation		CINAHL Author-subject		WoS citation	
SCOPUS-Qualitattive		CINAHL Author-subject		CINAHL Author-subject & CINAHL citation		WoK Author-subject & SCOPUS Author-subject		SCOPUS Author-subject	
CINAHL-Qualitative & CINAHL-Economics		SCOPUS Author-subject		SCOPUS-Qualitative		Subject search on e-library		SCOPUS-Qualitative	

Chapter Six
Direct Sibling Retrieval

6.1 Introduction

This chapter presents the results of direct sibling retrieval performance, using the sibling inclusion criteria. Direct siblings in this context refer to the studies that share at least one of the authors of the seed study (Section 4.9).

As in chapter five, the results will presented at the study level, indicating the performance of each search strategy and database for a particular seed study, followed by performance assessment comparison for all the seeds to explore whether there is a particular pattern for retrieving sibling studies regardless of the seed study's clinical topic, or whether the performance is strongly associated with the seed study's main clinical topic or type of intervention.

In this final stage of sophisticated searches were performed on PubMed using MeSH terms in order to explore and compare retrieval performance of these sophisticated search strategies and simple search strategies. Both sophisticated subject search, and sophisticated subject search with specialised Hedges search filters were performed. The best combination of databases and search strategies were also tested to examine how this might improve the recall.

6.2 Diabetes-Telemedicine - IDEATeL Seed Study

IDEATeL seed study search strategies and databases yielded 394 relevant on-topic studies. Those relevant studies were screened using the sibling study inclusion criteria and 70 sibling studies in total were identified from the various search strategies and from different databases. The following section presents the top three retrieval strategies in terms of recall, precision and odds estimator.

6.2.1 Recall

As shown in Table 30, WoK author-subject search identified 46 sibling studies scoring the best recall with value of 71%, next was SCOPUS author-subject search, which identified 39 direct sibling studies (recall 56%), which were both in the relevant set/category and siblings set/category, and the E-library author-subject strategy retrieved 36 siblings scoring the third best recall value of 51%.

6.2.2 Precision

As shown in Table 30, CINAHL Author-subject, CINAHL subject search with Hedges qualitative, economics or costs filters and SCOPUS subject search with Hedges qualitative filter searches scored the best precision value of 100% by not retrieving any non-sibling studies, followed by sophisticated author-subject with Hedges economics filter on PubMed with precision value of 83%, while SCOPUS Author-subject, SCOPUS subject search with Hedges costs filter and CINAHL citation searches scored the third best precision with value of 75%.

6.2.3 Odds Estimator

CINAHL Author-subject search did not retrieve any non-sibling studies causing some difficulties in odds estimator calculation due to division by zero mathematical error, but this problem was solved by substituting 1 to the zero (Section 4.5). CINAHL author-subject search scored the highest odds estimator indicating the highest possibility of retrieving sibling studies rather than non-siblings with value of 475.37. Subject search on SCOPUS with Hedges qualitative filter scored the second best odds estimator with value of 319, followed by WoK author-subject search with odds estimator value of 263.54 as shown in Table 30.

Table 30: The IDEATeL Search Strategies and Databases Retrieval

Diabetes Telemedicine (IDEATeL)								
Search Strategy	Siblings (A)	Siblings N/R©	Non Siblings (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
PubMed-Related articles	20	50	166	0.29	29%	0.11	11%	2.68
PubMed-Author-subject	22	48	135	0.31	31%	0.14	14%	3.88
Citation(Web of Science)	12	58	45	0.17	17%	0.21	21%	5.66
Subject search (e-library)	29	41	267	0.41	41%	0.1	10%	2.68
SCOPUS Author-subject (Author(s))	39	31	13	0.56	56%	0.75	75%	122.32
SCOPUS citation	18	52	46	0.26	26%	0.28	28%	9.26
CINAHL Author-subject	19	51	0	0.27	27%	1	100%	475.37
CINAHL citation	3	67	1	0.04	4%	0.75	75%	57.13
WoK-Author-subject	50	20	9	0.71	71%	0.74	74%	263.54
E-lib – Author-subject	36	34	19	0.51	51%	0.65	65%	70.11
PubMed-Economics	6	64	79	0.09	9%	0.07	7%	1.42
PubMed -Costs	7	63	70	0.1	10%	0.09	9%	1.92
PubMed -Qualitative	14	56	71	0.2	20%	0.16	16%	4.25
E-lib-Economics	18	52	211	0.26	26%	0.08	8%	1.75
E-lib-Costs	14	56	200	0.2	20%	0.07	7%	1.35
E-lib-Qualitative	14	56	254	0.2	20%	0.05	5%	1.01
WoK-Economics	3	67	59	0.04	4%	0.05	5%	0.92
WoK-Costs	6	64	93	0.09	9%	0.06	6%	1.19
WoK-Qualitative	16	54	78	0.23	23%	0.17	17%	4.55
Sophisticated	12	58	97	0.17	17%	0.11	11%	2.52
Sophisticated2 - Subject	33	37	471	0.47	47%	0.07	7%	1.53
Sophisticated2 – Author-Subject	32	38	36	0.46	46%	0.47	47%	29.03
Sophisticated-Author-Subject-Qualitative	13	57	13	0.18	18%	0.5	50%	6.68
Sophisticated-Subject- Economics	5	65	1	0.07	7%	0.83	83%	98.15
Sophisticated2-Author-Subject-Costs	4	66	3	0.06	6%	0.57	57%	25.74
SCOPUS-Qualitative	14	56	0	0.2	20%	1	100%	319
SCOPUS-Costs	3	67	1	0.04	4%	0.75	75%	57.13
SCOPUS-Economics	2	68	2	0.03	3%	0.5	50%	18.75
CINAHL-Qualitative	6	64	0	0.09	9%	1	100%	119.63
CINAHL-Costs	4	66	0	0.06	6%	1	100%	77.33
CINAHL-Economics	6	64	0	0.09	9%	1	100%	119.63

Total SIBLINGS R without duplicates	70
Total NON-SIBLINGS R without duplicate	324
Total Relevant R without duplicates	394
Total non-siblings without duplicates (D)	**1277**

Odds Estimator = A * (D - B)/ B * C

6.2.4 Sibling Studies Retrieval by Study Type

Sibling studies were investigated more thoroughly after identification in order to classify each sibling into one of our four sibling categories (RCT, qualitative, process evaluation and economics). According to Table 31, for RCT siblings, CINAHL author-subject search scored the best odds estimator with value of 708.89, with more likelihood of retrieving RCT rather than non-RCT siblings. WoK author-subject search scored the second best, with odds estimator value of 291.69, followed by SCOPUS author-subject search with odds estimator value of 129.641. SCOPUS subject search with Hedges economics or costs filter and CINAHL subject search with Hedges costs filter were the searches with the least possibility of retrieving RCT siblings rather than non-RCTs with odds estimator of zero.

As Table 31 shows, qualitative siblings were most likely to be retrieved by SCOPUS subject search with Hedges qualitative filter, with odds estimator value of 729.143. CINAHL author-subject and CINAHL subject with Hedges qualitative filter search odds estimator were next, with an odds estimator value of 283.56, followed by WoK author-subject search with odds estimator value of 208.35 as the third best search likely to retrieve qualitative siblings rather than non-qualitative siblings. Subject search with Hedges economics and costs filter on WoK, sophisticated author-subject with Hedges costs filter, SCOPUS subject search with Hedges economics or costs filter and CINAHL subject with Hedges costs filter search did not retrieve any qualitative siblings, with odds estimator value of zero.

The CINAHL subject search with Hedges economics or costs filters scored the highest odds estimator for retrieving economics siblings rather than the other sibling types with value of 3828. Sophisticated author-subject with Hedges economics search filter on PubMed was next with odds estimator of 1276 while SCOPUS subject scored with Hedges economics filter was the third best odds estimator with value of 637.5 indicating the third best likelihood of retrieving economics siblings rather than non-economics. And finally, there were several search strategies which did not retrieve any economics siblings scoring odds estimator value of zero, as illustrated in Table 31.

For process evaluation siblings, CINAHL author-subject search and SCOPUS subject search with Hedges qualitative filter scored the best odds estimator with value of 425.33 indicating the best likelihood of retrieving process evaluation rather than non-process evaluation siblings. WoK author-subject scored the second best odds estimator with value of 162.05, and SCOPUS author-subject search was third with odds estimator value of 97.23. Subject search with Hedges economics and costs filter on WoK and subject search on SCOPUS with Hedges economics filter did not retrieve any qualitative siblings indicating the searches least likely to retrieve process evaluation siblings rather than non-process evaluation siblings with odds estimator value of zero as shown in Table 31.

Table 31: IDEATeL Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCT			Qualitative			Economical Evaluation			Process Evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	6	22	1.83	6	16	2.51	1	3	2.23	7	9	5.21
PubMed-Author-subject	6	22	2.31	10	12	7.05	3	1	25.38	3	13	1.95
Citation(Web of Science)	3	25	3.29	5	17	8.05	2	2	27.38	2	14	3.91
Subject search (e-library)	12	16	2.84	10	12	3.15	2	2	3.78	5	11	1.72
SCOPUS Author-subject (Author(s))	16	12	129.64	13	9	140.44	2	2	97.23	8	8	97.23
SCOPUS citation	8	20	10.7	7	15	12.49	2	2	26.76	1	15	1.78
CINAHL Author-subject	10	18	708.89	4	18	283.56	1	3	425.33	4	12	425.33
CINAHL citation	1	27	47.26	1	21	60.76	0	4	0	1	15	85.07
WoK-Author-subject	20	8	291.69	14	8	208.35	3	1	291.69	9	7	162.05
E-lib - Author-subject	17	11	102.33	11	11	66.21	1	3	22.07	7	9	51.5
Economics-Hedges filter	1	27	0.56	2	20	1.52	1	3	5.06	2	14	2.17
Costs-Hedges filter	1	27	0.64	1	21	0.82	3	1	51.73	2	14	2.46
Qualitative-Hedges filter	4	24	2.83	6	16	6.37	0	4	0	4	12	5.66
E-lib-Economics	5	23	1.1	7	15	2.36	0	4	0	6	10	3.03
E-lib-Costs	4	24	0.9	4	18	1.2	2	2	5.39	4	12	1.8
E-lib-Qualitative	4	24	0.67	5	17	1.19	3	1	12.08	6	10	2.42
WoK-Economics	1	27	0.77	0	22	0	2	2	20.64	0	16	0
WoK-Costs	3	25	1.53	0	22	0	2	2	12.73	1	15	0.85
WoK-Qualitative	5	23	3.34	6	16	5.76	0	4	0	5	11	6.99
Sophisticated Author-subject	5	23	2.65	2	20	1.22	2	2	12.17	3	13	2.81
Sophisticated2 - Subject	13	15	1.48	10	12	1.43	3	1	5.13	7	9	1.33
Sophisticated2 – Author-Subject	13	15	29.88	10	12	28.73	2	2	34.47	7	9	26.81
Sophisticated-Author-Subject-Qualitative	3	25	11.55	4	18	21.39	1	3	32.08	5	11	43.74
Sophisticated-Subject-Economics	1	27	47.26	1	21	60.76	2	2	1276	1	15	85.07
Sophisticated2-Author-Subject-Costs	1	27	15.73	0	22	0	2	2	424.67	1	15	28.31
SCOPUS-Qualitative	2	26	98.15	8	14	729.14	0	4	0	4	12	425.33
SCOPUS-Costs	0	28	0	0	22	0	2	2	1276	1	15	85.07
SCOPUS-Economics	0	28	0	0	22	0	2	2	637.5	0	16	0
CINAHL-Qualitative	1	27	47.26	4	18	283.56	0	4	0	1	15	85.07
CINAHL-Costs	0	28	0	0	22	0	3	1	3828	1	15	85.07
CINAHL-Economics	1	27	47.26	1	21	60.76	3	1	3828	1	15	85.07

6.2.5 IDEATeL Missing Siblings

Based on the "gold standard" list provided for the IDEATeL seed study, by the IDEATeL author, there were some studies that did not appear in the pooled retrieval list. The studies were book chapters, conference papers or abstracts. PubMed and WoK were checked to see if the databases contained any of these studies although apparently they were missed by the search. In fact, the "gold standard" list provided by the IDEATeL seed study contact author only had 68 studies, but the number of siblings identified from the retrieval was 70 (there was some studies which share authors - apparently were not part of the IDEATeL project website publication list- but based on the topic were judged siblings as well). It is fair to assume that the list provided might not be fully up to date. This might also lead us to think that some authors carried out individual studies based on the concepts of the IDEATeL project where the main principle and aims were the same and therefore it remained under the main theme of the IDEATeL project. The following flowchart (Figure 7) demonstrates how the retrieved siblings and "missing" siblings were allocated.

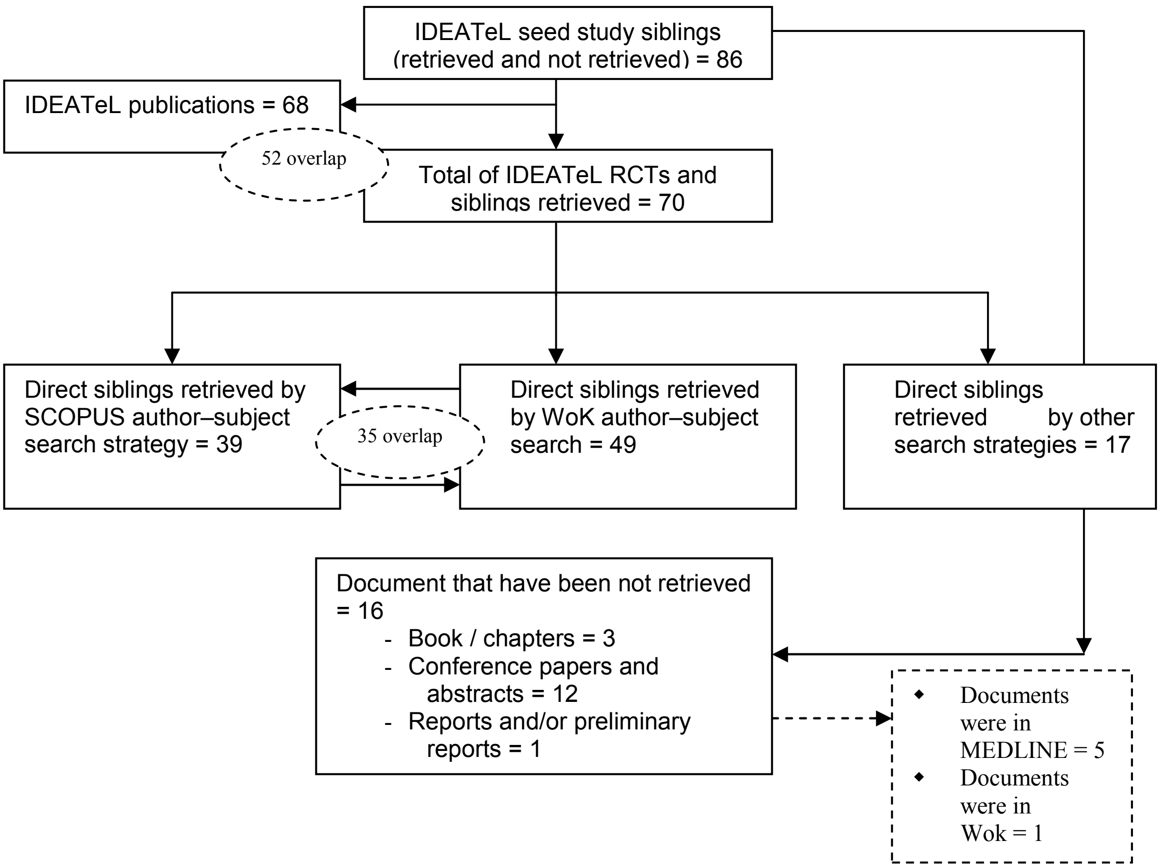


Figure 7: IDEATeL Siblings Retrieval

As indicated, six items were apparently missed in the retrieval although they could have been retrieved. Further inspection of the abstracts in MEDLINE confirmed that the six items did not fit the categories of RCT, qualitative or economic research; they appeared to discuss research techniques or technical know-how background to the research (rather than process evaluation). Closer inspection of items not retrieved showed that most of the lost studies (n=16) were conference papers and abstracts (n=12) which mean they are not completely lost and might appear later as journal articles; books or book chapters (n=3) and preliminary reports (n=1). In the lists provided there were in fact several pairs of items, a conference paper and a journal article, that had the same or very similar titles which suggested that the journal article was a revised and presumably improved version of the conference paper.

Figure 7 also suggests that more than one search strategy is necessary to boost recall. This is examined later in Section 6.9.

6.2.6 IDEATeL Siblings Overall Retrieval Process

Figure 8 demonstrates the overall retrieval process of the IDEATeL seed study. It shows that the overall retrieval was 1347 studies retrieved by all search strategies from different databases, the initial screening process extracted 394 candidate studies as possible siblings. In the final screening process, my inclusion criteria narrowed this to 70 siblings.

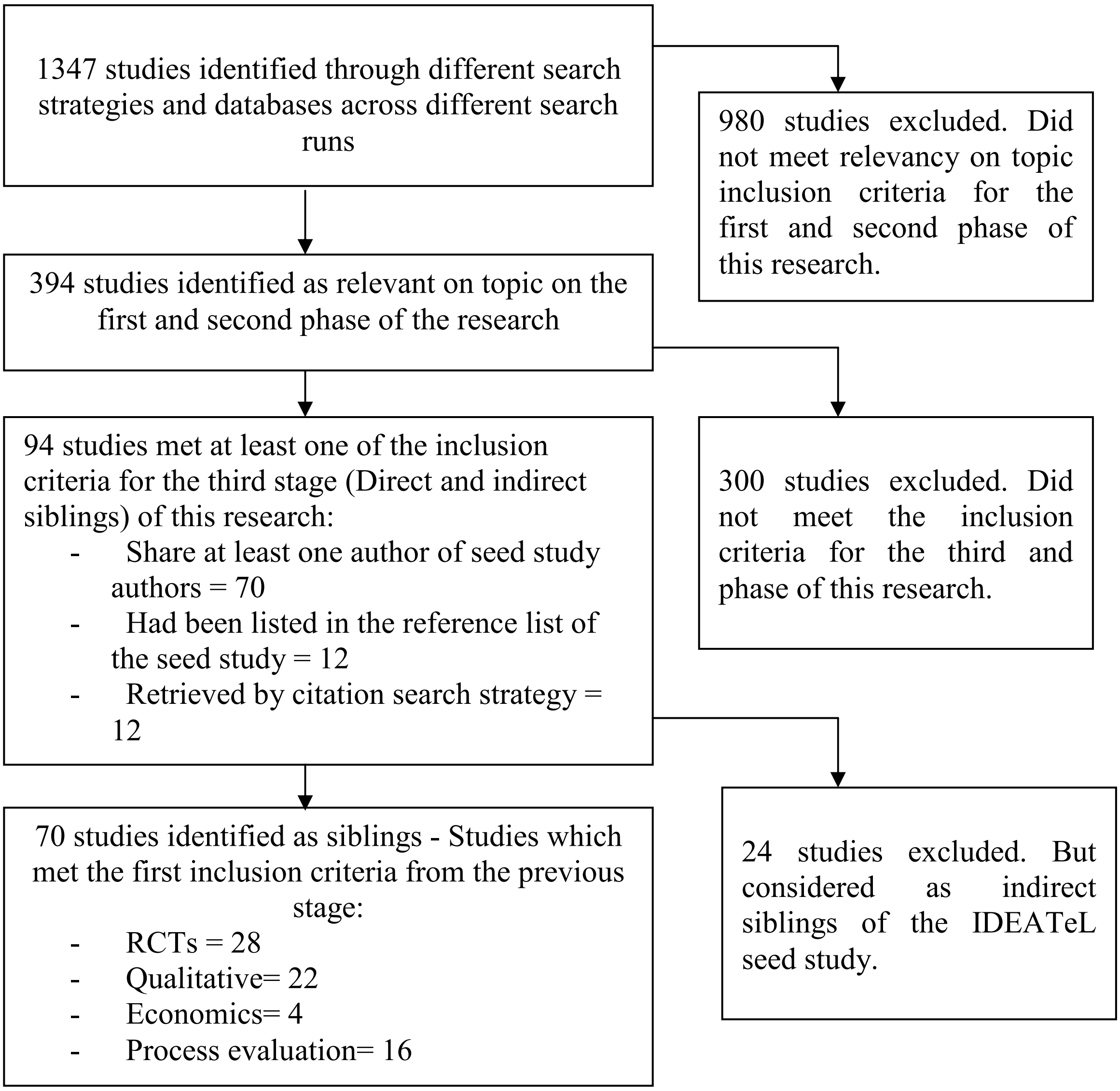


Figure 8: Overall Siblings Retrieval & Identification for IDEATeL Seed Study

6.2.7 Sophisticated Search vs Simple Search

Table 32 shows a comparison between simple search strategy and sophisticated search strategy (for PubMed only). The overall performance shows that the sophisticated search strategy provided a better recall than the simple search strategy, the precision values indicated that the sophisticated search strategy provided a better performance than the simple search. However, Table 28 shows that simple search strategy on different databases provided a better recall and precision, such as WoK (66% & 74%) or SCOPUS (56% & 75%).

Table 32: Comparison of Simple and Sophisticated Search Strategies for Siblings Retrieval on PubMed

Search Strategy – PubMed	Siblings	Total	Recall	Precision	Odds Estimator
Simple Author-Subject	22	157	31%	14%	4.49
Simple Author-Subject - Economics	6	85	9%	7%	1.59
Simple Author-Subject – Costs	7	77	10%	9%	2.12
Simple Author-Subject - Qualitative	14	85	20%	16%	4.69
Sophisticated Author-Subject	32	68	46%	47%	30.71
Sophisticated Author-Subject- Economics	5	6	7%	83%	103.08
Sophisticated Author-Subject-Costs	4	7	6%	57%	27.09
Sophisticated Author-Subject- Qualitative	13	26	19%	5%	6.68

6.3 Breast Cancer (Tamoxifen) Seed Study

The search strategies and databases retrieved 386 relevant on-topic studies of the Tamoxifen. 111 studies from the relevant set met the inclusion criteria, therefore were marked/categorised as siblings of the Tamoxifen seed study. The performance of each search strategy and databases was assessed and presented in the following section.

6.3.1 Recall

As Table 33 shows, SCOPUS author-subject search retrieved the highest number of siblings with value of 74 siblings out of 78 relevant studies, scoring the highest recall value of 67%, followed by WoK author-subject retrieving 71 siblings with recall value of 64% as the second best recall value. The third best recall was scored by sophisticated author-subject on PubMed retrieving 38 siblings and scoring recall value of 34%.

6.3.2 Precision

Sophisticated author – subject search and sophisticated author-subject search with Hedges qualitative search filter on PubMed scored the highest precision with value of 78%, followed by author-subject search on CINAHL with precision value of 51% as the second best precision. WoK author subject scored the third best precision with value of 48%as shown in Table 33.

6.3.3 Odds Estimator

According to Table 33 odds estimator calculations, sophisticated author-subject search on PubMed scored the highest odds estimator with value of 61.97, indicating the highest likelihood of retrieving siblings rather than non-siblings, followed by WoK, with odds estimator value of 38.7, and third was SCOPUS author-subject search with odds estimator value of 20.36. There were several search strategies and databases which did not retrieve scoring odds estimator value of zero; e.g. CINAHL citation.

Table 33: The Tamoxifen Search Strategies and Databases Retrieval

Search Strategy	Siblings (A)	Siblings N/R©	Non Siblings (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
Breast cancer (Tamoxifen)								
PubMed-Related articles	3	108	197	0.03	3%	0.02	2%	0.22
PubMed-Author-subject	23	88	428	0.21	21%	0.05	5%	0.8
Citation(Web of Science)	1	110	52	0.01	1%	0.02	2%	0.29
Subject search (e-library)	4	107	284	0.04	4%	0.01	1%	0.19
SCOPUS Author-subject (Author(s))	74	37	155	0.67	67%	0.32	32%	20.36
SCOPUS citation	1	110	58	0.01	1%	0.02	2%	0.26
CINAHL Author subject	21	90	20	0.19	19%	0.51	51%	19.99
CINAHL citation	0	111	4	0	0%	0	0%	0
WoK-Author - subject	71	40	76	0.64	64%	0.48	48%	38.7
E-lib - Author subject	3	108	40	0.03	3%	0.07	7%	1.18
Economics-Hedges filter	0	111	73	0	0%	0	0%	0
Costs-Hedges filter	0	111	66	0	0%	0	0%	0
Qualitative-Hedges filter	1	110	133	0.01	1%	0.01	1%	0.11
E-lib-Economics	0	111	213	0	0%	0	0%	0
E-lib-Costs	0	111	158	0	0%	0	0%	0
E-lib-Qualitative	0	111	215	0	0%	0	0%	0
WoK-Economics	0	111	139	0	0%	0	0%	0
WoK-Costs	0	111	99	0	0%	0	0%	0
WoK-Qualitative	0	111	206	0	0%	0	0%	0
Sophisticated-Author-Subject	38	96	11	0.34	34%	0.78	78%	61.97
Sophisticated-Author-Subject-Qualitative	25	109	7	0.23	23%	0.78	78%	12.42
Sophisticated-Subject- Economics	0	134	90	0	0%	0	0%	0
Sophisticated-Author-Subject-Costs	0	134	67	0	0%	0	0%	0
SCOPUS-Qualitative	0	111	134	0	0%	0	0%	0
SCOPUS-Costs	0	111	89	0	0%	0	0%	0
SCOPUS-Economics	0	111	40	0	0%	0	0%	0
CINAHL-Qualitative	0	111	22	0	0%	0	0%	0
CINAHL-Costs	0	111	11	0	0%	0	0%	0
CINAHL-Economics	0	111	30	0	0%	0	0%	0

Total SIBLINGS R without duplicates	111
Total NON-SIBLINGS R without duplicate	275
Total Relevant R without duplicates	386
Total non-relevant R without duplicates (D)	**1733**
Odds Estimator = A * (D - B)/ B * C	

6.3.4 Sibling Studies Retrieval by Study Type

As mentioned earlier, sibling studies were further investigated after identification in order to classify each sibling into one of the four sibling' categories (RCT, qualitative, process evaluation and economics). Therefore, as shown in Table 34, for RCT siblings, sophisticated author-subject search on PubMed scored the highest odds estimator of 87.76

followed by WoK author-subject odds estimator value of 76.31, third was sophisticated author-subject search with Hedges qualitative filter on PubMed, with odds estimator value of 70.45. As indicated in Table 34, there was many search strategies with zero odds estimator values.

As illustrated in Table 34, qualitative siblings were best retrieved by sophisticated author-subject search on PubMed, as it scored an odds estimator with value of 38.82, showing the best likelihood of retrieving qualitative siblings rather than non-qualitative siblings. Next were sophisticated author-subject search PubMed with Hedges qualitative filter, odds estimator value of 33.623, followed by Author-subject on E-Library with odds estimator value of 5.77. There was many search strategies with zero odds estimators, as shown in Table 31.

There was no economics direct sibling studies retrieved by any search strategy or database for the Tamoxifen seed studies.

For process evaluation siblings, the sophisticated author-subject search with Hedges qualitative filter on PubMed has the best odds estimator with value of 184.93, sophisticated author-subject search scored the second best odds estimator with value of 155.27, and WoK author-subject scored third, with odds estimator value of 54.51, as shown in Table 34.

There was more than one search strategy that did not retrieve any economics or process evaluation; hence there were several odds estimator values of zero, as shown in Table 34.

Table 34: Tamoxifen Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	1	71	0.11	1	24	0.33	0	0	0	1	13	0.6
PubMed-Author-subject	16	56	0.87	3	22	0.42	0	0	0	4	10	1.22
Citation(Web of Science)	0	72	0	1	24	1.35	0	0	0	0	14	0
Subject search (e-library)	4	68	0.3	0	25	0	0	0	0	0	14	0
SCOPUS Author-subject (Author(s))	56	16	35.63	7	18	3.96	0	0	0	11	3	37.33
SCOPUS citation	1	71	0.411	0	25	0	0	0	0	0	14	0
CINAHL Author-subject	17	55	26.47	0	25	0	0	0	0	4	10	34.26
CINAHL citation	0	72	0	0	25	0	0	0	0	0	14	0
WoK-Author-subject	56	16	76.31	5	20	5.45	0	0	0	10	4	54.51
E-lib - Author-subject	0	72	0	3	22	5.77	0	0	0	0	14	0
Economics-Hedges filter	0	72	0	0	25	0	0	0	0	0	14	0
Costs-Hedges filter	0	72	0	0	25	0	0	0	0	0	14	0
Qualitative-Hedges filter	1	71	0.17	0	25	0	0	0	0	0	14	0
E-lib-Economics	0	72	0	0	25	0	0	0	0	0	14	0
E-lib-Costs	0	72	0	0	25	0	0	0	0	0	14	0
E-lib-Qualitative	0	72	0	0	25	0	0	0	0	0	14	0
WoK-Economics	0	72	0	0	25	0	0	0	0	0	14	0
WoK-Costs	0	72	0	0	25	0	0	0	0	0	14	0
WoK-Qualitative	0	72	0	0	25	0	0	0	0	0	14	0
Sophisticated-Author-Subject	26	46	87.76	5	20	38.82	0	0	0	7	7	155.27
Sophisticated-Author-Subject-Qualitative	16	56	70.45	3	22	33.62	0	0	0	6	8	184.93
Sophisticated-Subject-Economics	0	72	0	0	25	0	0	0	0	0	14	0
Sophisticated-Author-Subject-Costs	0	72	0	0	25	0	0	0	0	0	14	0
SCOPUS-Qualitative	0	72	0	0	25	0	0	0	0	0	14	0
SCOPUS-Costs	0	72	0	0	25	0	0	0	0	0	14	0
SCOPUS-Economics	0	72	0	0	25	0	0	0	0	0	14	0
CINAHL-Qualitative	0	72	0	0	25	0	0	0	0	0	14	0
CINAHL-Costs	0	72	0	0	25	0	0	0	0	0	14	0
CINAHL-Economics	0	72	0	0	25	0	0	0	0	0	14	0

6.4.5 Breast Cancer (Tamoxifen) Overall Retrieval Process

Figure 9 demonstrates the overall retrieval process of Tamoxifen seed study. It shows that the overall retrieval was 1844 studies retrieved by all search strategies from different databases, the initial screening process identified 383 candidate studies to be siblings. In the final screening process and after applying the inclusion criteria, 111 sibling studies were identified.

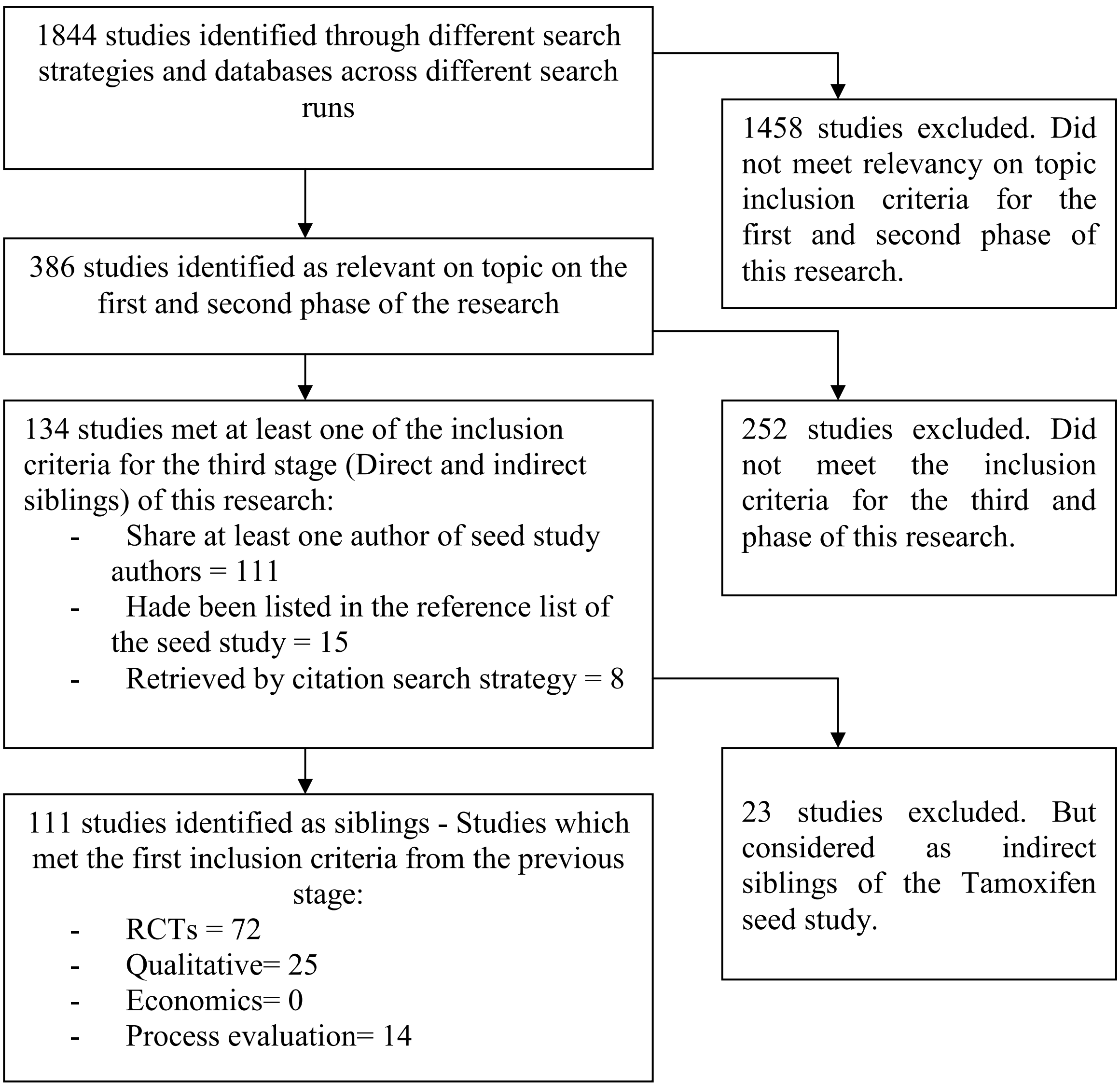


Figure 9: Overall Siblings Retrieval & Identification Process for Tamoxifen Seed Study

6.4 Breast Cancer (BMS) Seed Study

Search strategies and databases yielded 71 of relevant on-topic studies for breast cancer (BMS) seed study. There were 8 studies that met the inclusion criteria and were included as sibling studies. Recall, precision and odds estimator were calculated to investigate retrieval performance of each search strategy and database.

6.4.1 Recall

The recall values which are presented in Table 35 showing that the recall ratio is very low for this seed study for siblings retrieval. The highest recall value was 38% which was scored by WoK author-subject search and Sophisticated-Author-Subject, followed by a large number of strategies all retrieving 2 siblings, scoring a recall value of 25%.

6.4.2 Precision

According to Table 35, CINAHL author-subject scored the best precision value of 50%. The second best precision was scored by SCOPUS subject search with Hedges qualitative filter with precision value of 33%. And finally, author-subject search with Hedges qualitative search filter on WoK scored the third best precision with value of 25%.

6.4.3 Odds Estimator

The best likelihood of retrieving the siblings rather than non-siblings was achieved by CINAHL author-subject with odds estimator value of 86.43. SCOPUS subject search with Hedges qualitative filter was second with odds estimator value of 50.17, while author-subject search with Hedges qualitative filter on WoK was third with odds estimator value of 33.33 as shown in Table 35.

Table 35: The BMS Search Strategies and Databases Retrieval

Breast Cancer (BMS)								
Search Strategy	**Siblings (A)**	**Siblings N/R©**	**Non Siblings (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Estimator**
PubMed-Related articles	1	7	107	0.13	13%	0.01	1%	0.67
PubMed-Author-subject	2	6	65	0.25	25%	0.03	3%	2.77
Citation(Web of Science)	0	8	3	0	0%	0	0%	0
Subject search (e-library)	2	6	22	0.25	25%	0.08	8%	8.85
SCOPUS Author-subject (Group)	1	2	25	0.13	13%	0.04	4%	11.62
SCOPUS Author-subject (Author(s))	0	8	0	0	0%	0	0%	0
SCOPUS citation	0	8	3	0	0%	0	0%	0
CINAHL Author subject	1	7	1	0.13	13%	0.5	50%	86.43
CINAHL citation	0	8	2	0	0%	0	0%	0
WoK-Author - subject	3	5	18	0.38	38%	0.14	14%	19.6
E-lib - Author subject	2	6	81	0.25	25%	0.02	2%	2.16
Economics-Hedges filter	0	8	0	0	0%	0	0%	0
Costs-Hedges filter	0	8	0	0	0%	0	0%	0
Qualitative-Hedges filter	1	7	6	0.13	13%	0.14	14%	14.29
E-lib-Economics	0	8	0	0	0%	0	0%	0
E-lib-Costs	0	8	0	0	0%	0	0%	0
E-lib-Qualitative	0	8	0	0	0%	0	0%	0
WoK-Economics	0	8	0	0	0%	0	0%	0
WoK-Costs	0	8	0	0	0%	0	0%	0
WoK-Qualitative	2	6	6	0.25	25%	0.25	25%	33.33
Sophisticated-Author-Subject	3	5	68	0.38	38%	0.04	4%	4.75
Sophisticated-Author-Subject-Qualitative	2	6	9	0.25	25%	0.18	18%	2.3
Sophisticated-Subject-Economics	0	8	96	0	0%	0	0%	0
Sophisticated-Subject-Costs	0	8	66	0	0%	0	0%	0
SCOPUS-Qualitative	2	6	4	0.25	25%	0.33	33%	50.17
SCOPUS-Costs	0	8	0	0	0%	0	05	0
SCOPUS-Economics	0	8	0	0	0%	0	0%	0
CINAHL-Qualitative	2	6	12	0.25	25%	0.14	14%	16.5
CINAHL-Costs	0	8	3	0	0%	0	0%	0
CINAHL-Economics	1	7	11	0.13	13%	0.08	8%	7.73

Total SIBLINGS R without duplicates	8
Total NON-SIBLINGS R without duplicate	63
Total Relevant R without duplicates	71
Total non-relevant R without duplicates (D)	**606**

Odds Estimator = A * (D – B)/ B * C

6.4.4 Sibling Studies Retrieval by Study Type

After constructing the siblings set, each sibling was categorised into one of the four sibling types, and the odds estimator for each sibling category were calculated. Table 36 shows that for the RCT siblings, odds estimator values indicated that CINAHL author-subject search scored the best odds estimator with odds estimator value of 302.5, SCOPUS subject search with Hedges qualitative filter odds estimator was next with 75.25, followed by sophisticated author-subject search with Hedges qualitative filter, with odds estimator value of 33.17. One would not expect the RCTs to be retrieved preferentially with a qualitative filter!

As illustrated in Table 36, qualitative siblings behaved more as might be expected, with the Hedges qualitative filter on WoK scoring the highest odds estimator value of 200, followed by. SCOPUS subject search with Hedges qualitative filter, with odds estimator value of 75.25 and Author-subject search on WoK search third with odds estimator value of 65.33.

There were no economics siblings for the breast cancer with Body-Mind-Spirit therapy seed study retrieved by any search strategy or database.

For process evaluation siblings there were only two siblings were retrieved. Subject search with qualitative Hedges filter on WoK retrieved one of the siblings, scoring the best odds estimator with value of 100. Author-subject on PubMed retrieved the second one with odds estimator with value of 8.32 suggesting the second likelihood of retrieving process evaluation siblings (Table 36).

Table 36: BMS Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	1	2	2.33	0	3	0	0	0	0	0	2	0
PubMed-Author-subject	1	2	4.16	0	3	0	0	0	0	1	1	8.32
Citation(Web of Science)	0	3	0	0	3	0	0	0	0	0	2	0
Subject search (e-library)	1	2	13.27	1	2	13.27	0	0	0	0	2	0
SCOPUS Author-subject (Group)	1	2	11.62	0	3	0	0	0	0	0	2	0
SCOPUS Author-subject (Author(s))	0	3	0	0	3	0	0	0	0	0	2	0
SCOPUS citation	0	3	0	0	3	0	0	0	0	0	2	0
CINAHL Author-subject	1	2	302.5	0	3	0	0	0	0	0	2	0
CINAHL citation	0	3	0	0	3	0	0	0	0	0	2	0
WoK-Author-subject	1	2	16.33	2	1	65.33	0	0	0	0	2	0
E-lib - Author-subject	1	2	3.24	1	2	3.24	0	0	0	0	2	0
Economics-Hedges filter	0	3	0	0	3	0	0	0	0	0	2	0
Costs-Hedges filter	0	3	0	0	3	0	0	0	0	0	2	0
Qualitative-Hedges filter	0	3	0	1	2	50	0	0	0	0	2	0
E-lib-Economics	0	3	0	0	3	0	0	0	0	0	2	0
E-lib-Costs	0	3	0	0	3	0	0	0	0	0	2	0
E-lib-Qualitative	0	3	0	0	3	0	0	0	0	0	2	0
WoK-Economics	0	3	0	0	3	0	0	0	0	0	2	0
WoK-Costs	0	3	0	0	3	0	0	0	0	0	2	0
WoK-Qualitative	0	3	0	2	1	200	0	0	0	1	1	100
Sophisticated-Author-Subject	2	1	15.82	1	2	3.96	0	0	0	0	2	0
Sophisticated-Author-Subject-Qualitative	1	2	33.17	1	2	33.17	0	0	0	0	2	0
Sophisticated-Subject-Economics	0	3	0	0	3	0	0	0	0	0	2	0
Sophisticated-Subject-Costs	0	3	0	0	3	0	0	0	0	0	2	0
SCOPUS-Qualitative	1	2	75.25	1	2	75.25	0	0	0	0	2	0
SCOPUS-Costs	0	3	0	0	3	0	0	0	0	0	2	0
SCOPUS-Economics	0	3	0	0	3	0	0	0	0	0	2	0
CINAHL-Qualitative	1	2	24.75	1	2	24.75	0	0	0	0	2	0
CINAHL-Costs	0	3	0	0	3	0	0	0	0	0	2	0
CINAHL-Economics	0	3	0	1	2	27.05	0	0	0	0	2	0

6.4.5 Breast Cancer (BMS) Siblings Overall Retrieval Process

As illustrated in Figure 10 the overall retrieval for BMS seed study was 614 studies retrieved by all search strategies from different databases, applying the initial screening process produced 71 candidate sibling studies. In the last and final screening process, the inclusion criteria identified 8 studies as direct siblings.

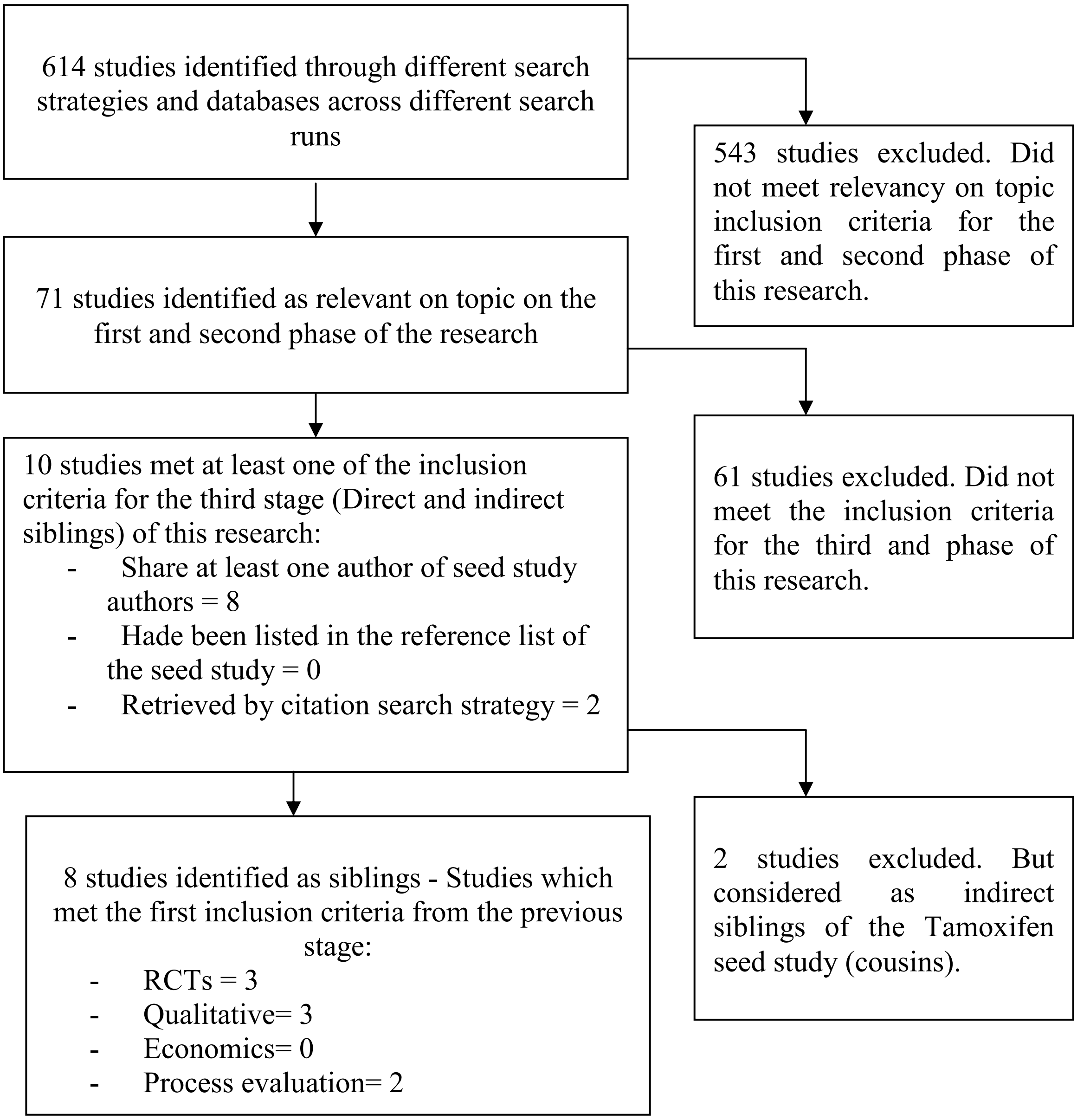


Figure 10: Overall Siblings Retrieval & Identification Process for BMS Seed Study

6.4.6 Checking Breast Cancer (BMS) studies

One author provided a list of studies on body-mind-spirit therapy. The list provided contained the eight studies I identified as siblings. In addition, 12 additional studies were listed (total list of 20 studies on BMS). The additional 12 studies were examined, and as four were published in year 2012 these were immediately excluded. The remaining eight missed studies were conference papers (n=2), book chapters (n=2) or addressing general cancer BMS (n=5). Some of them were in WoK (n=4) and/or SCOPUS (n=3) and some of them did not have abstract and accordingly it was difficult to determine the sibling relationship (See Appendices Six & Seven). However, six of the eight the missed studies addressed general cancer BMS and therefore were not deemed sibling studies according to the criteria set. The remaining two did not have abstract and accordingly the decision about their relationship with the seed study cannot be determined for certain. The criteria for inclusion as a sibling required that the title words made it obvious that the study was either a qualitative, process evaluation or economic sibling (see screening criteria in Section 4.5).

6.5 Chronic Lung Disease (Dexamethasone) Seed Study

The total number of studies that are relevant to CLD seed study was 153, after applying sibling studies identification and inclusion criteria 13 studies were identified and marked as siblings of CLD seed study.

The CLD seed study was done by 'The Vermont Oxford Network Steroid Study Group', and at the end of the article there was a list of the article committee. When searching for the CLD it was noticed that this seed was indexed sometimes under the group name only, i.e. in PubMed and CINAHL, and sometimes under group or author's names such as 'SOLL. R.F' in the author field providing them as two separate records, i.e. for WoK, while with SCOPUS it was indexed under 'SOLL. R.F' as the author and the only author with no reference to the group name.

6.5.1 Recall

The recall ratios for CLD siblings retrieval was slightly low as the highest recall value was 46% and was scored by subject search on e-library followed by PubMed related search and WoK author-subject scoring a recall value of 39%. And finally, SCOPUS author-subject (Group), SCOPUS author-subject (first author), SCOPUS citation, CINAHL author-subject & E-library subject (qualitative filter) scored the third best recall value of 31%, as illustrated in Table 37.

6.5.2 Precision

SCOPUS author-subject search scored the best precision value of 100%. Author-subject search on WoK scored the second best precision with value of 83% while CINAHL author-subject with a precision value of 27% scored the third best precision for CLD siblings retrieval, as shown in Table 37.

6.5.3 Odds Estimator

According to Table 37, odds estimator calculations revealed that WoK author-subject search is the most likely search to retrieve sibling studies rather than non-siblings by scoring odds estimator of value 623.13, second was CINAHL author-subject and SCOPUS author-subject, each with odds estimator value of 443.11. Finally, subject on e-library scored the third odds estimator value of 17.34.

Table 37: The CLD Search Strategies and Databases Retrieval

Search Strategy	Siblings (A)	Siblings N/R©	Non Siblings (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Estimator
Chronic Lung Disease (CLD)								
PubMed-Related articles	5	8	722	0.39	39%	0.01	1%	0.24
PubMed-Author-subject	3	10	36	0.23	23%	0.08	8%	8.02
Citation(Web of Science)	0	13	35	0	0%	0	0%	0
Subject search (e-library)	6	7	47	0.46	46%	0.11	11%	17.34
SCOPUS Author-subject (Group)	4	9	34	0.31	31%	0.11	11%	12.6
SCOPUS Author-subject (Author(s))	4	9	0	0.31	31%	1	100%	443.11
SCOPUS citation	4	9	40	0.31	31%	0.09	9%	10.64
CINAHL Author subject	4	9	1	0.31	31%	0.8	80%	443.11
CINAHL citation	0	13	2	0	0%	0	0%	0
WoK-Author - subject	5	8	1	0.39	39%	0.83	83%	623.13
E-lib - Author subject	2	11	36	0.15	15%	0.05	5%	4.86
Economics-Hedges filter	0	13	7	0	0%	0	0%	0
Costs-Hedges filter	0	13	5	0	0%	0	0%	0
Qualitative-Hedges filter	1	12	10	0.08	8%	0.09	9%	8.23
E-lib-Economics	0	13	57	0	0%	0	0%	0
E-lib-Costs	1	12	19	0.08	8%	0.05	5%	4.29
E-lib-Qualitative	2	11	88	0.15	157%	0.02	2%	1.88
WoK-Economics	0	13	1	0	0%	0	0%	0
WoK-Costs	0	13	4	0	0%	0	0%	0
WoK-Qualitative	1	12	17	0.08	8%	0.06	6%	4.81
Sophisticated-Author-Subject	1	12	14	0.08	8%	0.07	7%	5.86
Sophisticated-Author-Subject-Qualitative	0	13	8	0	0%	0	0%	0
Sophisticated-Subject-Economics	0	13	1	0	0%	0	0%	0
Sophisticated-Subject-Costs	0	13	1	0	0%	0	0%	0
SCOPUS-Qualitative	2	11	34	0.15	15%	0.06	6%	0.65
SCOPUS-Costs	2	11	23	0.15	15%	0.08	8%	1
SCOPUS-Economics	1	12	12	0.08	8%	0.08	8%	1
CINAHL-Qualitative	0	13	4	0	0%	0	0%	0
CINAHL-Costs	0	13	1	0	0%	0	0%	0
CINAHL-Economics	1	12	3	0.08	8%	0.25	25%	7.285

Total SIBLINGS R without duplicates	13
Total NON-SIBLINGS R without duplicate	985
Total Relevant R without duplicates	154
Total Non-Relevant R without duplicate (D)	**998**

Odds Estimator = A * (D – B)/ B * C

6.5.4 Sibling Studies Retrieval by Study Type

Table 38 shows that WoK author-subject was the search strategy most likely to retrieve the RCTs siblings rather than non-RCTs siblings as indicated by its odds estimator value of 1246.25, followed by SCOPUS author-subject (Author(s)) and CINAHL author-subject with odds estimator value of 797.6.The third best odds estimator value of 41.36 was scored by e subject search with economics filter on CINAHL.

There was no qualitative study retrieved for the CLD seed study. There was only one economics sibling for the CLD seed study which was retrieved only by author-subject on PubMed, scoring the only and obviously best odds estimator value of 26.72, while other search strategies/databases odds estimator was zero as shown in Table 38. Finally, according to Table 36, author-subject search on PubMed scored the best odds estimator value of 53.44 suggesting by that the highest likelihood of retrieving process evaluation rather than non-process evaluation siblings. PubMed related articles scored the next with odds estimator value of 0.19. The above search strategies were the only searches to retrieve process evaluation siblings.

Table 38: CLD Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	4	5	0.31	0	0	0	0	1	0	1	2	0.19
PubMed-Author-subject	0	9	0	0	0	0	1	0	26.72	2	1	53.44
Citation(Web of Science)	0	9	0	0	0	0	0	1	0	0	3	0
Subject search (e-library)	6	3	40.47	0	0	0	0	1	0	0	3	0
SCOPUS Author-subject (Group)	4	5	22.68	0	0	0	0	1	0	0	3	0
SCOPUS Author-subject (Author(s))	4	5	797.6	0	0	0	0	1	0	0	3	0
SCOPUS citation	4	5	19.16	0	0	0	0	1	0	0	3	0
CINAHL Author-subject	4	5	797.6	0	0	0	0	1	0	0	3	0
CINAHL citation	0	9	0	0	0	0	0	1	0	0	3	0
WoK-Author-subject	5	4	1246.25	0	0	0	0	1	0	0	3	0
E-lib - Author-subject	2	7	7.64	0	0	0	0	1	0	0	3	0
Economics-Hedges filter	0	9	0	0	0	0	0	1	0	0	3	0
Costs-Hedges filter	0	9	0	0	0	0	0	1	0	0	3	0
Qualitative-Hedges filter	1	8	12..35	0	0	0	0	1	0	0	3	0
E-lib-Economics	0	9	0	0	0	0	0	1	0	0	3	0
E-lib-Costs	1	8	6.44	0	0	0	0	1	0	0	3	0
E-lib-Qualitative	2	7	2.96	0	0	0	0	1	0	0	3	0
WoK-Economics	0	9	0	0	0	0	0	1	0	0	3	0
WoK-Costs	0	9	0	0	0	0	0	1	0	0	3	0
WoK-Qualitative	1	8	7.21	0	0	0	0	1	0	0	3	0
Sophisticated-Author-Subject	1	8	8.79	0	0	0	0	1	0	0	3	0
Sophisticated-Author-Subject-Qualitative	0	9	0	0	0	0	0	1	0	0	3	0
Sophisticated-Subject-Economics	0	9	0	0	0	0	0	1	0	0	3	0
Sophisticated-Subject-Costs	0	9	0	0	0	0	0	1	0	0	3	0
SCOPUS-Qualitative	1	8	3.54	0	0	0	0	1	0	0	3	0
SCOPUS-Costs	2	7	12.11	0	0	0	0	1	0	0	3	0
SCOPUS-Economics	1	8	10.27	0	0	0	0	1	0	0	3	0
CINAHL-Qualitative	0	9	0	0	0	0	0	1	0	0	3	0
CINAHL-Costs	0	9	0	0	0	0	0	1	0	0	3	0
CINAHL-Economics	1	8	41.46	0	0	0	0	1	0	0	3	0

6.5.5 CLD Siblings List

After attempting to cantact the authors of the CLD seed study one author responded and tried to verify the gold standard I created based on relative recall. According to him the relationship with the RCT (seed study) is not clear. The author considered some of the studies as not related to the seed study. However, after inspecting their abstracts again they are considerd to be on the main topic of the seed study and therefore it is fair to include them in the siblings list, according to the criteria derived on the basis of the literature review and with reference to the IDETeL list

Furthermore, the CLD was a large collaborative study and therefore was indexed either under author names or the collaboration name. Accordingly, in the siblings list there are three different representations of the seed study itself. I dealt with these as independent instances in order to emphasise and reflect the difference in the indexing process in different databases. Moreover, as a large collaborative study, many different researchers are on different learning curves, and the learning that one researcher brings to the project may not be appreciated or valued by someone of a different background. It seems that deciding what is a sibling study truly is a very subjective judgement, particularly for large collaborative studies. Again, this draws to the mind the problem of relevancy and relevancy judgement (See Section 3.6.2).

The author did not add any references to the provided list of siblings. This could emphasise that the relationship is not an easy one to determine, or confirm that the number of sibling studies for the CLD seed study is indeed low.

6.5.6 CLD Siblings Overall Retrieval Process

Search strategies and databases retrieved 1011 studies for the CLD seed study, from these 153 studies were extracted and tagged as candidate studies to be siblings and finally by applying the sibling inclusion criteria 13 sibling studies were identified as direct siblings as illustrated in Figure 11. It should be mentioned that there was different citations for the same study in different databases. Each record was handeled as an independent citation in this research to reflect and emphasis representation and indexing difference between databases especially with a callaprative research.

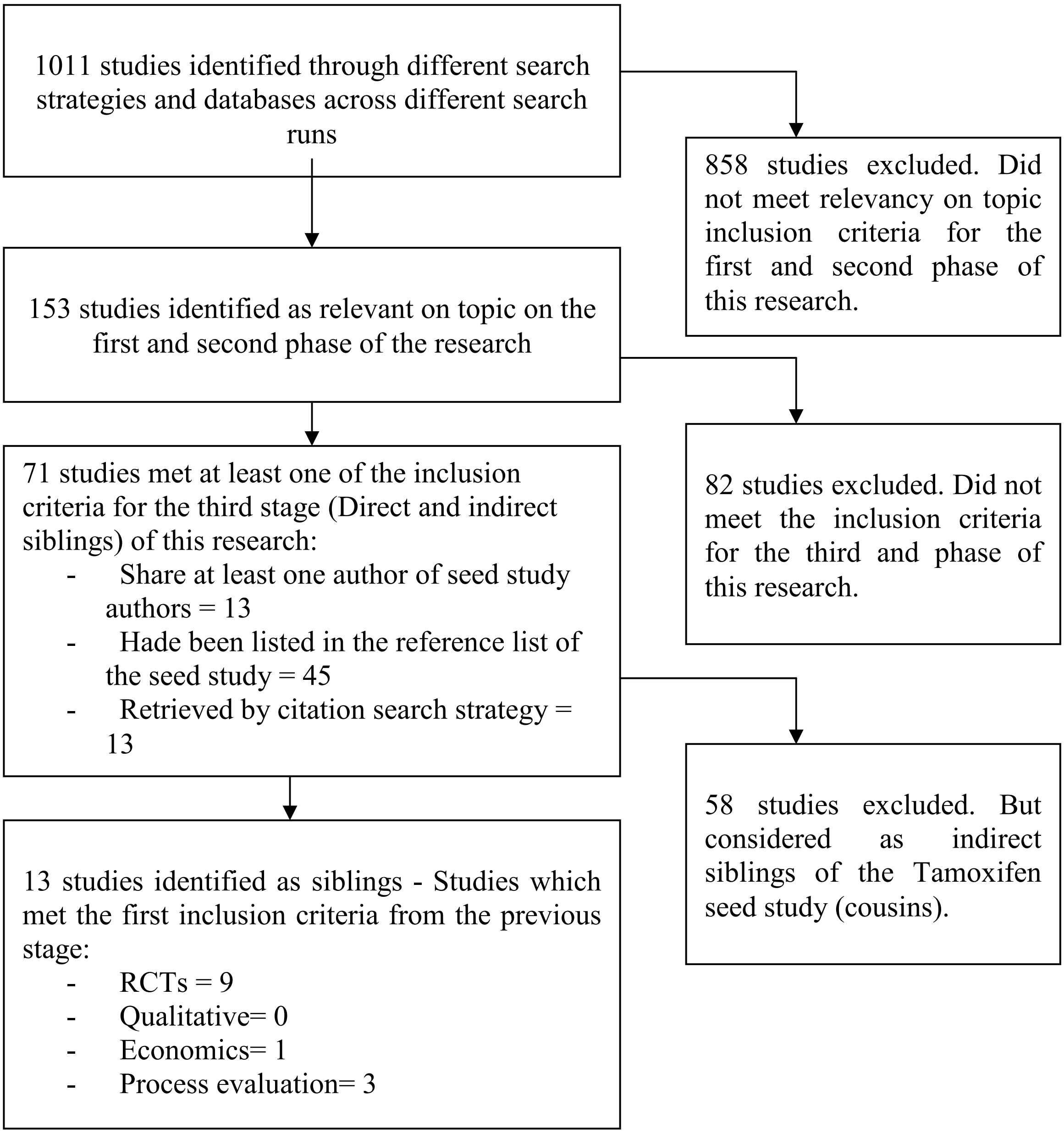


Figure 11: Overall Siblings Retrieval & Identification Process for CLD Seed Study

6.6 In Vitro Fertilisation (IVF) Seed Study

A total number of 70 relevant studies for the IVF seed study were identified, 13 studies met the inclusion criteria and extracted from the relevant set of studies as sibling studies.

For the IVF seed study, a full list of authors was available on WoK, where it was indexed under all authors' names and under the group name as well. This seed study was not included in the CINAHL database. For SCOPUS it was indexed under the first author only. In PubMed it was available under the group name only (but PubMed was searched using all of the author names as well, as done for the CLD seed study).

6.6.1 Recall

Table 39 shows that related articles search on PubMed scored the best recall with value of 62% by retrieving 8 of the sibling studies, followed by author-subject on PubMed and Sophisticated-Author-Subject with recall value of 39%. Several search strategies scored the third best recall value by retrieving 3 siblings scoring a recall value of 23%, i.e. Citation search on WoS.

6.6.2 Precision

According to Table 39, precision calculations indicated that subject search on e-library/MetaLib scored the best precision value of 27%, while SCOPUS subject search with Hedges economics filter scored the second best precision with value of 25%, followed by sophisticated author-subject search on PubMed with precision value of 14%.

6.6.3 Odds Estimator

As illustrated in Table 39, subject search on e-library/MetaLib is the most likely search to retrieve sibling studies rather than non-siblings by scoring odds estimator of value 61.1. Second was SCOPUS subject search with Hedges economics filter with odds estimator value of 45.39, followed by sophisticated author-subject search on PubMed an odds estimator of 31.35.

Table 39: The IVF Search Strategies and Databases Retrieval

IVF

Search Strategy	**Siblings (A)**	**Siblings N/R©**	**Non Siblings (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Estimator**
PubMed-Related articles	8	5	98	0.62	62%	0.08	8%	25.13
PubMed-Author-subject	5	8	909	0.39	39%	0.01	1%	0.5
Citation(Web of Science)	3	10	53	0.23	23%	0.05	5%	8.97
Subject search (e-library)	3	10	8	0.23	23%	0.27	27%	61.09
SCOPUS Author-subject (Author(s))	3	10	47	0.23	23%	0.06	6%	10.15
SCOPUS citation	0	13	0	0	0%	0	0%	0
CINAHL Author subject	0	13	0	0	0%	0	0%	0
CINAHL citation	0	13	0	0	0%	0	0%	0
WoK-Author - subject	1	12	13	0.08	8%	0.07	7%	10.41
E-lib - Author subject	1	12	47	0.08	8%	0.02	2%	2.82
Economics-Hedges filter	0	13	9	0	0%	0	0%	0
Costs-Hedges filter	0	13	9	0	0%	0	0%	0
Qualitative-Hedges filter	0	13	28	0	0%	0	0%	0
E-lib-Economics	2	11	240	0.15	15%	0.01	1%	1.06
E-lib-Costs	2	11	166	0.15	15%	0.01	1%	1.61
E-lib-Qualitative	3	10	175	0.23	23%	0.02	2%	2.51
WoK-Economics	0	13	8	0	0%	0	0%	0
WoK-Costs	0	13	8	0	0%	0	0%	0
WoK-Qualitative	0	13	31	0	0%	0	0%	0
Sophisticated-Author-Subject	5	8	32	0.39	39%	0.14	14%	31.35
Sophisticated-Author-Subject-Qualitative	1	12	8	0.08	8%	0.11	11%	0.8
Sophisticated-Subject-Economics	0	13	9	0	0%	0	0%	0
Sophisticated-Subject-Costs	0	13	6	0	0%	0	0%	0
SCOPUS-Qualitative	0	13	4	0	0%	0	0%	0
SCOPUS-Costs	0	13	9	0	0%	0	0%	0
SCOPUS-Economics	1	12	3	0.08	8%	0.25	25%	45.39
CINAHL-Qualitative	0	13	0	0	0%	0	0%	0
CINAHL-Costs	0	13	0	0	0%	0	0%	0
CINAHL-Economics	0	13	0	0	0%	0	0%	0

Total SIBLINGS R without duplicates	13
Total NON-SIBLINGS R without duplicate	72
Total Relevant R without duplicates	85
Total non-relevant R without duplicates (D)	**1637**

Odds Estimator = A * (D - B)/ B * C

6.6.4 Sibling Studies Retrieval by Study Type

Again, each sibling study was categorised into one of the four sibling types, accordingly, Table 40 shows that subject search on e-library/MetaLib odds estimator with value of 101.4 was most likely to retrieve the RCT siblings rather than non-RCT siblings, followed by SCOPUS subject search with Hedges economics with odds estimator value of 60.52 and finally, sophisticated author-subject search on PubMed had the third best odds estimator with value of 40.13.

There were no qualitative siblings retrieved for the IVF seed study. As for the economics siblings, author-subject on E-library with the Hedges qualitative filter and author-subject on PubMed search were the only search strategy/database to retrieve one economics sibling for the IVF seed study scoring odds estimator value of 9.33 and 1.8 respectively , as shown in Table 40.

According to Table 40, sophisticated-Author-Subject on PubMed Hedges qualitative filter with odds estimator value of 101.81, had the best likelihood of retrieving process evaluation rather siblings than non siblings followed by sophisticated author-subject search on PubMed with odds estimator value of 25.08, and third was SCOPUS Author-subject with odds estimator value of 17.37.

Table 40: IVF Search Strategies and Databases Odds Estimator per Sibling Type

Search Strategy	RCTs			Qualitative			Economical evaluation			Process evaluation		
	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE	RR	R/N	OE
PubMed-Related articles	6	3	33.14	0	0	0	0	1	0	1	2	8.31
PubMed-Author-subject	3	6	0.9	0	0	0	1	1	1.8	1	2	0.9
Citation(Web of Science)	2	7	8.8	0	0	0	0	1	0	1	2	15.41
Subject search (e-library)	3	6	101.4	0	0	0	0	1	0	0	2	0
SCOPUS Author-subject (Author(s))	2	7	9.92	0	0	0	0	1	0	1	2	17.37
SCOPUS citation	0	10	0	0	0	0	0	1	0	0	2	0
CINAHL Author subject	0	10	0	0	0	0	0	1	0	0	2	0
CINAHL citation	0	10	0	0	0	0	0	1	0	0	2	0
WoK-Author - subject	1	8	15.72	0	0	0	0	1	0	0	2	0
E-lib - Author subject	1	8	4.35	0	0	0	0	1	0	0	2	0
Economics-Hedges filter	0	10	0	0	0	0	0	1	0	0	2	0
Costs-Hedges filter	0	10	0	0	0	0	0	1	0	0	2	0
Qualitative-Hedges filter	0	10	0	0	0	0	0	1	0	0	2	0
E-lib-Economics	2	7	1.94	0	0	0	0	1	0	0	2	0
E-lib-Costs	2	7	2.81	0	0	0	0	1	0	0	2	0
E-lib-Qualitative	1	8	1.17	0	0	0	1	0	9.33	1	2	4.66
WoK-Economics	0	10	0	0	0	0	0	1	0	0	2	0
WoK-Costs	0	10	0	0	0	0	0	1	0	0	2	0
WoK-Qualitative	0	10	0	0	0	0	0	1	0	0	2	0
Sophisticated-Author-Subject	4	5	40.13	0	0	0	0	1	0	1	2	25.08
Sophisticated-Author-Subject-Qualitative	0	10	0	0	0	0	0	1	0	1	2	101.81
Sophisticated-Subject-Economics	0	10	0	0	0	0	0	1	0	0	2	0
Sophisticated-Subject-Costs	0	10	0	0	0	0	0	1	0	0	2	0
SCOPUS-Qualitative	0	10	0	0	0	0	0	1	0	0	2	0
SCOPUS-Costs	0	10	0	0	0	0	0	1	0	0	2	0
SCOPUS-Economics	1	9	60.52	0	0	0	0	1	0	0	2	0
CINAHL-Qualitative	0	10	0	0	0	0	0	1	0	0	2	0
CINAHL-Costs	0	10	0	0	0	0	0	1	0	0	2	0
CINAHL-Economics	0	10	0	0	0	0	0	1	0	0	2	0

6.6.5 IVF Siblings Overall Retrieval Process

The overall retrieval for the IVF seed study was 1650 studies, retrieved by all search strategies from different databases, and after applying the initial screening process, 72 candidate sibling studies were extracted. After applying sibling inclusion criteria 13 direct sibling studies were identified, see Figure 12.

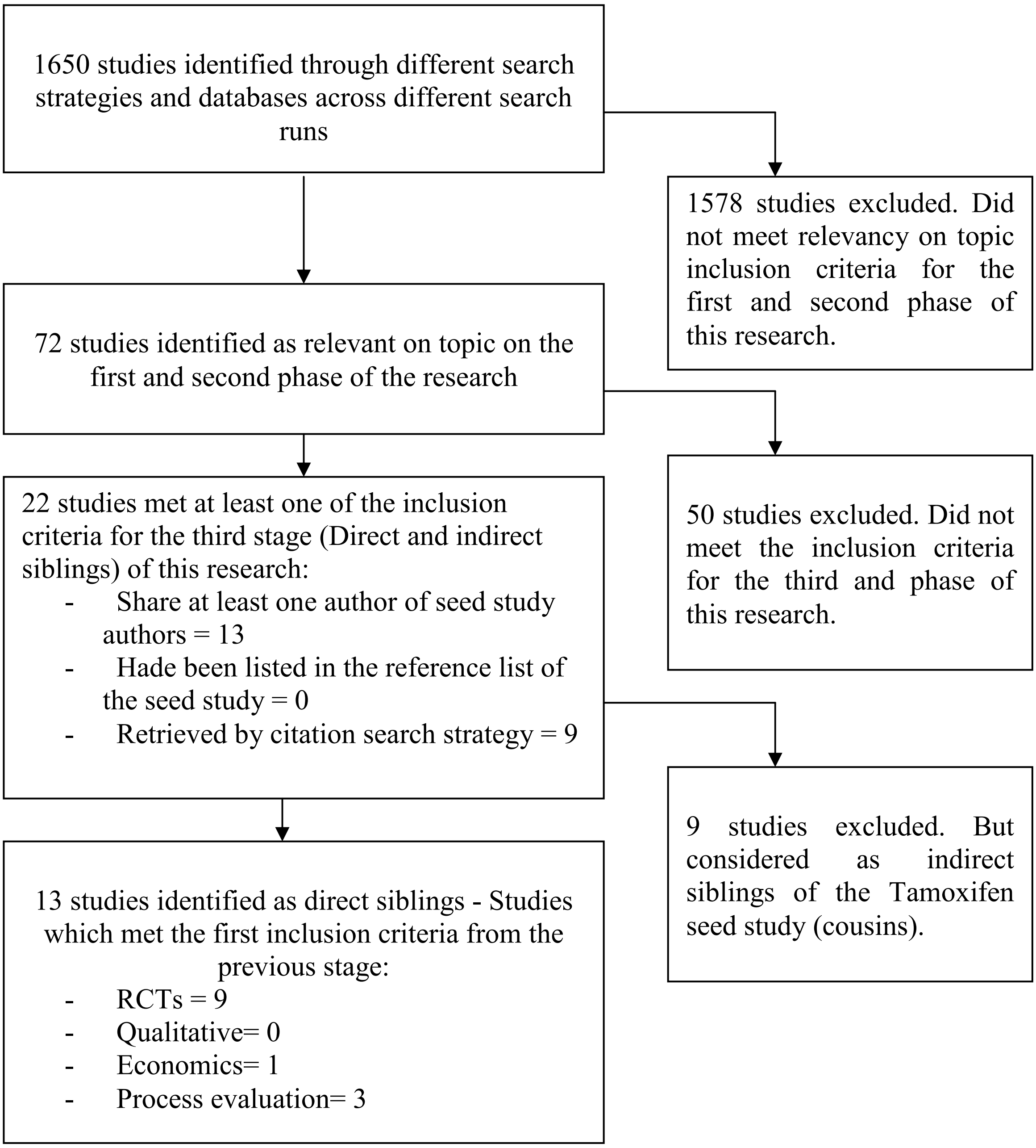


Figure 12: Overall Siblings Retrieval & Identification Process for the IVF Seed Study

6.7 Research Hypotheses and Hypotheses Test

As explained in the methodology Section 4.10, a set of hypotheses were generated. Both Friedman and Kruskal-Wallis tests were used to test various hypotheses.

The first set of hypotheses examines whether search strategy or database might affect the retrieval performance of sibling studies, therefore it will examine the existence of a relationship between recall or precision, and either search strategy or database. The second set of hypotheses aims to examine the relationship between the number of siblings retrieved and search strategy and database. Moreover, it aims to investigate the relationship between retrieval likelihood (Odds Estimator) and search strategy or database A Kruskal-Wallis test was conducted to test the general behaviour of recall, precision, odds estimator and number of siblings retrieved and dependency on search strategy and/or databases but blinding seed study clinical area (by pooling all the values together). Table 41 (search strategy) shows that search strategy does significantly affect the values of recall, precision, odds estimator and number of direct siblings retrieved (as one might expect). However, Table 42 shows that the performance is independent of database.

Table 41: Test of Significance of Retrieval Measurements and Search Strategy

Test Statistics[b,c]

			Recall	Precision	Odds Estimator	Siblings
Chi-Square			52.214	33.110	37.678	32.915
Df			11	11	11	11
Asymp. Sig.			.000	.001	.000	.001
Monte Carlo Sig.	Sig.		.000[a]	.000[a]	.000[a]	.000[a]
	99% Confidence Interval	Lower Bound	.000	.000	.000	.000
		Upper Bound	.000	.000	.000	.000

a. Based on 10000 sampled tables with starting seed 1314643744.

b. Kruskal Wallis Test

c. Grouping Variable: SearchStg

Table 42: Test of Significance of Retrieval Measurements and Database

Test Statistics[b,c]

			Recall	Precision	Odds Estimator	Siblings
Chi-Square			8.291	3.353	3.091	5.217
Df			6	6	6	6
Asymp. Sig.			.218	.763	.797	.516
Monte Carlo Sig.	Sig.		.211[a]	.782[a]	.812[a]	.529[a]
	99% Confidence Interval	Lower Bound	.201	.771	.802	.516
		Upper Bound	.222	.793	.822	.542

a. Based on 10000 sampled tables with starting seed 1535910591.

b. Kruskal Wallis Test

c. Grouping Variable: DataBase

The third set of hypotheses aimed to investigate the relationship between seed study clinical area and search strategy performance and/or database performance. For this set of hypotheses, both the Kruskal-Wallis and Friedman tests were used to test the effect which clinical area (as a grouping variable) might have on search strategy and database performance. Table 43 shows that seed study clinical area significantly affects measurements performance in general.

Table 43: Test of Significance of Retrieval Measurements and Clinical Area

Test Statistics[b,c]

			Recall	Precision	Odds Estimator	Siblings
Chi-Square			21.254	44.888	31.842	52.368
Df			4	4	4	4
Asymp. Sig.			.000	.000	.000	.000
Monte Carlo Sig.	Sig.		.000[a]	.000[a]	.000[a]	.000[a]
	99% Confidence Interval	Lower Bound	.000	.000	.000	.000
		Upper Bound	.001	.000	.000	.000

a. Based on 10000 sampled tables with starting seed 303130861.

b. Kruskal Wallis Test

c. Grouping Variable: ClinicalArea

The Friedman test was conducted to evaluate the differences between the median of measurements matrices between the five seed studies. In other words, it is a useful test to estimate the dependency of performance behaviour and clinical area. According to Tables 44, 45, 46 and 47 recall, precision, odds estimator and number of siblings retrieved changed significantly between different seed studies, with $\chi 2$ (4) = 32.206, $\chi 2$ (4) = 37.623, $\chi 2$ (4) = 26.826 and χ^2 (4) = 43.046 at $p< 0.001$, which matched the results obtained by Kruskal-Wallis according to Table 43 (as might be expected, given the relationship between the Kruskal-Wallis and Friedman tests).

Table 44: Recall and Clinical Area Dependency

Test Statistics[a]

N	30
Chi-Square	35.252
Df	4
Asymp. Sig.	.000

a. Friedman Test

Table 45: Precision and Clinical Area Dependency

Test Statistics[a]

N	30
Chi-Square	55.097
Df	4
Asymp. Sig.	.000

a. Friedman Test

Table 46: Odds Estimator and Clinical Area Dependency

Test Statistics[a]

N	30
Chi-Square	37.350
Df	4
Asymp. Sig.	.000

a. Friedman Test

Table 47: Number of Siblings Retrieved and Clinical Area Dependency

Test Statistics[a]

N	30
Chi-Square	55.656
Df	4
Asymp. Sig.	.000

a. Friedman Test

6.8 Sibling Studies Publication Time

Publication time of sibling studies was investigated in order to detect possible publication pattern of sibling studies in relation to seed study publication date. Each sibling publication date was investigated and classified into after, before or on the same date category as following:

Table 48: Sibling Studies Publication Time

	After				***Before***				**On the same year as the seed†**			
	RCT	Q	E	P	RCT	Q	E	P	RCT	Q	E	P
IDEATeL	**37**				**19**				**14**			
	18	12	2	5	4	6	1	8	6	4	1	3
Tamoxifen	**62**				**40**				**8**			
	49	5	--	8	27	7	--	6	6	2	--	1
BMS	**7**				**0**				**1**			
	2	5	--	--	--	--	--	--	1	--	--	--
CLD	**6**				**3**				**6**			
	3	--	--	3	2	1	--	--	4	--	1	1
IVF	**8**				**4**				**1**			
	6	--	1	1	2	--	--	2	1	--	--	--

† The seed study is included.

Table 48 shows that there is no stable publication time for sibling studies in relation to the seed study, however it appears that the proportion of sibling studies published after the seed study is greater than sibling studies published before or at the same time as the seed study. Most of the siblings published after the seed study are RCTs for all of the seed studies except for the BMS seed study where almost all of the siblings retrieved are published after the seed study itself and they were qualitative siblings mostly. The qualitative, economics and process evaluation siblings proportion was almost the same before and after the seed study (+/- 1-2 siblings on each side of the time scale).

6.9 Multiple Database Retrieval

The recall was low for all seed studies with the highest recall of value of 71% (WoK author-subject search for the IDEATeL seed study) which would not be sufficient for systematic reviewers. Evidence suggests that for qualitative and policy research it is necessary to choose more than one database for searching. Inspecting each search strategy and databases performance individually, combining some of the best performers appears to be the most promising approach, although it is also important to consider the unique retrievals. (Section 6.10)

According to Table 49, searching more than one database is resourceful, simple author-subject search alone provided a good (but not very good) and different recall for each seed study from different databases. The retrieval appeared to be influenced by seed study clinical topic. WoK and PubMed were the databases which contributed to retrieval for all seed studies. This might be because of the different coverage of each database and therefore they complement each other's performance. There was a lot of overlap between WoK and SCOPUS (as both are very close in their coverage) and therefore combining both databases does not seem to boost recall much (only some cases, i.e.Tamoxifen seed study). E-library author-subject search provided some unique siblings and therefore enhanced the recall, this demonstrates the potential of a federated search strategy. The combination of search strategies varied among seed studies, which means that database selection is heavily influenced by clinical area and database coverage, for example for the IVF seed study more than half of the sibling were located by searching PubMed (both related articles and author-subject search). Moreover, some of the searches contributed to the retrieval by one sibling only. This might increase the recall but the precision will be dramatically affected.

Table 49: Multiple Databases Retrieval Performance

Seed Study	Search Strategy	No. of siblings retrieved	Total no. siblings	Recall	Precision
IDEATeL	WoK AS , PubMed AS, PubMed RA, E-Library AS & E-library with costs filter	66	70	94%	15%
Tamoxifen	WoK AS, SCOPUS AS, CINAHL AS, PubMed AS & E-Library AS	110	111	99%	13%
CLD	PubMed AS (simple), PubMed RA, WoK AS, SCOPUS AS & E-Library AS	12	13	92%	2%
BMS	PubMed AS, WoK AS, & E-lib AS	8	8	100%	5%
IVF	PubMed RA, PubMed AS, E-library with filters&	13	13	100%	1%

6.9.1 Unique Retrieval and Overlap

Unique retrievals were identified manually between the three top performing search strategies and databases based on the number of siblings which each strategy or databases retrieve (Table 50). The three top performing databases and/or search strategies that retrieved the highest number of siblings were investigated thoroughly and overlap was investigated. The overlap indicates that when the same record is retrieved by multiple strategies and/or databases signals its relevance, (Saracevic and Kantor 1988). In other words, records found in the intersection have a higher odds of being relevant (the matter in question is searched independently and then the overlap is examined). Venn diagrams illustrated in Figures 13-18 show such overlap for the five seed studies.

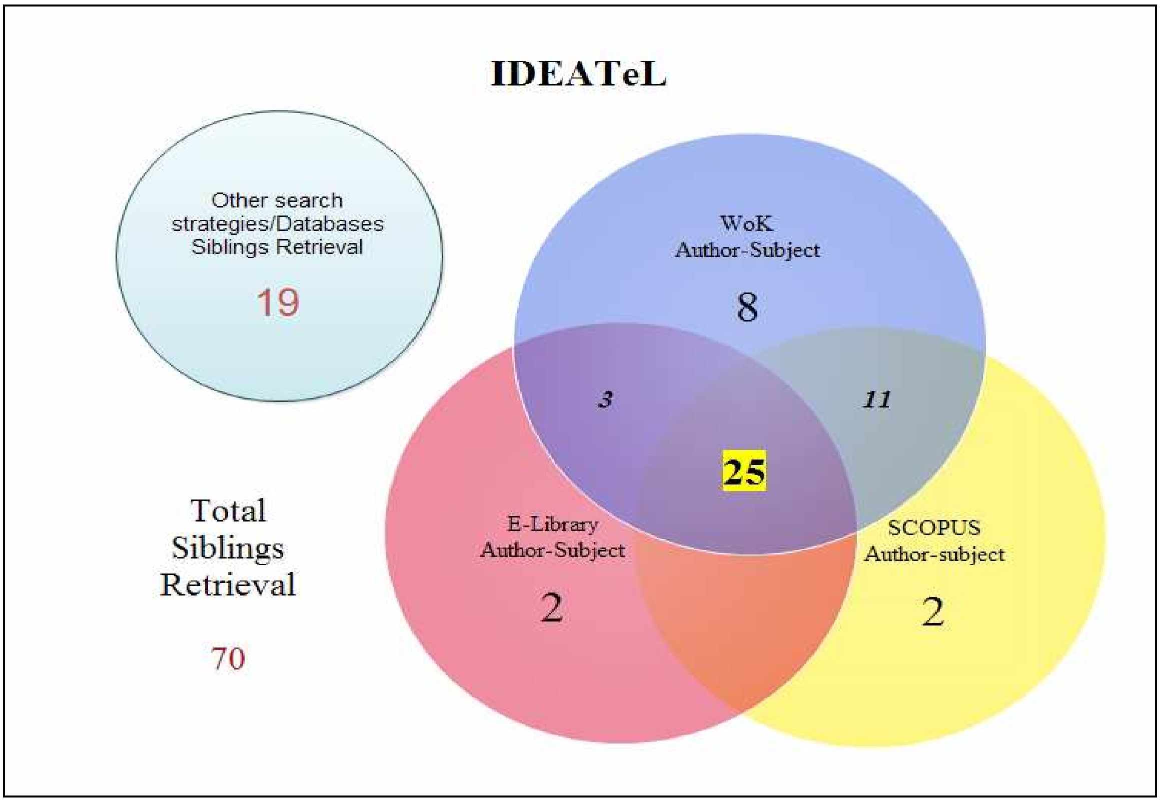


Figure 13: Databases Retrieval and Unique Contribution for IDEATeL Siblings

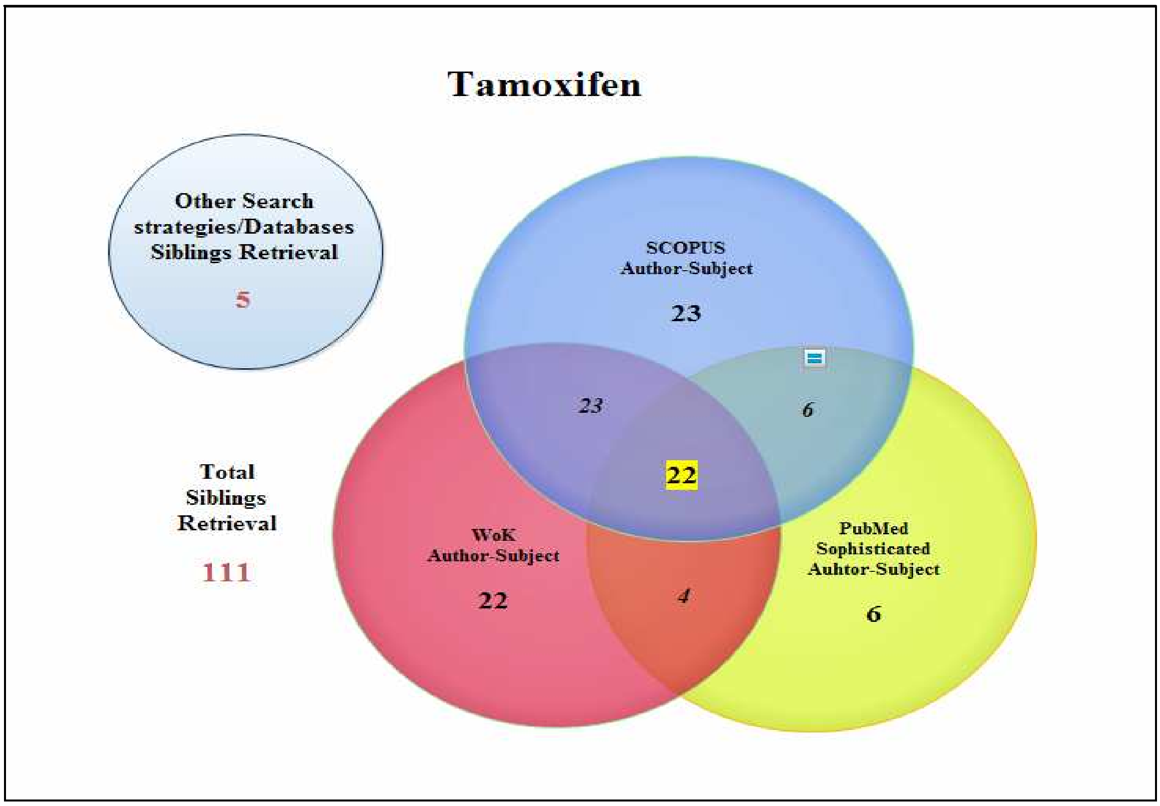


Figure 14: Databases Retrieval and Unique Contribution for Tamoxifen Siblings

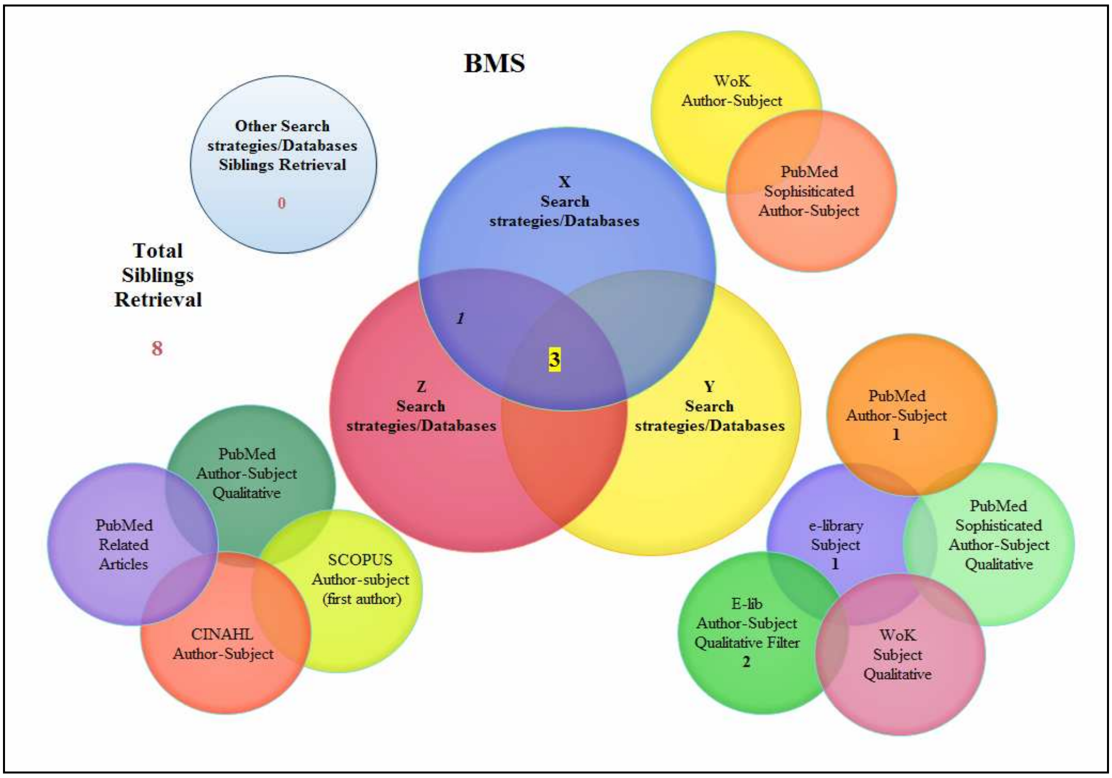


Figure 15: Databases Retrieval and Unique Contribution for BMS Siblings

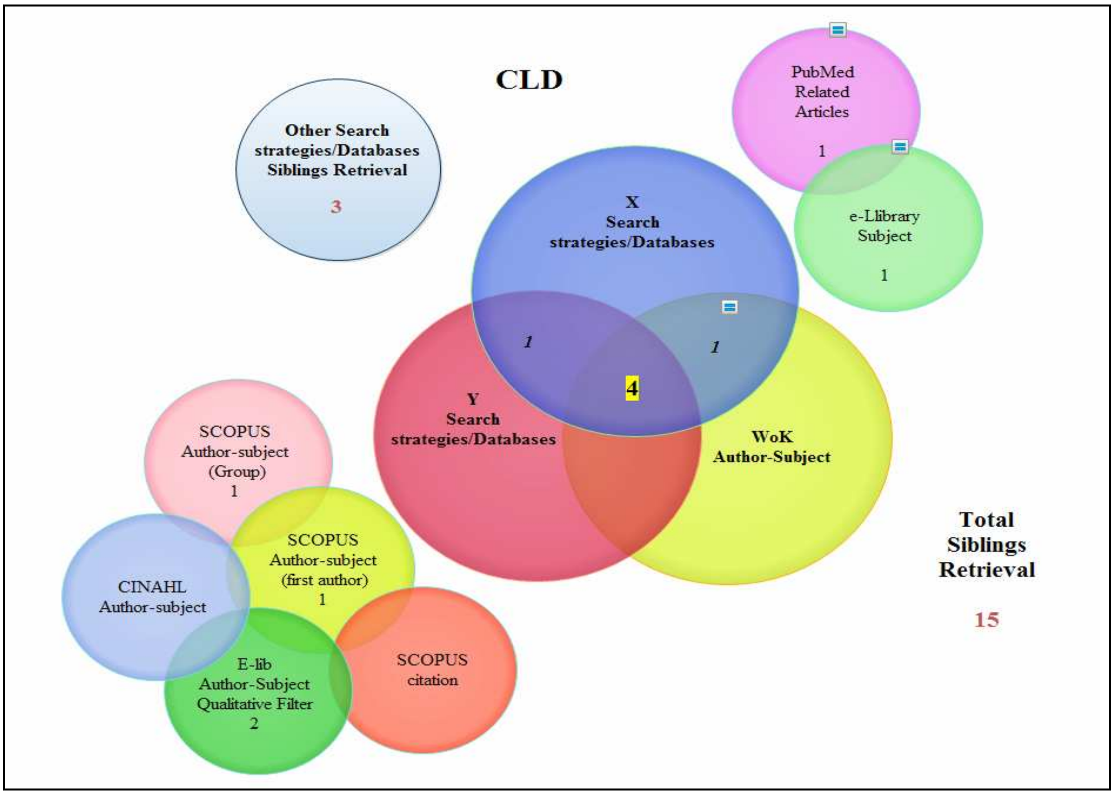


Figure 16: Databases Retrieval and Unique Contribution for CLD Siblings

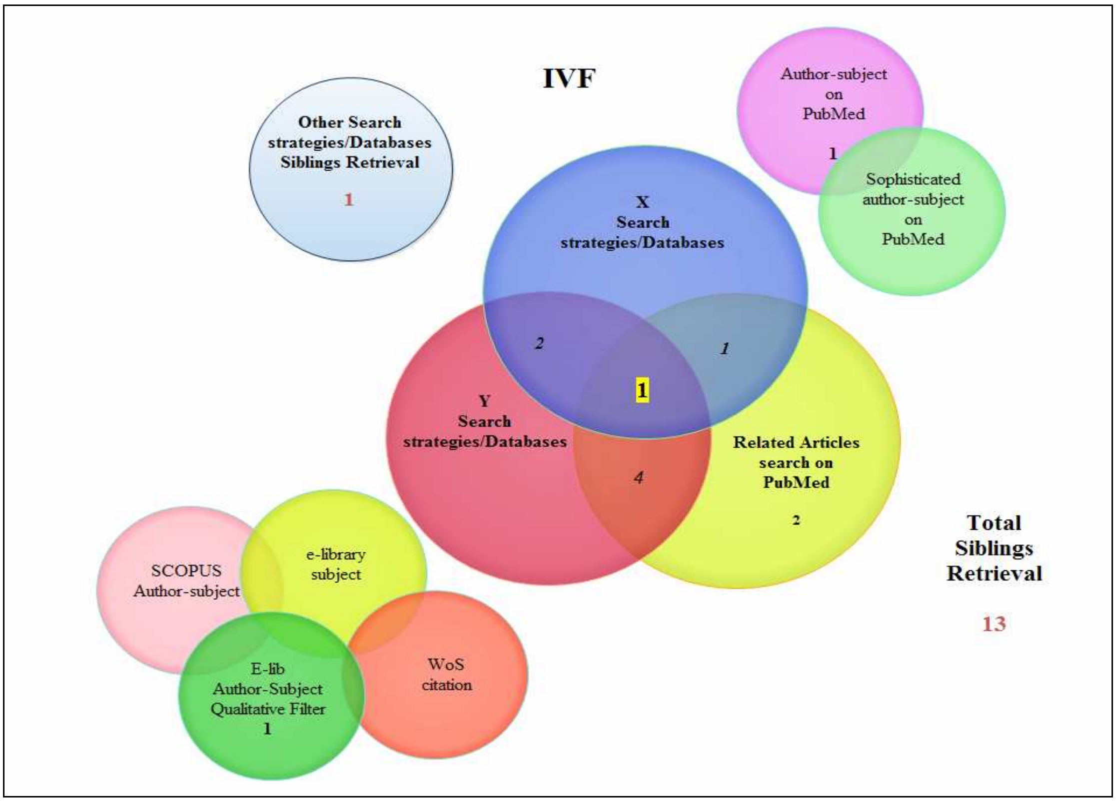


Figure 17: Databases Retrieval and Unique Contribution for IVF Siblings

PubMed related articles, author-subject search on PubMed, SCOPUS, CINAHL, WoK and E-Library search all retrieved some studies not retrieved by other search strategies. Citation reference search did not identify any unique siblings. The sophisticated author-subject on PubMed only retrieved unique studies for the Tamoxifen seed study (Table 48).

It should be mentioned that for Figures 15-17 the group of small circles at the sides represent the group of search of the central group of search strategies (Y, X & Z), for example; Y search strategies and databases in Figure 17 refers to SCOPUS Author-subject, E-Library author-subject (Qualitative filter), WoS citation search and subject search on e-library while X group of search and databases refers to author-subject search on PubMed and sophisticated author-subject search on PubMed. This is because in these cases the top performers were a group of search strategies, not a single search as in the IDEATeL (Figure 13) and the Tamoxifen (Figure 14).

Table 50: Unique Contribution of Search Strategies and Databases

Search strategy	IDEATeL	Tamoxifen	BMS	CLD	IVF
PubMed-Related articles	--	--	--	1 (6.6%)	2 (15.3%)
PubMed-Author-subject	4 (5.7%)	2 (1.8%)	1 (12.5%)	3 (20%)	1 (7.7%)
Citation(Web of Science)	--	--	--	--	--
Subject search (e-library)	--	--	--	1 (6.6%)	--
SCOPUS Author-subject (Group)	--	--	--	--	--
SCOPUS Author-subject (Author(s))	1 (1.4%)	16 (14.4%)	--	2 (13.3%)	--
SCOPUS citation	--	--	--	--	--
CINAHL Author subject	1 (1.4%)	3 (2.7%)	--	--	--
CINAHL citation	--	--	--	--	--
WoK-Author - subject	3 (4.3%)	19 (17%)	--	--	--
E-lib - Author subject	3 (4.3%)	1 (0.9%)	2 (25%)	--	--
Economics-Hedges filter	--	--	--	--	--
Costs-Hedges filter	--	--	--	--	--
Qualitative-Hedges filter	--	--	--	--	--
E-lib-Economics	--	1 (0.9%)	--	--	--
E-lib-Costs	2 (2.9%)	--	--	--	--
E-lib-Qualitative	1 (1.4%)	--	--	--	1 (7.7%)
WoK-Economics	--	--	--	--	--
WoK-Costs	--	--	--	--	--
WoK-Qualitative	1 (1.4%)	--	--	--	--
Sophisticated-Author-Subject	--	6 (5.4%)	--	--	--
Sophisticated-Author-Subject-Qualitative	--	4 (3.6%)	--	--	--
Sophisticated-Subject- Economics	--	--	--	--	--
Sophisticated-Subject-Costs	--	--	--	--	--
SCOPUS-Qualitative	--	--	--	--	--
SCOPUS-Costs	--	--	--	--	--
SCOPUS-Economics	--	--	--	--	--
CINAHL-Qualitative	--	--	--	--	--
CINAHL-Costs	--	--	--	--	--
CINAHL-Economics	--	--	--	--	--
Total Number of siblings retrieved	**70**	**111**	**8**	**15**	**13**

6.10 Summary: Reflection on Retrieval Performance

The metrics used included the standard information retrieval performance measures (for search strategies/databases) of recall and precision, with a third measure, the odds estimator that focuses on ranking search strategies and/or databases based on their efficiency of retrieving the siblings rather than non-siblings. This section summarises, and presents the top three performing strategies for each independent seed study; based on the results of the three measures used in this research as shown in Table 51.

According to Tables 51 and 52 the simple author-subject search performed better than either related articles search or citing search for all of the seed studies. The simple author-subject search strategy was the top search strategy compared to the other two strategies, although, the PubMed related articles search was one of the top performing search strategies for three of the seed studies where for these three it scored the top recall score for two of seed studies. It can be concluded that author-subject search strategy can be considered a winner over the related articles search strategy and/or citing reference search strategy.

The sophisticated author-subject search provided a good recall with a relatively manageable retrieval, however it did not outperform the simple author-subject search strategy on SCOPUS and WoK, but for the PubMed it provided a better recall than the simple author-subject for some of the seed studies.

The performance of the author-subject search strategy varied among different databases. WoK and SCOPUS databases provided a good recall with a readable retrieval size. The related articles search and author-subject search on PubMed retrieval size were slightly bigger than Wok and SCOPUS. Although both searches did not outperform either SCOPUS or WoK recall, they did provide a good recall for some of the seed studies suggesting that related articles search in particular have good chances in retrieving sibling studies, the direct siblings at this final stage.

In terms of precision, the author-subject search provided good precision, often high precision values. The CINAHL database was the winner in most of the seed studies when precision was the criterion. And it was ranked among the best databases to have the best likelihood of retrieving direct sibling studies rather than non-siblings. WoK and SCOPUS provided good chances of retrieving direct sibling studies rather than non-siblings with a good precision value, but not as much as likelihood as the CINAHL database could have. The sophisticated author-subject search strategy provided a good precision as well, however when compared with the simple search strategies it cannot be considered a winner, specially when looking at the odds estimator ranking as the ranks always indicated the superiority of the simple author-subject search however on databases other than PubMed (Tables 51 and 52).

Incorporating the filters with the author-subject search strategy either simple or sophisticated showed a good performance in retrieving the siblings, on different databases especially for CINAHL and SCOPUS. To make for fairer and easier comparisons, I used a general search filter based on the PubMed search filters for SCOPUS, WoK and E-Library). The qualitative filter was the search filter to have achieved good performance compared to the other search filters. It might be convenient to further develop search filters for these databases (SCOPUS and WoK) as they appear to provide a good performance.

Relative performance of the search strategies and databases showed a fairly stable pattern of performance over the three stages of sibling studies retrieval process among the five seed studies. In other words, the top performing strategies in the first phase were usually the top performing strategies for the final phase. For example, WoK author subject remained the top performer for the IDEATeL seed study across all phases, while SCOPUS author subject for the Tamoxifen and PubMed related article for the CLD seed study. For some of the seeds studies the top three performers remained the same but the internal order of them (at the seed level) was changed, i.e. the IDEATeL. Moreover, this stable pattern was more notable in the last two phases (direct and indirect siblings phases and direct siblings phase). The performance differences seems to be directly associated with the type of the clinical area which the seed study involves, but nevertheless it was stable for the individual clinical area. It should be mentioned that the addition of the search filters to the databases would alter the performance order, as expected; however the main pattern (at the seed level) remains fairly stable (Tables 18; 29 and 51).

Table 51: Summary of Search Strategies and Databases Retrieval Performance

IDEATeL		Tamoxifen		BMS		CLD		IVF	
Recall									
WoK Author-subject	71%	SCOPUS Author- subject	67%	WoK Author-subject & Sophisticated Author-subject on PubMed	38%	PubMed related articles & e-library subject search	40%	PubMed related articles	62%
SCOPUS Author-subject	56%	WoK Author- subject	64%	PubMed Author-subject, e-lib- subject, E-lib – Author-subject WoK subject-Qualitativ, Sophisticated-Author-subject-Qualitative, SCOPUS-Qualitative & CINAHL-Qualitative	25%	WoK Author-subject	33%	PubMed Author-subject & Sophisticated-Author-subject	39%
E-library Author-subject	51%	Sophisticated Author-subject on PubMed	34%	PubMed related articles, SCOPUS Author-subject, CINAHL Author-subject, PubMed subject- Qualitative	13%	SCOPUS Author-subject (Group), SCOPUS Author-subject (first author), SCOPUS citation, CINAHL Author-subject & E-library subject- Qualitative	27%	WoS citation e-lib- subject SCOPUS Author-subject & E-library Author-subject-Qualitative	23%
Precision									
CINAHL Author-subject, SCOPUS-Qualitative & CINAHL-(qualitative, economics & costs)	100%	PubMed Soph-Author subject& PubMed Soph-Author-subject-Qualitative	78%	CINAHL Author-subject	50%	WoK Author-subject	83%	e-lib- subject	27%
SCOPUS Author subject	83%	CINAHL Author-subject	51%	SCOPUS-Qualitative	33%	SCOPUS Author-subject & CINAHL Author-subject	80%	SCOPUS-Economics	25%
PubMed Soph-Author-subject – Economics, SCOPUS-Costs &CINAHL citation	75%	WoK-Author – subject	48%	WoK subject-Qualitative	25%	PubMed Author-subject-Qualitative	27%	PubMed Soph-Author-subject	14%

(Continuous)

Table 51: (Continued)

DEATeL	Tamoxifen	BMS	CLD	IVF
		Odds Estimator		
CINAHL Author-subject	PubMed Soph-Author-subject	CINAHL Author-subject	WoK Author-subject	e-lib- subject
SCOPUS – subject - Qualitative	WoK Author-subject	SCOPUS-Qualitative	CINAHL Author-subject & SCOPUS Author-subject	SCOPUS-Economics
WoK Author-subject	SCOPUS Author-subject	WoK subject- Qualitative	PubMed Author-subject- Qualitative	Soph-Author-subject on PubMed

Table 52: Search Strategies and Databases Average Performance for Siblings Retrieval (Average Recall, Precision and OE)

Search strategy	Average Recall	Average Precision	Average Odds Estimator
PubMed-Related articles	29%	5%	5.79
PubMed-Author-subject	27%	6%	2.92
Citation(Web of Science)	8%	6%	2.98
Subject search (e-library)	27%	11%	17.25
SCOPUS Author-subject (Group)	20%	8%	10.96
SCOPUS Author-subject (Author(s))	35%	39%	102.93
SCOPUS citation	11%	8%	3.64
CINAHL Author subject	17%	56%	188.72
CINAHL citation	1%	15%	11.43
WoK-Author - subject	42%	45%	158.62
E-lib - Author subject	20%	16%	16.07
Economics-Hedges filter	2%	1%	0.28
Costs-Hedges filter	2%	2%	0.38
Qualitative-Hedges filter	11%	12%	9.91
E-lib-Economics	8%	2%	0.56
E-lib-Costs	8%	3%	1.33
E-lib-Qualitative	14%	2%	1.47
WoK-Economics	1%	1%	0.18
WoK-Costs	2%	1%	0.24
WoK-Qualitative	14%	12%	10.85
Sophisticated-Author-Subject	33%	30%	26.42
Sophisticated-Author-Subject-Qualitative	15%	31%	4.44
Sophisticated-Subject- Economics	1%	17%	19.63
Sophisticated-Subject-Costs	1%	11%	5.15
SCOPUS-Qualitative	12%	28%	73.94
SCOPUS-Costs	3%	17%	11.6
SCOPUS-Economics	4%	16%	13
CINAHL-Qualitative	8%	28%	27.95
CINAHL-Costs	1%	20%	15.47
CINAHL-Economics	6%	27%	26.19

6.10.1 Sibling Studies Retrieval

It appears that simple author subject search strategy can be effective in retrieving sibling studies. According to the results the recall values were relatively good, even if the figures seem to be low, the recall is per sibling type, and therefore this a partial recall from the overall recall of the search strategy.

The results shows that, it is not necessary that the search strategy or databases is winner for overall siblings retrieval to be among the best performing for specific sibling retrieval, and therefore the search strategy or database might have a low overall recall and a high per sibling study type recall depending on the portion of sibling type it retrieves.

According to Table 53, the general performance pattern suggested that the simple author subject search is the best performing strategy among all other search strategies for all seed studies, while the sophisticated search strategy performed well for some seed studies. The filters appeared to provide a good performance with retrieving siblings, especially the qualitative filter to retrieve qualitative and process evaluation siblings. WoK, SCOPUS and PubMed were the databases that retrieved most of the siblings, but the performance varied among different seed studies as well as different sibling type.

The precision values shows that author subject search strategies appeared to be the best performing strategy for all seed studies again, while the sophisticated search strategy was the top performing search strategy for some seed studies. The filters appeared to have improved the precision of search strategies, i.e. costs and economics for economics siblings. The performance appeared to be depending or influenced by the seeds study clinical area and the type of siblings it retrieves, (Section 6.7). The databases performance appears to be associated with clinical area and the type sibling as well, Table 54.

The odds estimator results did not show any difference from the results above, the simple author subject provided the most likelihood of retrieving sibling studies, but for some seed studies the sophisticated search appeared to be the most likely search strategy to retrieve some siblings of specific type, Table 55. Generally, the performance patterns seem to be dependenton the clinical area (section 6.7). However the results suggested that the simple search strategy appeared to be useful to retrieve sibling studies, and if siblings with specific type are required the filters can be useful.

Moreover, creating a gold standard for sibling studies retrieval appears to be complicated. At an early stage of this research, authors of the seed studies were contacted to confirm on the initial list of siblings that was identified during the pilot sudy. However, only one author of the IDEATeL seed study responded and provided a list of the IDEATeL publications. The provided list was invistegated in depth. It was noticed that all of the studies shared at least one author of the seed study which provided a good base factor to decide on direct siblings relationship for the other seed studies. Later, authors of other seed studies were contacted and were provided again with an information sheet to explain the research problem, aims and objectives. Only one author each of the BMS and CLD responded this time. From this, appears that contact authors find the concept of sibling studies ambiguous and subject to personal perception, aims and objectives and experience. Even more, for collaporative projects it might be more difficult to come to a consensus on their siblings. In other words, the sibling relatiionship is basically a subjective matter and therefore creating a gold standard is more complicated.

Table 53: Summary of Top Three RecallThree Recalls of Search Strategies/Databases per Sibling Type

	IDEATeL		Tamoxifen		BMS		CLD		IVF	
					Recall					
RCT	WoK-Author-subject	29%	SCOPUS Author-subject WoK-Author-subject	51%	PubMed Soph-Author-Subject	25%	PubMed Soph-Author-Subject	47%	PubMed related articles	46%
	E-lib - Author-subject	24%	PubMed Soph-Author-Subject	23%	*All search strategies retrieved one sibling. (See Appendix Three)*	13%	e-library Subject	40%	PubMed Soph-Author-Subject	31%
	SCOPUS Author-subject	23%	CINAHL Author-subject	15%			WoK-Author-subject	33%	PubMed Author-subject e-library Subject	23%
Qualitative	WoK-Author-subject	20%	SCOPUS Author-subject	6%	WoK subject-Qualitative & WoK Author-subject	25%	All the search strategies (8). *All search strategies retrieved one sibling. (see Appendix Three)*	7%	--	--
	SCOPUS Author-subject	19%	WoK Author-subject & PubMed Soph-Author-Subject PubMed Author-Subject	5%	All the other search strategies. i.e. PubMed subject-Qualitative	13%			--	--
	E-lib - Author-subject	16%	E-lib - Author-subject & PubMed Soph-Author-Subject-Qualitative	3%					--	--
Economics	6 search strategies. *(see Appendix Three)*	4%	--	--	--	--	PubMed Author-subject	7%	PubMed Author-subject E-lib Authorr-subject-Qualitative	8%
	12 search strategies. *(see Appendix Three)*	3%	--	--	--	--				
	5 search strategies. *(see Appendix Three)*	1%	--	--	--	--				
Process evaluation	WoK-Author-subject	13%	WoK Author-subject	10%	PubMed Author-Subject WoK subject-Qualitative	13%	PubMed Author-subject	13%	*All search strategies retrieved one sibling. (see Appendix Three)*	8%
	SCOPUS Author-subject	11%	SCOPUS Author-subject	9%			4 Search strategies	7%		
	PubMed related articles & E-lib - Author-subject	10%	PubMed Soph-Author-Subject	6%						

Table 54: Summary of Top Three Precision of Search Strategies/Databases per Sibling Type

	IDEATeL		Tamoxifen		BMS		CLD		IVF	
					Precision					
	CINAHL Author-subject	53%	PubMed Soph-Author-Subject	53%	CINAHL Author-subject	50%	SCOPUS Author-subject	100%	e-library subject	27 %
RCT	WoK Author-subject	32%	PubMed Soph-Author-Subject-Qualitative	50%	SCOPUS subject-Qualitative	17%	WoK-Author-subject	83%	SCOPUS subject-Economics	25 %
	SCOPUS Author-subject & E-lib - Author-subject	31%	CINAHL Author-subject	42%	PubMed Soph-Author-Subject-Qualitative	9%	CINAHL Author-subject	80%	PubMed Soph-Author-Subject	11 %
	CINAHL subject-Qualitative	67%	PubMed Soph-Author-Subject	10%	WoK subject-Qualitative	25%	CINAHL subject-Qualitative	25%	--	--
Qualitative	SCOPUS subject-Qualitative	57%	PubMed Soph-Author-Subject-Qualitative	9%	SCOPUS subject-Qualitative	17%	WoK subject-Qualitative & PubMed subject-Qualitative	17%	--	--
	SCOPUS Author-subject	25%	E-Library Author-Subject	7%	PubMed Author-Subject	14%	--	--	--	--
	CINAHL subject-Costs	75%	--	--	--	--				
Economics	SCOPUS subject(Costs & Economics) &CINAHL subject-Economics	50%	--	--	--	--	PubMed Author-subject	3%	E-lib Authorr-subject-Qualitative	1%
	PubMed Soph-Author-Subject Economics	33%	--	--	--	--				
	SCOPUS subject-Qualitative	29%	PubMed Soph-Author-Subject-Qualitative	19%	WoK subject-Qualitative	13%	WoK subject-Qualitative PubMed subject-Qualitative	17%	PubMed Soph-Author-Subject-Qualitative	11 %
Process evaluation	CINAHL subject-Costs & SCOPUS subject-Costs	25%	PubMed Soph-Author-Subject	14%	PubMed Author-Subject	2%	PubMed Author-subject	5%	PubMed Soph-Author-Subject	3%
	CINAHL Author-subject	21%	CINAHL Author-subject	10%	--	--	PubMed related articles & E-Library Author-subject-Qualiitative	2%	SCOPUS Author-subject & WoS citation	2%

Table 55: Summary of Top Three Odds Estimator of Search Strategies/Databases per Sibling Type

	IDEATeL	Tamoxifen	BMS	CLD	IVF
			Odds Estimator		
RCT	CINAHL Author-subject	PubMed Soph-Author-Subject	CINAHL Author-subject	WoK subject-Qualitative	e-library subject
	WoK Author-subject	PubMed Soph-Author-Subject-Qualitative	SCOPUS subject-Qualitative	CINAHL Author-subject & SCOPUS Author-subject	SCOPUS subject - Economics
	SCOPUS Author-subject	WoK Author-subject	PubMed Soph-Author-Subject-Qualitative	e-library subject	PubMed Soph-Author-Subject
Qualitative	SCOPUS subject-Qualitative	PubMed Soph-Author-Subject	WoK subject-Qualitative	CINAHL subject-Qualitative	--
	CINAHAL Author-Subject & CINAHL subject-Qualitative	PubMed Soph-Author-Subject-Qualitative	SCOPUS subject-Qualitative	PubMed subject-Qualitative	--
	WoK Author-subject	E-lib - Author-subject	WoK Author-subject	WoK subject-Qualitative	--
Economics	CINAHL subject-(Costs &Economics)	--	--	PubMed Author-Subject	E-lib Authorr-subject-Qualitative
	PubMed Soph-Author-Subject-Qualitative & SCOPUS subject-Costs	--	--		PubMed Author-subject
	SCOPUS-subject-Economics	--	--		--
Process evaluation	CINAHL Author-subject & SCOPUS subject-Qualitative	PubMed Soph-Author-Subject-Qualitative	WoK subject-Qualitative	PubMed subject-Qualitative	PubMed Soph-Author-Subject-Qualitative
	WoK Author-subject	PubMed Soph-Author-Subject	PubMed Author-subject	PubMed Author-subject	PubMed Soph-Author-Subject
	SCOPUS Author-subject	WoK Author-subject	--	WoK subject-Qualitative	SCOPUS Author-subject

Chapter Seven
Discussion

7.1 Introduction

In the discussion, the major findings of this research are reviewed and the implications considered for retrieval of sibling studies. First, the top three performing search strategies and/or databases are discussed, with the implications for retrieval performance considered, in particular how efficient those search strategies and/or databases were in retrieving relevant indirect sibling and direct siblings. The performance patterns found are related to existing evidence, although there is no directly equivalent research published. This research investigated five seed studies from different clinical areas, and therefore it is also important to discuss how clinical area or research type might affect retrieval performance of different search strategies and/or databases. This can only be indicative as there were only five seed studies.

Additionally, the potential benefits of simple search (key search terms chosen from title and abstract of the seed studies) approaches compared to more sophisticated search strategies are considered (MeSH terms were used in combination of key terms from title and/or abstract) (See Section 4.3 and Appendix one), as well as the characteristic performance of the databases themselves.

Considerable effort has been put into the development of search filters and their capabilities and expectations in sibling study retrieval and identification are examined. The effectiveness of search filters is considered. Furthermore, there is the problem of the definition of sibling studies, the difficulties of judging a study as a direct sibling or judging it as an indirect sibling or just a relevant item (on topic) – in other words, how to decide on the relationship between seed study and any study that is relevant to it. This requires consideration of the networks of scholarly communication and authorship.

Finally, the possible role of trial registration number as a search aid for finding sibling studies is considered.

Throughout the discussion, the limitations of this research are considered, as these affect the possible generalisation of the findings.

7.2 Search Strategies and Databases: Main Findings

The results show no overall winner among search strategies or in databases. Search strategies performed differently among the five seed studies. Some of the search strategies performed better than others in one clinical topic, but again that particular search strategy might perform badly for sibling retrieval for another seed study. Database performance does not follow a particular pattern in retrieval, and therefore there was no overall winner among databases either. This section reviews the top three performing search strategies and databases in terms of recall, precision, and odds estimator. The possible effect of clinical area on performance is discussed later in section 7.3 in this chapter. Section 7.5 examines how combinations of search strategies might improve recall.

7.2.1 Effective and Precise Search Strategy(s)

The retrieval performance varied between seed studies, however the author-subject search strategy which was based on simple subject search provided a good recall and precision compared to the other search strategies, as it scored one of the top three places for all seed studies (see Section 6.11, Tables 51 and 52). The sophisticated author-subject search strategy on PubMed perfroamed well for some seed studies. This suggests the possibility of this search strategy for retrieving sibling studies, though it was placed lower than the simple author-subject.

The related articles search on PubMed as a search based on similarity score was another search strategy expected to perform well. It provided a fairly good recall but not precision. (Sections 6.5 and 6.6). This suggests the potential of this search strategy when recall is the main concern, however the clinical area should be taken into consideration. The overall performance was low compared to the simple or sophisticated author-subject strategy. Its performance was closer to the sophisticated search strategy rather than the simple search strategy.

Citation reference search did not identify a good portion of sibling studies which may be due to fact that some of the siblings might have been completed before the seed study was conducted and sometimes because it takes time for a new work to be published, recognised and cited by another work. The overall results suggested that a combination of simple author-subject with a careful choice of search terms is the best approach to use in order to retrieve sibling studies. This is discussed in more detail in Section 7.5. The citation search performed as one of the top performing search strategies in terms of precision as well for one seed study only, the IDEATeL seed. Overall, this tells us that when user is interested in precise retrieval of sibling studies, a simple author-subject search is the best search to do, (See Section 6.10, Table 51).

The results appeared to depend mainly on clinical area (Sections 6.7 and 7.3). This is might be due to the different indexing scheme each which each database employs; the nature of the topic itself and the focus of the database itself. For example the CINAHL database mainly focuses on nursing and allied health literature and therefore it will perform better in retrieving general clinical practice topics such as diabetes (that are the concern of several types of health professional). More specialist medical topics are not likely to be covered well by CINAHL. For example, the IVF seed study was not even indexed in the database and accordingly the CINAHL search strategy failed to achieve results for this seed study, asserting again the influence of the clinical area/topic on the overall performance of the database and even the search strategies.

According to Howes et al. (2004) and Oglivie et al. (2005), searching for social interventions associated with health is known to be problematic and it is possible that the best comprehensive search strategy may depend on the topic and its match with available databases. In this research, it seems that a simple search across a wide variety of databases and a careful choice of the databases within an E-library search should be cost-effective in time, although some decisions need to be made about the stopping point. Overall, there is no clear winner as might be expected, therefore, a need to combine some search strategies is probably indispensable.

Royle and Waugh (2005) claimed that simple search strategies might be useful for busy clinicians, who are interested in rapid and reliable answers for their query. They suggest that a comprehensive search is no longer feasible nor cost-effective, and in their study, they found that using simple search strategy in MEDLINE affected only a small percentage of the total outcome of a few systematic reviews and did not affect the final conclusion significantly. In other words, Royle and Waugh (2005) suggested that it is cost-effective to use simple search strategies and retain precise and reliable systematic review conclusions. In this doctoral research, the simple author subject search strategy (the top performers) produced a mix of sibling studies (RCT, qualitative, process evaluation and economics where available) for each seed study, and these siblings are of assumed high quality, given the databases which retrieved these siblings. Moreover, the quality of the retrieved siblings could be demonstrated clearly by the overlap between the top three performers (See Section 6.9.1, Venn diagrams).

On the other hand, Day et al. (2005) stated that simple search strategies are effective when searching for clinical trials but their sensitivity is not adequate for conducting systematic reviews. However, they can achieve a good level of sensitivity to search for trials of pharmaceutical interventions and for trials of well-defined physical interventions. In conclusion, it seems that a careful use of a simple search strategy might be sufficient for many purposes, but at least one more sophisticated search, aimed at maximising recall should be done if the main aim is to support systematic reviewers (Day et al., 2005).

Odds estimator values also supported the previous conclusion about simple author-subject search strategy being the best performing search strategy to retrieve sibling studies.

Odds estimator tests the likelihood of particular search strategies or databases retrieving sibling studies rather than non-sibling studies. Stokes et al. confirmed that the odds estimator proved to provide a good performance indicator (Section 3.6.3). Therefore, based on Stokes et al. (2009) findings, odds estimator measurements were applied in this research.

Odds estimator values appeared to be associated mostly with precision, i.e. the search strategy/database with the highest precision value has the highest or near to highest odds estimator. There is an association between precision and odds estimator (See Section 6.10, Table 52), for example simple author-subject search strategy scored the highest precision values and the highest odds estimator among other search strategies. CINAHL database scored the highest precision and were ranked first in terms of likelihood of retrieving siblings rather than no-siblings as indicated by its odds estimator value. The WoK database precision values was the second and its odds estimator was again the second highest among other databases, while SCOPUS scored the third best precision and was ranked third in terms of likelihood of retrieving siblings rather than non-sibling studies. This could be due to the fact that both measurements have the B (the number of siblings not retrieved) in the denominator. The same performance pattern was noticed in Stokes et al. (2009) findings. In the end, the odds estimator provided a useful comparison of the odds for group of searches in retrieving siblings rather than non-siblings taking recall and precision into account. In some cases there was a division by zero problem which was dealt with in section (4.5) (See Appendix 8).

7.2.2 Database Performance Comparisons

In this research there was no stable performance pattern for databases. Each database provided a different retrieval performance among different clinical areas (seed studies). SCOPUS and PubMed appeared to be winning over the other databases. MEDLINE is one of the leading databases to use when conducting systematic reviews, it is quick and easy to use, besides it is available freely under PubMed. PubMed's search strategies showed some potential which might suggest that a combination of these search strategies might be useful for retrieving sibling studies; but no more than adequate i.e. the related articles search. And therefore, they should be combined with other search strategies and/or databases to get an optimal performance, (See Section 7.5), The PubMed database focuses on medicine and biomedical sciences and allows a large number of keywords per search and it uses automatic mapping procedure. Several studies have attested to the value of PubMed (Falagas et al., 2008; Stokes et al., 2009). But again each database performed differently among the five seed studies, emphasising differences in clinical area. The CINAHL database was the most precise database.

Variation in database performance is not surprising. Stokes et al. (2009) stated that when nursing studies are needed to be retrieved, both CINAHL and MEDLINE are effective in terms of recall and precision, and are considered to be good performers although MEDLINE performance is less good than that of CINAHL (Section 3.6.3). Again, it is difficult to generalise these findings to other types of study and clinical areas (Stokes et al., 2009). In another study by Baykoucheva (2010), MEDLINE performance was compared with SCOPUS and WoS databases, this time the databases were tested to retrieve drug studies, using selected drugs name(s) as search terms individually or using the drug group name. The results showed that SCOPUS was a good performer in terms of both the number of documents it retrieved and journal coverage, followed by WoS, while MEDLINE retrieval was less good than the other two databases. The results for the drug topic search suggested that both SCOPUS and WoS complemented each other in terms of journal coverage and therefore in order to get a comprehensive retrieval of drug studies both databases should be searched. However if only one database is available, SCOPUS is the best one to use, and if none of these databases is available, then MEDLINE can be used satisfactorily (Baykoucheva, 2010). It can be said that the performance found in this research follows the same pattern as Baykoucheva in terms of databases' general performance, especially for SCOPUS. SCOPUS author-subject search provided a good recall for all of the seed studies (was one of three top performers), while WoK author-subject was one of the top performers for four seed studies.

Other studies have indicated that it is not always adequate to search MEDLINE alone to get a comprehensive retrieval (See Section 7.5). Some recommended the use of CINAHL especially when methodological aspects are the issue (Evans, 2002). The PubMed related articles search function has been tested and found to provide a good proportion of relevant studies when it aims to find new studies to update an existing systematic review (Sampson, 2009) and therefore the strength of this special feature offered from PubMed seems promising for finding sibling studies. In this research the related articles search provided a different performance on the different seed studies, but it yielded a good recall for one of the seed studies – the IVF one (See Section 6.6). This could be explained by the interaction of terminology differences in reporting studies, among principal authors, with the way the related article algorithm works , as it is based on the contents of title and abstract to calculate the relationship – all siblings share terms (See Section 3.5.2)

SCOPUS and WoK have not received that much attention for systematic review work, although they are both considered a rich source for social science research. The coverage of these databases is more general than that of PubMed, but they do include clinical research, and therefore they can be taken into consideration when searching for studies of clinical nature, especially as there are citation reference features provided by both (Sampson et al., 2006b). However, the citation search did not appear to add much value in this research, which agrees with other observations (Sampson, 2009). According to Sampson the newer studies would not have had enough time to be cited and hence a window of 10 years for citing search is necessary to be useful, moreover, Bernstram et al. (2001) noted that citing reference tends to be incomplete. They found that not all the important relevant studies are cited by related work.

In this research the result suggested the selection of databases to search for optimal retrieval depends on the clinical area, however generally, searching WoK, with either PubMed or SCOPUS can produce more comprehensive retrieval results (See Section 7.5).

7.2.3 Retrieval Performance for Particular Type of Siblings

As this research aims to retrieve sibling studies, search strategies and/or database(s) performance to retrieve particular types of sibling was checked. It appears that search strategies and databases' performance are profoundly influenced by the type of siblings being retrieved, indicating the association between the type of sibling and the search strategy retrieval performance as well as database(s) retrieval performance pointing out the influence of research type and/or design on retrieval process.

7.2.3.1 RCT Siblings

The low recall (See Appendix Three, Table 3) overall suggests that there is no clear winner here, which is not surprising. However, averaging recall values from different clinical area brings simple author-subject to the top of all other search strategies on different databases. The PubMed related articles average recall was better than citation search and it was relatively good for the RCT sibling type. Citation search overall performance was not good, however it yielded a comparatively good recall in one of the seed studies on SCOPUS database. Perhaps with different combinations of databases, it may be possible to recommend search strategies that are better at retrieving particular types of siblings, but the results may depend on the topic. In terms of database recall, PubMed, WoK and SCOPUS outperformed the other databases.

Both precision and odds estimator values varied very much among different seed studies, again indicating that the differences among clinical areas might influence the performance of search strategies and databases. Averaged precision values suggested the superiority of simple author-subject search strategy, (see Appendix Three, Tables 2). On the other hand, both related articles and citation search precision values were very low, and it would not seem advisable to use either one if precision is the objective. CINAHL, WoK and SCOPUS databases were the databases that yielded good precision values (See Appendix Three, Tables 2). The Odds estimator follows the same pattern as precision, ranking the simple author-subject search strategies as the most likely strategy to retrieve RCTs rather than non-RCTs (See Appendix Three, Table 1).

For the RCT sibling type, the PubMed database has a randomised clinical trial publication type feature which will assist the retrieval of this type of sibling, if the search is particularly targeting RCTs only. In this research, there are sibling types of interest other than only RCTs, and therefore this facility is not useful for this research, but it clarifies the great focus which the RCTs receive rather than other studies of different research design.

In conclusion, the simple author-subject search strategy appears to provide a good recall and precision for retrieving RCT siblings, with different performance on different databases. This suggests that retrieval performance might be associated with the type of siblings, the search is targeting and the characteristic of the databases being searched, mainly its focus and coverage and the indexing scheme it employs.

7.2.3.2 Qualitative Siblings

The recall (See Appendix three, Table 6) values were lower in the case of qualitative siblings type than they were with RCTs. There is no clear winner again, but the resultsthat the simple author-subject search strategy performance was the best among other search strategies, but the low precision value suggests that combining simple subject search (e-library or WoK with authors) with a more sophisticated author-subject (PubMed) might provide better performance for retrieving qualitative sibling studies as the sophisticated search was a generally good performer (Appendix Three, Table 5).

Furthermore, the qualitative filter appeared to perform well with this type of sibling. The results showed that combining simple author-subject search (eg. WoK) with the Hedges qualitative filter provided a good performance, especially in terms of precision compared to other search strategies except for the simple author-subject search on WoK. This should be expected as this filter used a focused search with terms targeting studies with qualitative research (See Section 6.10.1, Table 54). On the other hand, neither related articles nor citation search would be recommended if qualitative studies are the objective.

The qualitative search filter is designed to target qualitative studies but it picked up other studies with other research design i.e. RCTs (See Section 6.10.1, Table 55).

However, all the search strategies and/or databases retrieved qualitative siblings for four seed studies except for the IVF seed study, which implies the influence which clinical area have over search strategies and databases performance, (See Section 7.3). Also perhaps, a careful selection of databases to incorporate into an E-library/Metalib type search might result in good recall as well. However WoK is the best database to search to retrieve qualitative siblings, as WoK has an extended coverage of topics for the social sciences compared to PubMed (for example). The author-subject search either simple or a more sophisticated form on PubMed database might be a good start.

CINAHL, SCOPUS or WoK with the simple author-subject (with or without qualitative filter) or any combination of these databases are recommended to yield good precision values. The combination of databases here is reasonable as each one of these databases focuses on different subject areas. CINAHL mainly focuses on nursing and midwifery subjects and it is famous for its almost precise indexing of qualitative research in particular and therefore it is not surprising that CINAHL was the most precise database for retrieving qualitative siblings for most of the seed studies, whereas WoK and SCOPUS cover most scientific fields and social sciences. Therefore the scope of coverage of either WoK or SCOPUS is broader than some of the other databases i.e. PubMed, and hence general (rather than just nursing) social science research studies are more likely to be captured (Falagas et al., 2008).

Furthermore, the simple author-subject search strategy and simple author-subject combined with a qualitative filter has the best likelihood of retrieving qualitative siblings rather than non-siblings specially when combined with SCOPUS, CINAHL and WoK. Odds estimator performance again reflects the database's main focus and the implication on retrieval performance. MEDLINE's performance was generally good (though not very good) for retrieving qualitative studies, either with use of simple or sophisticated search strategy combined with qualitative filter - odds estimator values of simple and sophisticated searches varied at clinical area level giving no definitive performance indicator (See Section 6.10.1), for example the simple search strategy (PubMed) outperformed the sophisticated with the CLD and BMS seed study while with the IDEATeL and Tamoxifen it was the other way round. The overall recall and precision of both were almost similar although the odds estimator indicated that combining the simple search strategy with the qualitative filter has a greater likelihood of retrieving sibling studies rather than non-siblings compared to the sophisticated search strategy even when combined with qualitative filter (See Appendix Three, Table 4). There is no definitive explanation for this since for some seed studies there were no siblings for the odds estimator to be calculated, but odds estimator calculations get influenced by the non-siblings being retrieved, and accordingly the bigger the proportion of number of siblings retrieved to number of non-siblings retrieved the better the odds estimator will be.

Qualitative research proved to be both difficult and complex to identify and retrieve, as found in other studies (e.g. Evans, 2002) which might be due to the inappropriate usage or choice of search terms to be used in the studies' titles or abstracts as reported by author(s)

(See Section 3.5.5). In other words the poor reporting of qualitative studies will affect the choice of indexing terms, moreover the difference in indexing process that each database adopts to handle this type of research will have implications on the retrieval process. But the results from this research are consistent with other research findings and recommendations, in that a simple search strategy can provide an optimal performance to retrieve qualitative research or siblings and can be as effective as a complex search (Flemming & Briggs, 2007). Moreover, Flemming and Briggs (2007) claimed that searching CINAHL for qualitative research can identify most of the relevant qualitative research, and this is due to the CINAHL indexing method, as it uses more indexing terms to index qualitative research as well using precise methodological terms compared to other databases. Subirana et al. (2005) recommended searching MEDLINE and CINAHL for qualitative nursing studies for optimal retrieval. However, the study by Flemming and Briggs (2007) recommended that searching CINAHL solely when the query is of specific nursing focus is sufficient. This indicates that the nature of the clinical area will influence the choice of database(s) to search in addition to database(s) retrieval performance (See Section 7.3).

Shaw et al. (2004) evaluated three search strategies' performances for retrieving qualitative research, and simple broad-based terms was one of the evaluated search strategies. The results showed that the simple search strategy yielded most of the potentially relevant records compared to the other search strategies – thesaurus and free text – however the precision was low. The final conclusion and recommendation of the Shaw study was that either one of the search strategies tested in their study can identify relevant qualitative studies but the precision will be poor. Again, this is could be because of the poor reporting in qualitative research description from authors which affects the indexing in databases and therefore, they recommended the usage of at least two search strategies to achieve as much effective retrieval as possible.

Furthermore, Lewin, Glenton and Oxman (2009) and Glenton, Lewin and Scheel (2011) affirmed that qualitative research is less common than expected and seemed to be mostly conducted before the trial (uncommon alongside the trial). The qualitative findings appeared to be poorly integrated with the trial findings and often had major methodological shortcomings. Moreover, the association between qualitative studies and trials that are reported separately is still unclear and need to be explicitly linked to one another to facilitate retrieval (Section 2.4). This might explain the findings of this research – the low number of siblings which were identified

The result above highlighted some difficulties in searching and retrieving qualitative research that arose in this research. When comparing performance of WoK, PubMed, SCOPUS and CINAHL. CINAHL was the best databases to consider when the target of search is qualitative research retrieval as indicated by odds estimator value. The performance detected for CINAHL is a reflection of its indexing process for the CINAHL is known to utilise methodological terms to index qualitative researches (Evans, 2002; Flemming & Briggs, 2007). PubMed database's odds estimator was comparatively low even with a careful choice of the MeSH terms - if the suitable terms available – to search for qualitative research.

In this context, Evans (2002) compared the index terms used to index the same qualitative study in both CINAHL and MEDLINE. It was obvious that MEDLINE index is lacking compared to the CINAHL index. The former indexes qualitative research under the quantitative framework while the latter is focused on using methodological terms that accurately describe qualitative research. Evans' conclusion can support and explain the performance of PubMed here. It appears that NLM indexers really try to do some good indexing by providing appropriate MeSH terms for qualitative research – but CINAHL has a better choice. However even when MeSH terms are available and were used in this research, the search did not perform as expected which only might be because of the qualitative research itself which is the core factor of the indexing process, and therefore both the sophisticated and simple search strategy performance was relatively low and similar. SCOPUS or WoK performance was fairly good but not as efficient as CINAHL in terms of precision or odds estomator.

7.2.3.3 Economic Evaluation Siblings

Most of search strategies and databases either did not perform as expected or did not retrieve economics siblings at all (See Section 6.10.1) which resulted in a very low recall and precision. There was no clear winner among search strategies and databases as the retrieval performance varied among the seed studies (See Appendix Three). However, the overall recall suggested that the simple author-subject search strategy with or without either costs or economics filters were the best performing search strategy, though the IDEATeL seed study was the study with the most economics siblings retrieved (See Section 6.2 and Appendix Three, Table 9). Both citation and PubMed related articles searches perform badly. Overall, PubMed when combined with simple author-subject search can perform well in retrieving economics sibling. Both overall precision and overall odds estimator value suggested the superiority of the simple subject search strategy when combined with either economics or costs search filter, and the sophisticated author-subject with economics filter to retrieve economics siblings (See Section 6.10.1 and Appendix Three, Tables 7 and 8).

On the whole, subject search (the simple version) with specialised filters provided a good performance the results suggest the economics siblings are even more difficult to identify than qualitative siblings. But the influence of clinical area should not be neglected.

The choice of database is not clear either, however, CINAHL and SCOPUS with simple subject search combined with either costs or economics filters and PubMed sophisticated author-subject search with economics filter are the best databases to search for this type of siblings.

The CRD's guidance for undertaking reviews in health care recommended searching specialised databases such as NHS Economic Evaluation Database (NHS EED), as this database contains abstracts of full economics evaluation and lists the bibliographic details about partial evaluations. The other database to search is the Health Economics Evaluation Database (HEED) which contains either a full or partial summaries of economics evaluation. Both these databases are supposed to provide a full coverage of economics studies, however for a more realistic and updated retrieval additional sources such as MEDLINE should be searched, though the choice of database depends on the clinical topics as this research illustrated. Moreover, the NHS CRD (2009) emphasise the use of a combination of subject topics as well as relevant economics terms. This agrees with the results from this research as the results revealed that both simple and sophisticated author-subject search strategies perform well with the use of the economics terms. However, the PubMed simple (author-subject) search strategy provided the highest recall (See Appendix three, Table 9), the precision was very low and the odds estimator ranked the simple subject search combined with economics filter as the winner to retrieve economics siblings.

The Royle and Waugh (2003) study examined different databases' effectiveness in retrieving economics evaluation studies to make searching for technology assessment reports (TARs) more efficient by advising on an optimal economics studies retrieval strategy. In their study, they created a gold standard of recent - 20 TARs which make 424 studies in total – against which to compare retrieval performance and characteristics.

When Royle and Waugh (2003) analysed the characteristics of the economics studies, they found that among TAR studies, 14 reviews mentioned the use of search filters when searching for economics studies. A deeper analysis of the reviews revealed that the studies that contributed to making the TARs were classified into published studies (80%), abstracts (11.3%) and unpublished (8.7%). The TARs described the economics studies as having either RCT study design or non-RCT design. For searching databases, they recommended using a combination of suitable keywords to search title, abstract and subject heading as well as using the terms cost* or economics* in all of the above fields. MEDLINE and NHS EED were among the databases being assessed. The results revealed that MEDLINE was the best performing database retrieving the highest number of economics studies (published studies), whereas NHS EED only identified two extra studies which were not indexed in MEDLINE. This suggests that searching only MEDLINE can suffice for an optimal retrieval of the economics studies especially when time scale is tight. The analyses revealed that at least third of the included studies (in the gold standard and did not retrieved by MEDLINE) were either unpublished or grey literature, so subsequently if they removed these unpublished data then MEDLINE performance estimator will be better (Royle & Waugh, 2003).

The results from the Royle and Waugh (2003) agreed with the results from this research as the PubMed performance was the best among databases in terms of recall. The odds estimator recommended the use of search filters with suitable subject heading for CINAHL and SCOPUS. However the simple author-subject search recall indicates that this strategy is better with good odds of retrieving economics siblings. Perhaps with different combinations of databases and different search strategies it may be possible to retrieve this particular type of siblings, but the results may depend on the topic. Sassi, Archard & McDaid (2002) argued that economics studies retrieval tends to be challenging due to the lack of uniform conception about what can be considered economics research and this will affect the indexing process producing inconsistent indexes among databases (Sassi, Archard & McDaid, 2002). And this is might be the case in this research, as the recall was relatively low for all seed studies which might suggest that either some seed studies do not have published economics siblings or that due to indexing issues it was very difficult to identify some of the siblings. Another explanation might be that economics siblings need economics experts for the work so that they will be done by different research team even though they are siblings (based on an RCT) and therefore do not share author(s) (or many authors) from the seed, and therefore, the best possible option is to not use author names when searching for sibling studies, though another approach for sibling relationship identification should be developed.

7.2.3.4 Process Evaluation Siblings

The results here supported the argument made before. Process evaluation research sometimes can be confusing, as this type of research employs so much of the qualitative approach. It is difficult to disentangle the process evaluation from the purely qualitative studies (when designing a search strategy). Again there is no stable performance pattern to be deduced, different search strategies performed differently on different databases and different clinical area as well (See Sections 6.10.1), though the performance varies among seed studies reflecting clinical area effect. It appears that when time is limited the simple strategy seems sufficient, compared to the more sophisticated strategy.

PubMed simple author-subject search is the best choice for searching for process evaluation siblings. Other search strategies on PubMed – related articles search and sophisticated author-subject search – provided a good recall (see Appendix Three, Table 12). The precision of the sophisticated strategy is better than the simple strategy especially when combined with the qualitative filter (see Appendix Three, Table 11). However, WoK, SCOPUS and CINAHL databases provided a good performance with the simple search strategy with or without the qualitative filter (See Section 6.10.1). This suggests that databases with social science coverage and are most likely to have process evaluation studies in their collection. Federated search using E-library/Metalib database and simple author-subject search with qualitative search appeared to provide a good recall indicating that careful selection of databases on federated search might be beneficial for studies with different research design taking advantage of different indexes and subject representation between databases.

In general a simple author-subject search can provide a good recall. The qualitative search filter provided a good recall when combined with simple search strategy, though it did not outperform the simple author-subject search. But it provided a better precision than the simple author-subject search. This reflects the qualitative aspects of process evaluation siblings. Citation search overall performance was not good, while PubMed related articles average recall was better than citation search but low compared to the above author-subject search.

7.2.4 Sibling Studies Publication Time

Sibling studies publication time was investigated (See Section 6.8). The results showed no clear trend about the time when a sibling is published compared to the seed publication date. However it suggests that the main focus of the medical research is still on RCTs rather than other type of siblings (more than half of the siblings were RCTs and most of them were published after the seed study).

7.2.5 Search Strategies for Systematic Review

The recommendations of the Cochrane handbook (Chapter 6) focus on a comprehensive search for systematic reviews, they stated that:

> In order to identify as many relevant records as possible searches should comprise a combination of subject terms selected from the controlled vocabulary or thesaurus ('exploded' where appropriate) with a wide range of free-text terms.

The main issue here that the authors might not accurately describe their methods or objectives which in return will reflect on the indexing process and in the end the search will be influenced deeply (Higgins & Green, 2011). Therefore, comprehensive search should maximise the sensitivity of the search – maximising recall to ensure that there are very few missing but relevant studies. The yield is very high, and therefore they are not always practical for quicker searching.

Nevertheless, simple searching has recently started to receive more attention as such approaches may be quick, easy and inexpensive in time. In this research, the results suggested the potential of a simple search strategy (using a menu of three search terms from title or abstract) compared to the more complex, the sophisticated search strategy in retrieving sibling studies. On some, very rare occasions, the sophisticated/complex search provided a good yield but still, the overall retrieval performance suggested the superiority of the simple search strategies for efficient searching. Others have also concluded that a simple search strategy can provide good performance (e.g. Day et al., 2005; Royle & Waugh, 2005), (See Section 7.2.1). Evidence mapping rsearches is another example that proved the effictiveness of simple search strategy to produce an evidence map within the time and budgets restrains compared to HSSS search strategy (See Section 3.3.3).

The simple search strategy used here is based on key terms derived form the seed study title and/or abstract (See Section 4.3.). However, even with ATM, PubMed retrieval does not improve as expected. The initial search string is simple (short search string) but will be mapped into a more long search string mapping each search term on to its equivalent/appropriate indexed search term i.e MeSH term (See Appendix Two), so the search might not actually be a simple search in the end, but to the user it is still a simple search as the search entered a few key terms (See sections 3.5.3 and 7.6.1). This might mean the PubMed does not index some of the siblings (different research design). Furthermore, search filters appeared to work better on other databases; WoK and SCOPUS (for example), a finding that reflect PubMed's medical orientation which even the filter does not help to overcome.

Table 56 provides a comparison between SR and sibling studies:

Table 56: Comparison Between Systematic Review and Sibling Studies

Systematic review	**Sibling studies**
• RCTs only.	• Studies with different research design than RCTs beside the RCTs.
• Comprehensive and highly sensitive search strategy.	• Simple broad search strategy.
• Muliple databases search.	• Multiple databases search.
• PubMed could provide a high recall.	• PubMed not generally associated with high recall.

Sibling studies search seems to go more with evidence mapping (Section 3.3.3). With the use of search filters the search will be more focused on retrieving specific sibling types according to users' information needs, however, a broader search terms to target more then one research type in one clinical area might be more desirable (See Appendix one).

7.3 Clinical Area and Research Type

7.3.1 Overview: Main Findings

The seeds studies in this research were from clinical areas, but the confounding factor is that these seeds are different types of research as well. With some areas of clinical research, the emphasis is much more on biochemical/biomedical research – whereas for telemedicine/diabetes the questions are more around delivery. This might suggests that some siblings might be done by different author(s) on one hand. On the other hand, different research objectives and emphasis might affect the retrieval as well. This can be because of different databases coverage and interest, the IVF study, for example, is far more clinically oriented at this stage and accordingly it was not indexed in CINAHL and therefore it is unlikely to retrieve any siblings on the IVF topic unless its focus is on social aspect of the clinical area for instance. Another factor that should be considered is the indexing process of databases and the reporting of studies themselves.

This can be viewed as a computational linguistic problem, and it has emerged mainly from the different interpretations of study meanings and design. For example, Sassi, Archard and Mcdaid (2002) concluded that it was difficult to retrieve economics studies as there was no standard definition of what they are or how to interpret economics studies, which then affect the indexing process and therefore the retrieval process. Moreover, researchers (even economics experts) differ on how to interpret economics studies or inclusion criteria when considering a systematic review. And this might be the case with the other siblings types especially process evaluation type as there are no clear boundaries between process evaluation and RCT or qualitative studies. The different perception and interpretations of each type will influence the choice of indexing terms, search terms and consequently the retrieval and relevancy judgments. In other words, reporting, interpretation and relevancy judgments (knowledge, experience and cognitive state) are the core issue here, and normal as all retrieval based on human cognition, knowledge organisation and representation. The indexers cannot easily forecast what the future uses of a document might be.

The statistical tests (See Section 6.7) show a significant relationship between clinical area and search strategy and/or database performance which was clear from the performance differences among the seed studies. The results suggested that clinical area has the major influence on retrieval performance, as should be expected. The retrieval performance seems to depend heavily on reporting and indexing, and it was proved through the literature how the reporting, interpretation and terms used to describe the subject on indexing and retrieval profoundly affect document processing. Again it is up to users to decide on the relevancy, and in this research it was even more complicated to decide on sibling studies due to the lack of standards and guidelines to help to draw the decision as individuals differ on what means what or what is relevant or important. Moreover, a question is needed to be asked about the type of health professional involvement (other than medical involvement) that would be expected and this may affect the choice of databases searched. For example, there are specialist allied health professional databases such as OTseeker (for occupational therapy) and PEDRO (for physiotherapists) and these might need to be added to the list of databases to search.

It should be emphasised that existing literature (search strategy/database) focused on one clinical area, i.e. Day et al. (2005) focused on clinical trials in the field of musculoskeletal disorders and pain or one type of research design (RCT and Qualitative research), i.e. Glanville et al. (2006) for RCT searching and Evans (2002) and Shaw et al. (2004) for qualitative research searching and retrieval.

7.3.2 Authorship and Group Authorship

The IDEATeL seed study was the only seed study for which one of its authors responded when initially contacted, to provide a complete list of sibling studies. It was a very large project, with considerable emphasis on implementation questions, and lots of different aspects covered. It is not clear why it was possible to find all the direct siblings through a number of author names, but presumably this may have something to do with the strict organisation of contributions for publication. It appears that the IDEATeL team were the only people who fully understood the the objectives of this research and this might be because of their focus on health delivery.

There is emphasis now in many journals for declarations of authorial contribution to the paper (under ICMJE) but up to date there is no solid foundation of how to decide on authorship or contribution. The Journals tried to use a form asking each author to state or describe their contribution to the manuscript, however, the evidence showed inconsistency and unreliability from the responses, consequently, it was conclude that the decision about authorial contributions is not a straightforward one to make and it is not for the editors to make such decision, it is more likely a moral than technical issue. Marušić and Marušić (2010) suggested that asking individual authors the question “Why do you think you deserve to be the author of this manuscript?” might be the only option to solve this dilemma.

In this study, author names were combined with simple subject terms of seed studies since the initial assumption was that the sibling studies could share one of the authors of the seed study at least, for example, the Tamoxifen seed study was based on two RCTs, where both share authors. This strategy appeared to provide a good performance (See Section 7. 2).

Breast cancer with holistic therapy seed study, was chosen in order to examine the influence of having Chinese author names. Sibling retrieval was very low in this case, as all search strategies together only identified 8 direct siblings. This is might be due to the fact that either there are not many studies associated with this seed or to the nature of the authors' names. For example 'CHAN' as a surname is very common in China, therefore it was very difficult to precisely retrieve siblings based on authors' name, and the yield was very low though the retrieval is high. This seed study raises the problem of authorship, especially with the Chinese names as there is so many authors who share the same surname (in English versions of the Chinese name) which will create ambiguity and users might not have time to go through all name variations. Smalheiser and Torvik (2009) argued that the main source of ambiguity is when many different individuals have the same name, but there are others i.e. the variation of a single author name due to spelling error, change of name due marriage, using a pen name and finally spelling variants and multi-authored, multi-disciplinary and multi-institutional efforts. All this might lead to problems when works from specific author is needed.

The BMS seed study introduces an example of such ambiguity, being a work done by Chinese authors who are known to have thousands of individuals sharing the same name. This problem will affect precision values, author name ambiguity will result in a large retrieval with a large proportion of non-relevant items, and accordingly the precision values for this seed study were relatively lower than for other seed studies. WoK via WoS has an author finder functionality that might help to solve the problem of similar author names and locate article written by same author but this requires searching for each author individually (using Last name or initials) (Thomson Reuters, 2012). PubMed do ATM for authors in the same way it does for search terms (See Section 3.5.3). It uses different tables to match author names against. The Full authors translation table is the first to use, it includes full authors names for articles published from 2002. If the author is not found in the translation table and is not a single term, PubMed searches a full collaboration translation table, and if the author is not found then the author index is searched next and if the entire above tables do not produce any match then the collaboration index is searched. If all of the above fails, PubMed breaks down the term and redoes the search again in the same order (NLM, 2012a; NLM, 2012b).

Another issue that was addressed in this research was that of articles with corporate authors (IVF and CLD). The two seed studies were done by a group of authors, in these cases, there were two forms of searches to use, first; using individual authors' names, second; to use the group name (See Sections 6.5 and 6.6).

The IVF seed study was a collaborative work which was done by *European and Israeli Study Group on Highly Purified Menotropin versus Recombinant Follicle-Stimulating Hormone.* It was a multicentre RCT which accordingly is expected to have a number of other RCT siblings, however, the number of siblings identified was low (See Section 6.6). The performance pattern is different here from any other seed studies. With the IVF seed study, related articles search on PubMed was the winner retrieving the highest number of siblings and two uniquely identified siblings. However, the precision value was very low, but on the other hand a high likelihood of retrieving siblings rather than non-sibling studies as indicated by the odds estimator.

The CLD seed study was done by the Vermont Oxford Network (*a non-profit voluntary collaboration of health care professionals*) that aims to help newborns with chronic lung disease. The seed study is a quality improvement intervention based on each unit's better practice, so it is a multi approach assessment with no single RCT to be tested alone. Subsequently, multiple RCTs could be associated with it as it depends on recommendations from the units. Moreover, as a quality improvement intervention there is a possibility of having process evaluation siblings, even qualitative siblings as well, as this RCT involved families.

This reflects the fact that different databases handle this type of article differently, and that will affect the retrieval performance when relying on "author" name to help find siblings. It is reflecting the indexing differences between databases' policies when indexing an article with group name instead of using individual authors' names. These differences will affect the citation ratio associated with each article, as each time the article is being cited, it will be cited in two different ways and might result in losses during citation tracking. In this research, it was noticed that for the WoS citation search, and with the IVF seed study, that there was no citation for this seed when it was indexed under the group name, while there was 56 citations for the IVF seed study when it was indexed under authors' names. The same was noticed for the CLD seed study as well. This reflects findings by other researchers on the effect of group authors on the search retrieval and performance (Dickersin et al., 2002). PubMed help currently provides no advice on corporate authors.

Old MEDLINE and PubMed did not recognise group authors, which caused problems of course (NLM, 2011), although now PubMed MEDLINE is using group names in the same way as it uses individual author names which makes the problem easier, however the inconsistency in indexing is still a lingering problem (NLM, 2010). WoK retrieved several instances for the same record, it handles each instance as an independent retrieval when it come from a different database which indicates the original indexing mechanism the source database used, for example PubMed indexes the CLD and IVF seed studies under group name.

These problems brought into mind the problem of double author name such as the case with Spanish authors (that was not a problem in this study but it is worth mentioning). Usually Spanish people have two family names that make their surnames; first family name which comes from the father, and the second family name which comes from the mother. What happened is that the double family name is usually being misunderstood during the indexing, that is the second family name is considered as the family name while the first became a middle initial. MEDLINE became aware of such problem and consequently double family names are dealt with properly.

However, there is another problem with Spanish names, the first name this time, where composite names are common which again might be misunderstood and handled wrongly, in this scenario the last two names were treated as the family name although the first two names is a composite first name and the last name is the family name but MEDLINE has understood this situation and handled it correctly (Fernández & García, 2003). This may lead to inconsistency in indexing between different databases, a problem that was investigated and addressed, where inconsistency was recognised between MEDLINE, Science Citation Index (SCI) and Indice Médico Español (IME). In addition, the authors themselves are inconsistent in the way they cite their names on their work which will lead to inconsistency in indexing, and will affect databases retrievability (Ruiz-Pérez et al., 2002).

7. 3.3 Derivation of Gold Standard According to Authors

Authors of the seed studies were contacted to confirm on siblings list that was created for each seed study using relative recall. Three author of three seed studies responded. One author of the IDEATeL seed study was the fisrt and the only one to reply initially and provided a complete list of all publications on the IDEATeL project. Based on the IDEATeL publication list it was noticed that all publications shared at least one author and from that it was decided that shared authors is the criterion for direct siblings identification. After that, the siblings lists for the remining seed studies were created using relative recall. The authors of the other four seed studies were contacted again to confirm the list of siblings that was created. One author of the BMS seed study responded and provided a list of all BMS publications (Section 6.4.6). One author of CLD seed study processed the list of siblings list that was provided by the e-mail (Section 6.5.5).

The authors' varied response suggests that the conception of siblings relationship is still unclear for most of the researchers. For example the CLD seed study author stated that he is not aware what sibling studies are (despite being given an explanation). According to him the relationship with the RCT (seed study) is not clear (Section 6.5.5). Furthermore, the CLD was a large collaborative study and accordingly many different researchers may perceive relevancy in many different ways and may not appreciate or value other researchers from a different background. It seems that deciding what is a sibling study is a very subjective judgement. Again, this draws to the mind the problem of relevancy and relevancy judgement (See Section 3.6.2).

7. 3.4 Reference Lists

A complementary search was conducted by examining the reference list of each seed study to see if any sibling study which was missed by the main search strategy was cited on the reference list. The reference lists did not identify any additional direct siblings based on the inclusion criteria advised for direct sibling studies (Sections 4.8 and 4.9). Some of studies appeared in the reference list were either published before 1992 or not on the direct subject of the clinical area, for example for the Tamoxifen seed study there was a study by one of the main authors but its subject revolved around ovarian cancer, so it was not considered as direct sibling. In other words, some of studies in the reference list addressed general topics that are related to one aspect of seed study. Therefore the reference lists did not add any value for locating sibling studies.

7.4 Search Filters

Specialised search filters were examined to assess their performance in retrieving sibling studies. The overall performance of these filters did not add any value to siblings' overall retrieval in terms of recall, precision and odds estimator. However, at the individual level, seed study level and per sibling type, some of these filters performed well in terms of precision and odds restimator, i.e. economics filter for IDEATeL seed study and qualitative filter for Tamoxifen and BMS seed studies, (See Section 6.10, Table 51). Investigating the results based on sibling type, the results show that specialised search filters performed well for individual sibling types in terms of recall, precision and odds estimator. Although, in general, these search filters did not contribute much to the overall performance of search strategies, their contribution at the individual level, of type of sibling is valuable (as expected).

Coiera, Westbrook and Rogers (2008) stated that filters can help clinicians to get an answer quickly and appear to increase the rate at which a decision is made. Investigating the results it seems that for a particular type of siblings the use of specialised filter is useful and effective (Sections 7.2.3 & 7.4), however, for general retrieval of siblings of different types, the results degraded, confirming that simple search strategy with broad-based terms is more feasible (Pope, Royen & Baker, 2002).

Moreover, it was observed that these filters performed well with databases other than PubMed, even though they were designed for MEDLINE (Section 6.10, Table 51). This suggests that on one hand the MEDLINE indexing might not be as accurate as it should be for research designs other than randomised controlled trials. This combined with the lack of consistency in reporting qualitative research would mean that devising a Boolean search strategy (to search the title, abstract and descriptor (index term) fields) could not be expected to be wholly successful. On the other hand, other databases have a broader coverage for social science topics, i.e. SCOPUS which uses both MeSH and Emtree terms for indexing without restricting the number of index terms for records, although in SCOPUS the indexer only uses index terms that directly describe the topic to focus the retrieval process to relevant records as much as possible, but using both MeSH and Emtree for indexing may add more flexibility to the search.

Qualitative studies are problematic and difficult to retrieve, even though MEDLINE and CINAHL are using “qualitative research” (MeSH term) and “qualitative studies” respectively. CINAHL introduced this term earlier than MEDLINE reflecting the importance of this type of research for nurses (Evans, 2002; Noyes et al., 2008). However, the results of this research show that WoK performance was considerably better for qualitative studies compared to other databases, at both general and individual level (See Sections 6.10.1 and 7.2).

Checking was necessary when some of the seed studies did not appear to have certain types of sibling, as it is possible that the filters were excluding some relevant items. The Tamoxifen seed study, for example, only appeared to have RCTs, qualitative and process evaluation siblings with no economics siblings, but checking the references of siblings retrieved suggested that there were indeed no published economics siblings. NHS EED retrieved 6 relevant studies on the topic but further investigation revealed they were not direct siblings. For the BMS seed study, this seed study discussed a complementary therapy, it did not have a large number of siblings and accordingly the yield was low. Most of the siblings were retrieved using the qualitative filters and were mostly on WoK database (See Section 6.4).

The CLD seed study siblings were mostly retrieved by searching PubMed and WoK. The filters did not appear to provide a better performance than the simple search strategies (See Section 6.5). But for individual siblings type the qualitative filter did retrieve qualitative and process evaluation studies which show that the filters can retrieve specific sibling types if the focus is only on one specific type. However, the value of simple author subject is greater as it might retrieve more than one type of siblings in one search. Based on the sibling types retrieved, this seed study appears to be of interest of different health professionals and policy makers and therefore it was indexed in PubMed and WoK, two databases with different coverage and scope.

The retrieval trend that emerged in this research is that (mostly) the databases with social sciences coverage are more likely to retrieve most of the siblings associated with seed studies – WoK and SCOPUS for example – and that PubMed might complement the retrieval for seed studies that have aspects of health service research of interest to clinicians (such as IDEATeL, Tamoxifen and BMS). PubMed mainly contributed to the IVF and CLD seed study siblings retrieval, this might be because both these seed studies are dealing with clinical (and specialised) techniques of intervention delivery. In conclusion, the filters work well for specific sibling retrieval and accordingly the filters should be added if only certain types of siblings are required, with some consideration of the nature and the interest of the clinical area, for choice of databases to use (Sampson & McGowan, 2011).

7.5 Multiple Database(s) Search

It is to be expected that neither one single search strategy nor single database will perform perfectly due to indexing inconsistencies and authors' choice of words to describe their methods and results. The results suggest that searching a single database might limit identification and retrieval of existing literature. This is consistent with recommendation made by researchers and meta-analysts (mainly targeting the RCTs), as in order to implement a comprehensive search more than one source is needed (Lemeshow et al., 2005; Papaioannou et al., 2009; Parkhill et al. 2011) and it is not enough to search only MEDLINE (Avenell, Handoll & Grant, 2001; Royle & Waugh, 2005). Accordingly, based on the results, it seems necessary to advise a combination of search strategies and databases for a more comprehensive search (See Section 6.9). For example, the Tamoxifen seed study a combination of simple author-subject search on SCOPUS and WoK, will yield 92% of the siblings and can be increased to 99% if simple author-subject on CINAHL, PubMed and E-Library were added to the previous two databases (See Section 6.9, Table 49).

Naturally, searching multiple databases might improve recall, however, precision will be sacrificed. So far, the precision value was considerably acceptable, and the fact that studies identified by multiple databases are more likely to be siblings than studies identified by one database, therefore, it would be expected that the incremental yield from the databases will result in reduction of precision. Also, when considering the value of some databases in retrieval of items not retrieved from other databases, it must be remembered that not all uniquely identified studies may be siblings and therefore this will enhance recall slightly but affect precision drastically (See Sections 6.9, Table 49 & 6.10, Table 51). However, for the IDEATeL and the Tamoxifen seed studies the precision was better than the other seed studies.

MEDLINE is said to have the most discriminating power compared to the indexing of several other biomedical databases, (except, perhaps, CINAHL in terms of qualitative research), and therefore, a PubMed search is expected to offer the most precise retrieval, due to MeSH and automatic mapping features (Sampson, 2009), however this does not seem to be the case here. This might be because PubMed is known to have randomised controlled trials publication type for reports of such trials, and so it is expected to be precise if the search is targeting RCTs only. However for studies of other research types the situation might and will be different, and consequently, even with the automatic mapping feature, it was difficult to retrieve those studies. MEDLINE indexers use the most precise term in the hierarchy as much as possible but some misunderstandings might occur causing the lack of suitable MeSH terms, made worse by the lack of appropriate or inconsistent reporting of the study in the abstract, as the major searchable field. If there is no appropriate MeSH term, or the text term that searchers enter differs from that used by the authors in the abstract, the yield might be high at the expense of precision.

In the end, the number of siblings retrieved for each seed study varies. For some of the seeds this number was very low, i.e. 8 siblings for the BMS and 15 for the CLD, indicating either the difficulty of locating sibling studies (although further checks were made on the reference lists) or that there is in fact a very low number of siblings. For example, Lewin et al. (2009) and Glenton et al. (2011) found fewer qualitative studies than they expected (See Section 2.4), with less than a third of the trials having associated qualitative research, and it must be emphasised that Lewin et al. (2009) and Glenton et al. (2011) both were searching in an area of health services research (Cochrane Effective Practice and Organisation of Care register), where such studies would be expected. Finding siblings for all the seed studies seems to suggest that the search strategies (or a combination of them) are effective. The difficulty of locating siblings is because of different research design for each type of sibling, and also, different databases with different coverage will have different interests.

In other word, qualitative databases such as CINAHL will mainly focus on qualitative research and will consequently its indexing scheme will be more accurate in terms of qualitative research indexing. PubMed is considered to be more RCTs oriented. In the conclusion, all this mean that because the siblings have differences, combining different databases (different coverage and orientation) will help to overcome these differences.

7.6 Databases Selection Implication on Siblings Retrieval

Databases selection was pragmatic, meaning that the difference in databases characteristics databases different coverage and indexing scheme and terms - might influence the search performance and hence enhance the retrieval results. Each one of the selected has pros and cons that affect the retrieval performance positively and negatively at the same time. Table 57 demonstrate some of these databases pros and cons from sibling studies perspective.

Specialised databases might have some potential for locationg specific types of siblings such as economics databases NHS EED and HEED. However, for comprehensive retrieval non specialised databases with social scienece coverage are more suitable. Both NHS EED and HEED are known for their strict and selective inclusion of peer reviewed economics studies and therefore locating siblings for specific RCTs might be difficult if not impossible. For example, in this research NHS EED was searched for economics siblings but no economics siblings were identified, while other non economics databases contributed some economic siblings. Other databases such as EMBASE might be considered in retrieving sibling studies. Embase is another biomedical database that covers different clinical topics. However, it does not have social science coverage and therefore it might not contribute very much to sibling retrieval.

Table 57: Some of the Selected Databases Pros and Cons

PubMed	
Pros:	*Cons:*
– Free access database.	– Biomedical databases, the social science in minority.
– No limit of search terms	– The main focus is randomised controlled trials.
– Link to all online journals available	– Other type of studies is marginalised.
– Special features; PubMed related articles and PubMed ATM	– No citation analysis.
– A variety of limit option	
– Export search results into reference management software.	
Federated search	
Pros:	*Cons:*
– The scope depends on the selected databases.	– Advances search features are lost
– Allow searching across multiple resources (one stop search).	– Database specific features are lost.
– Allow the user to select which databases to search.	– Databases selection is subject to organisational choice.
– Benefit from databases deferences (coverage and indexing).	– Retrieval redundancy.
– Users do not need to be familiar with different databases interfaces.	– No clear relevancy ranking (too many variables for reliable rank).
	– Performance issues can occur if the federator waits for the slowest remote search engine to respond.

(Continuous)

Table 57: (Cointinued)

SCOPUS	
– The scope is life and health sciences, hard sciences, social sciences, and earth and agricultural sciences.	– Needs subscription.
– Citation analysis. full cited reference information that are searchable.	– Allow up to 30 keywords per search.
– Links to full articles (if available).	
– Beside journals it indexes books; conferences proceedings, web sites and patents.	
– Use both MeSH terms and Emtree.	

7.7 Simple Search *vs* Sophisticated Search – With Reference to IDEATeL

It was necessary to compare simple search strategy retrieval effectiveness with sophisticated search strategy effectiveness. Sophisticated search strategies were conducted using the PubMed database only. The results show that simple search strategies on different databases often performed better then the sophisticated search (See Sections 6.10 and 7.2). At clinical area level the sophisticated search strategies gave a good performance, but still it did not outperform the simple search strategies. Sophisticated search strategy performed well for individual types of siblings (See Section 7.2.3) and even better when it was combined with the specialised filter for that type of siblings, i.e. qualitative filter for qualitative researches.

The highly sensitive search strategy (HSSS) is the main search strategy recommended to use for systematic review preparation and was developed mainly for MEDLINE database and was designed to retrieve controlled trial studies using all possible headings and terms which might be used in the study reporting to achieve the maximum sensitivity possible (Robinson & Dickersin, 2002). The HSSS received great attention and was adopted as the main search strategy and HSSS will be combined with subject specific terms to use to prepare a systematic review. When Sampson et al. (2006c) tested HSSS, they noted that known item searching of MEDLINE only provided 72% of included studies and it was therefore recommended for comprehensive searching to use other databases besides MEDLINE.

Another study compared HSSS different phases performance, the study found that even HSSS $_{123}$ will achieve the highest sensitivity but the reviewing time is 1086 hours, another variation to $HSSS_{12}$ by adding the text word "*versus*" can balance precision and recall and minimised reviewing time to be 823 hours (Zhang, Ajiferuke & Sampson, 2006). The decision between either HSSS strategy will often depend on the possible influence of missing items on systematic review conclusions. However these highly sensitive strategies may produce too large an output to scan easily and effectively and recently there has been a trend toward simple search strategies (See Section 7.2).

In this research, a simple search approach provided a reasonable performance (author-subject). There are, however, differences in performance from different databases, and therefore it is recommended that multiple databases are searched for a more optimal performance (See Section 7.5). Sampson and Mcgowan (2011) states that there might be situations where some items cannot be retrieved unless using a broad search and sibling retrieval seems to be one of these situations, as there is no known relation to link them all together aside from the subject of the seed study. As mentioned before, the filters with simple subject search (no authors were used with the search) performed well to retrieve specified type of siblings, but it did miss some of the siblings and sometimes it retrieved siblings of different types.

The IDEATeL research project could be described as programme evaluation as the research was undertaken by team of specialists, who worked to cover all the aspects of the project; barriers to overcome (attitudes, satisfaction); feasibility (social/technical/economic); implementation feasibility; and costs, in order to assess what worked, why and how telemedicine could benefit particular groups of patients. Being a project done by a large team of specialists, it was known to have many siblings. Contacting authors, through the corresponding author, provided a list of all published IDEATeL work.

This was studied thoroughly in order to deduce a theme, and determine which of the publications were genuine siblings, to generate a “gold standard” in the conventional sense. Some of the publications on the list were more concerned with the validation of particular research methods used. These were excluded from the final gold standard list used for evaluation. From this list it was obvious that all publications shared at least one author from the seed study (and from this it was decided to make this as the main inclusion criterion to decide on sibling relations for the other seed studies). This evaluation was of course carried out after the retrieval operations for the seed studies.

WoK and SCOPUS retrieved most of the siblings, where both databases overlapped yet each retrieved unique siblings. However the best combination of databases to retrieve most of the siblings is WoK and PubMed with simple author-subject search, they appeared to complement each other changing the recall, precision and odds estimator to (81%, 71%, 232.58) respectively (See Sections 6.2; 6.9 and 7.2).

PubMed database is a free database and one of the main sources to search for medical data and consequently it was used as a case study to compare simple and sophisticated search strategies against. In the sophisticated search strategies, MeSH terms were used, although the PubMed has an automatic mapping technique and accordingly it is expected to retrieve more records, however it did not perform as expected (See Sections 3.5.3 and 7.6).

The comparison between simple search and sophisticated search strategies on PubMed, showed that sophisticated search strategies, with or without search filters did not seem to be able to provide a comprehensive retrieval of siblings, at least on PubMed, and therefore, the choice between the sophisticated and simple search (on other databases) is for the researcher to make, for example WoK provided a better performance, (See Sections 6.2, Table 30). Moreover, PubMed simple search identified four unique siblings while the sophisticated strategy did not retrieve any unique siblings (See Section 6.9.1, Table 50).

In conclusion,and compared to other databases, the sophisticated search on PubMed does not performwell enough to be considered a better choice than the simple search strategy to retrieve sibling studies as the added value was not large enough to be of significance.

The comparison of the items retrieved against the list of publications provided by the IDEATeL team confirmed that book chapters, conference proceedings and abstracts were the most difficult to retrieve as these were mostly not in MEDLINE (or if present, not a relevant sibling) and many were not in Web of Knowledge either (See Section 6.2.5). Visual inspection of the journal item titles against the titles of the conference papers suggested that for IDEATeL, most of the conference papers on the research work contributed to later journal articles. Therefore, for a comprehensive search, the E-Library/MetaLib set of databases may need to include a database that indexes conference proceedings, as well as databases of theses and dissertations and therefore, another simple author-subject was conducted with a different selection of databases (See Section 4.3).

However, even with the new database selection, those siblings were still not identified. Those siblings which were preliminary feasibility reports and technical background research and which can be considered as process evaluation studies proved to be more difficult to retrieve and all appeared in conference proceedings only. Unless such conference papers reappeared in a journal article, they may not appear in bibliographic databases which cover biomedical research. This implies that works contributed in proceedings receive less attention and less credits from databases than journal articles. For IDEATeL this was not a problem (except perhaps in timing, as the journal articles followed the conference papers) (Section 6.2.5).

The same was noticed from BMS seed study list that was provided by one author. Most of the missed studies were book chapters or conference papers (Section 6.4.6). This support the previous arguments about how works contributed in proceedings receive less attention and less credits from databases than journal articles and therefore the chances of retrieving theses papers are very slim.

7.7.1 PubMed Automatic Term Mapping (ATM)

PubMed ATM is supposed to help in retrieval process however, compared to other databases' performance – SCOPUS, CINAHL and WoK – the PubMed performance did not seem to have been improved by PubMed ATM, leaving room for debate. These results can be either due to clinical area characteristics or that PubMed does not index all the expected sibling studies or at least are indexed differently specially the qualitative, economics and process evaluation siblings. However, PubMed can complement other databases retrieval providing some unique retrieval, and if only free access databases are available, then PubMed can be useful.

For both the simple author-subject and sophisticated author-subject search PubMed automatically mapped search terms into the appropriate search terms for better performance based on the database indexed terms (See Appendix Two). Comparing both translation of simple and sophisticated revealed that simple search query did not miss many of the search terms after translation, making both simple and sophisticated almost compatible in terms of search terms. For some of the seed studies, the sophisticated search strategy seems to add more restrictive retrieval when adding more MeSH terms (e.g. the CLD and IVF seed studies), while in other cases it appeared that the sophisticated strategy exploded the search resulting in higher (but still on topic) retrieval outputs (such as the BMS and Tamoxifen seed studies). The sophisticated search narrows down the retrieval by adding more terms (more terms for mapping mean more specified MeSH) and therefore a more precise performance than the simple search.

From the translation query, it appeared that beside the appropriate MeSH terms, PubMed tries all term variations to search all fields in the database, which might add some ambiguity to the search and hence some noise to the retrieval results, especially with filters translation. For example, with the economics filter the terms "costamerogenesis [Title/Abstract] OR costamers [Title/Abstract] OR costami [Title/Abstract] OR costamp [Title/Abstract]" were used.

In conclusion, PubMed ATM does not appear to add much value to the retrieval process for sibling studies retrieval. Obviously, the terminology is changing continuously making searching for studies a difficult one to generalise. In order to accommodate these changes, descriptors have to be added, changed or deleted from MeSH with adjustments in the related hierarchies. For example, PubMed added 302 new MeSH headings, changed 26 MeSH headings and deleted 30 MeSH headings in year 2013 (NLM, 2012c).

7.8 Text Mining

Recently there is an increased awareness of text mining techniques ability to retrieve hidden knowledge and discover possible associations and patterns in texts. Systematic reviews have their established standards and steps in doing the review which are up to date still manually performed. The main issue with systematic review is the quantity of potentially relevant literature that can become unmanageable and therefore becomes a burden for systematic reviewers. Subsequently text mining can help in searching, screening and synthesising in systematic review (Section 3.7).

Text mining techniques were investigated in this research (the searching and retrieval functionality). As aforementioned, systematic reviews requires time and efforts to complete. Up to date, text mining can not fully support SRs. However it can provide a semi-automated support and therefore the review can be completed more quickly and maybe more systematically as more data can processed and summarised. Equally important, searching, screening and synthesising can be more customised, focusing on pertinent terms, retrieving relevant documents and synthesising prominent information.

Unfortunately, text mining is still limited to MEDLINE only. Although MEDLINE is one of the leading biomedical databases, this limits search options when conducting systematic reviews. It is well established that searching MEDLINE only cannot be considered sufficient for comprehensive retrieval which was demonstrated in research as well (Section 7.5). Therefore, extending text mining to other databases is required; especially as the health services field has a social element that is covered by other databases that have more coverage in social sciences such as SCOPUS or WoK. Additionally, biomedical literature is multidisciplinary and accordingly text mining tools need to be flexible to handle the different formats of information that are available. Moreover, it will be more useful to extend current TM tools to full text rather than abstract only.

In all text mining opportunities for the application of text mining in systematic reviewing and in the social sciences in general are growing. Text mining has the potential to help systematic reviewer in their job. Text mining researchers aim to apply TM tools more widely in social science. Moreover, multiple databases should be taken into consideration for search comprehensiveness, to reduce bias and provide for robust results.

7.9 Indirect Sibling Studies Retrieval

The previous section explored and explained direct sibling studies retrieval and identification. It showed that simple author-subject search strategy across multiple databases is feasible and can produce a good retrieval performance. However, some researchers might find indirect siblings of as much interest as the direct sibling studies.

Comparing search strategies performance between the two search phases (direct and indirect siblings retrieval – phase three and direct siblings retrieval – phase four) showed the same performance pattern detected and investigated in direct sibling studies retrieval (See Sections 5.4.6 and 6.10). In other words, the simple author-subject search strategy on multiple databases i.e. WoK and SCOPUS can produce a good retrieval performance in terms of recall and precision. Again, retrieval performance depends on clinical area.

Some indirect siblings appear to be done by authors or research teams other than the seed study research team (if the association is explicitly stated then it will be a direct sibling) and therefore complementary search strategies might be able to locate economics sibling studies as well as qualitative and process evaluation siblings. However, this did not emerge in this study. Complementary searches such as reference list consultation did not add unique indirect or direct sibling studies to the retrieval list. Therefore such additional strategies in addition to the citation search strategy are not considered to be of value for sibling studies retrieval (both direct and indirect).

Chapter Eight
Conclusion

8.1 Introduction

The previous chapter discussed the performance of search strategies and databases in retrieving sibling studies. It provided possible explanations for some of the results and discussed the main issues and challenges that emerged in this research and how those might have affected the retrieval performance.

This chapter revisits the main findings of the research, in relation to the objectives. It discusses the limitations of this research which might be considered as barriers to implementation of some of the ideas proposed. In addition, the generalisability issue is discussed in terms of the applicability of the findings outside this research and the topics considered. Further research and development are advised based on the results, challenges and limitations of this research.

This thesis set out to address the following overall research question:

Is there efficient search strategy(s) to retrieve qualitative, economic, and/or process studies that may be associated directly with the seed RCT?

Chapter One set the scene, providing a brief overview of the current state of the art in information retrieval for systematic reviews. It described the motivation, the research problem, the relationship with the Cochrane IR methods group and the objective of this thesis.

From the above main research question, more specific questions emerged after further consideration of the literature and examination of the role of each type of sibling. Each chapter of this thesis provided some answers to the sub-questions that emerged from the main question.

Chapter Two discussed sibling studies by analysing several studies of each type of siblings in order to explore the main characteristics of each type. Accordingly it helped to answer the first sub-question:

1) *How can sibling studies be identified? Are there common characteristics that make the studies siblings?*

Each type of sibling can contribute some additional value to an RCT. Economics siblings set out to assess intervention effectiveness and cost-effectiveness. Qualitative research helps to explore factors that affect the intervention outcomes while process evaluation explore how the intervention was implemented and delivered and what factors attributed to the outcomes. Chapter 2 established that each type of sibling has different aims and therefore different research designs to realise these aims. Accordingly, a relationship is a very hard one to recognise unless explicitly implied or stated.

It was difficult, if not impossible, to decide on ways of verifying the sibling relationship. The IDEATeL seed study was the only seed study associated with a complete list of siblings provided by one of its authors at, at the early stage of this research, and therefore it was used as a case study to help to deduce the ways a sibling relationship could be verified. In the end, the only clear option was to use authors as the basis of the relationship. This assumes that a sibling study is likely to share one of its authors with the RCT seed study. This worked for IDEATeL but might not work for other seed studies. Other seed studies authors were contacted, only one author of BMS seed study and another author of the CLD seed study responded. The BMS author provided a list of all BMS publication where all of the identified siblings appeared in the provided list which supported the intial claim that was derived from the IDEATeL seed study list of publications. The rest of listed studies did not match the inclusion criteria (See Section 6.4.6), while the CLD author only processed the provided list of siblings (See Section 6.5.5).

Chapter Three addressed the following three questions:

2) *Are subject searching, author searching, related articles and citation searching search strategies effective in retrieving sibling studies?*

3) *Which database is considered to be more productive and comprehensive and which provides more unique or reliable studies within specific time frame?*

4) *How best to measure retrieval performance and effectiveness of both search strategies and databases?*

In this chapter user information needs, the indexing process and the interactive implications on retrieval performance, user satisfaction and expectations were discussed. Several search approaches were introduced and discussed, including federated search. In this chapter, debates about search strategies and databases retrieval performance were discussed in order to investigate what has been done in the field and the extent of contributions made. The literature review on IR performance measurements demonstrated how relevancy judgements govern performance efficiency measures. Other tools, to complement searching or to make searching easier, such as text mining were explored. The potential usefulness of the clinical trial registry number was noted. Such a link would be useful for identification of sibling studies, but there are many registries and little clear incentive or mandates to cite trial numbers.

Chapter Four readdressed the fourth sub-question in more detail to help decisions on measurement metrics to be used in this research. This chapter introduced the methodological aspects of information retrieval, how research has been conducted and the paradigms that exist in information retrieval research. It appears that much information retrieval research follows a quantitative approach using recall and precision which depends on human judgements of relevance, and there is considerable debate about the difficulty of assessing relevance – it is inevitably subjective. Federated searching is another issue that received attention due to the large collections involved and therefore the complications such searching adds to measurement metrics, when probabilistic searching means decisions need to be made about cutoff points for retrieval outputs. The potential number of items that could be retrieved is huge. Ranking algorithms are essential for IR systems efficiency, however information about them may not be transparent enough to understand IR system performance.

Chapter Five, provided the analysis of relevant "on topic" and indirect siblings. It provided a brief summary of the pilot study (Section 5.2) and how it helped to plan the second stage. In the **Second** retrieval stage (Section 5.3), which involved relevant "on topic" retrieval performance assessment studies, the result was a pooled list of relevant items which was ready for further examination for direct and indirect siblings identification. Performance analysis examined the top performing search strategies and databases for each seed study, to check whether there were any patterns, although it was not altogether surprising that performance was very variable. The **Third** stage (Section 5.4) of this research was identification of indirect siblings, using specific inclusion criteria (See Section 4.8). At this stage the best search strategies and databases were provided as some researchers might be interested in indirect siblings as much as they are interested in the direct siblings.

Moreover, because of the multidisciplinary nature of siblings it was imperative to search databases other than PubMed as the social perspective of some type of siblings (e.g. qualitative studies), to utilise indexing variation between databases which might affect the retrieval performance and therefore, sibling retrieval can be optimised taking into account those two factors. Related article searching was tested for its capability to complement subject (or author/subject) search, other research indicated the possibility of combining different type of searches, i.e. Boolean and similarity search (See Sections 3.5.1 & 3.5.2)

Chapter Six, focused on the retrieval of direct siblings. Performance analysis for all search strategies and databases was performed, and the best search strategies and databases were recognised. This stage considered the core of this research and accordingly, more analysis was carried out. The idea of the unique contribution of search strategies and/or databases was introduced, and the publication time for sibling studies was investigated to check on assumed time associations between the seed study and its siblings. Furthermore, multiple databases performance analysis was conducted.

Chapter Seven discussed the results of this research, and compared it against the existing evidence if available. It provided some possible explanations to the results and explored the factors that might have influenced the retrieval behaviour. In this chapter sub-questions five, six and seven were answered;

5) *How often is the sibling published before/simultaneously/after the seed study*?

6) *Does the clinical area affect the retrieval performance of search strategy or database?*

7) *Is there any pattern or information to associate the seed studies and its siblings, i.e. clinical trial number?*

The research questions were readdressed one more time in Chapter Eight.

8.2 Main Findings - Summary

The key message for information specialist is that simple author-subject search appeared to provide a good recall for sibling studies retrieval. Search filters provided a good retrieval recall in some instances (Section 7.4), i.e. qualitative siblings, so it performed well when specifically qualitative siblings are required (Section 7.2.3.2). Related articles search on PubMed did not contribute much to identifying siblings studies and therefore is a worthwhile search strategy for sibling identification, however it can be combined with the simple search strategy to complement search retrieval since it did provide a good performance for some seed studies (Section 6.10). Citing search is not effective in retrieving sibling studies. In the end, relative performance of the simple search strategy seems to be stable across the five seed studies (in terms of providing a good recall), however the performance varied across databases depending on database coverage.

On the whole, relative performance of search strategies and databases was fairly stable across phase three (direct and indirect siblings retrieval) and phase four (direct siblings retrieval). There are several plausible perspectives on recall and precision, and the results indicated which strategies/databases could be used if recall was the main objective, and which could be used if precision was the main objective.

The low recall called for testing different combination of databases for each seed study (different clinical area). The recall was maximised but on the account of precision (Sections 6.9 & 7.5). Simple author subject search on WoK and PubMed and PubMed related articles combination provided a good start point to boost recall but not as required. SCOPUS might add value to retrieval performance as well (e.g. Tamoxifen). Both the CINAHL and E-Library search helped to boost the recall for some seed studies (Section 6.9, Table 49). Again, the clinical area influence performance and some professional judgments required.

In conclusion, for clinicians wanting some ideas about subject searching and with a tight time limit WoK and SCOPUS can provide an optimal retrieval performance.

8.3 Research Limitations

Simple search strategies were the main focus of this research, but it has to be acknowledged that PubMed searching with automatic term mapping "helps" to make a simple search more sensitive. The research could not evaluate the effectiveness of automatic term mapping itself, merely report that with author-subject searching it seems to work well. Federated searching assumes that the search strategy is not likely to be complex, but decisions have to be made about the cutoff point for performance calculations. This was a pragmatic judgement, based on observations about the steep drop off in retrieval of relevant "on topic" items beyond a particular point. More research seems to be necessary on how well federated searching works. As noted in the research, the choice of databases probably has to be pragmatic, based on the topic and the likelier databases where relevant research might be found.

The choice of PubMed was a pragmatic choice. First, PubMed is freely available. Second, Aberystwyth University does not subscribe to other versions of MEDLINE. The effect of database interface could not therefore be investigated. Arrangements had to be made to search SCOPUS and CINAHL at another university. This limited flexibility and the ability to follow up later queries.

The IR performance efficiency was measured by recall, precision and odds estimator metrics. Odds estimator was a useful measurement for comparing the likelihood, the chances, of a search strategy/database retrieving siblings rather than non-sibling studies. Comparisons of the odds estimators can provide a way of taking recall and precision into account. The number of studies in the database which are not siblings and not retrieved (D) would be calculated for each search strategy/database separately, however, it impossible to estimate the D value accurately, therefore, a pooled number of non-relevant items and retrieved by all search strategies/databases after subtracting the number of relevant/sibling studies which was retrieved by specific search strategy/database was substituted for the D value (Section 4.5). In the end, it is a mean to provide a rank for search strategies and databases based on the other search strategies and databases retrieval performance. The pooled sibling lists were created based on relative recall concept (Section 4.4) as it was difficult to create a gold standard using a conventional way. The time period used in this research (10 year before and after) limited using a conventional gold standard. Furthermore, when a division by zero problem arose, the zero was substituted by 1 as a neutral value to overcome such mathematical problems. However, this is a mathematical manipulation that might seem a bit suspect. Furthermore, authors were contacted in order to verify the list of siblings that was created. According to their response it was obvious that deciding on siblings relationship is a very subjective judgement and accordingly the decision on siblings is not clear cut (Sections 6.5.5; 7.3.3).

The screening process used might have some implications for retrieval efficiency calculations as the judgements for relevancy were made based on titles and abstracts only, eliminating the records when the abstract was unobtainable. Therefore it was not possible to assess all retrieved records, leaving a possibility that some siblings were missed. Moreover, the inclusion criteria used here – deciding primarily on a sibling relationship based on sharing at least one author – was made based on a theme driven from one seed study gold standard (IDEATeL). This assumption may not be valid if there are siblings with no shared authors with the seed study. On further checking, it appears that there are not many direct siblings for the seed studies. One might think that there are other studies which do not share author(s), which is actually expected, however the link is still a weak one to recognise.

The selection of the sample may have implications for generalisability, yet selection criteria ensured that the five seed studies topics were clinically important. Each one of the five seed studies was chosen with care, and with advice from the Cochrane IR methods retrieval group to represent topics of current importance. In addition, two of the seed studies were from collaborative clinical networks and/or multi-centre trials, which could pose implications for author/corporate author name searching.

Simple search strategies on SCOPUS and WoK showed consistent recall across four clinical areas which can be considered a very positive sign that these results may be generally applicable, however, the effect of clinical area on search/database(s) performance cannot be neglected (as indicated in the very variable results for related article searching).

8.4 Contribution to Knowledge

It was statistically challenging to conclude which search strategy or database could consider being more efficient in retrieving sibling studies. Previous studies focused on retrieval performance for a particular clinical area, however this thesis targeted five different clinical areas (each one has different characteristics and target population) and the retrieval of sibling studies associated with it.

Retrieval and identification of siblings was suspected and demonstrated to be problematic. First, the link between the seed study and its siblings is not clear, but common author(s) is a way to link siblings together, and perhaps guidelines for authorship contributions (ICMJE), and transparency in reporting mean that shared authors will be more common in future than the literature review indicated. Second, the retrieval performance varied among databases depending on database coverage and clinical area of the seed study, which was expected, and accordingly there is no winnng search strategy or database. Previous similar studies have often been done on one particular clinical topic or area. This research shows that much depends on the clinical topic chosen.

The simple combination of terms appears to work effectively assuming more than one database is searched, with appropriate and probably pragmatic choice of databases. Odds estimator calculations can help in decisions about which database to consider for sibling studies retrieval by assighning performance ranks to the database considering other databases retrieval performance. The results shows that PubMed related articles search and Boolean search (simple form) can provide an optimal retrieval of siblings, especially since related articles search retrieved at least one unique sibling for all seed studies. PubMed automatically maps search terms to thesaurus terms, however the retrieval does not appear to be much different between simple search and more complex combinations of thesaurus terms.

On the whole, the clinical area and topic influenced the retrieval performance and should be taken into consideration when selecting database(s) to search, especially when E-library/MetaLib is used. Consideration of the type of studies required is important as well, for example when qualitative research is the focus, then CINAHL is the database to use.

Apparently, sibling study retrieval depends on the reporting, indexing and database coverage. For instance, qualitative research is challenging to identify in many electronic databases, and this might be due to the lack of suitable terms which clearly describe the research in titles or abstracts, reflecting the problems of reporting of qualitative research, which in turn affect the indexing consistency in databases. The results in this research suggests that this type of research, as well as process evaluation, still receive less attention than they need, although there is increased awareness of its importance. And until we have a better reporting and indexing of sibling studies an effective generic search might not be feasible yet.

Moreover, information professionals might consider combining similarity search and the simple Boolean search on multiple databases. This approach appeared to be efficient in sibling retrieval, considering the different research design for all siblings. Accordingly, with a more unified studies (single research design) this approach might efficiently retrieve relevant studies for SRs. Searching a combination of different databases (different coverage, i.e. WoK and PubMed) appears to ease this problem taking into account different indexing approaches and therefore increasing the recall (Sections 6.9 & 7.5).

Interestingly, PubMed search filter works for both WoK and SCOPUS (See Section 7.4), however after careful investigation, it was confirmed that SCOPUS uses MeSH terms as entry terms as PubMed while WoK links to PubMed (beside other resources) when searching for studies which means that some aspects of PubMed are inherited in WoK.

8.5 Future Work and Recommendations

It was interesting to find the potential of simple search strategies in retrieving sibling studies, as it outperformed sophisticated search strategies with a reasonable recall and precision. Related articles on PubMed and citation search were initially considered as potential search strategies which should perform well for sibling retrieval. However, this was not the case in this research. Related articles appeared to be popular between PubMed searchers, however there is little performance evaluation of this strategy, and it has not been widely adopted for systematic review searching. Perhaps in the case of systematic review updating it can be useful, but with siblings it was efficient with two seed studies only (Section 6.10). It seems that the PubMed related article algorithm cannot effectively detect the relationship from the text itself. PubMed related article searching should receive more attention from researchers to evaluate its performance and contribution to systematic reviews.

The PubMed related article performance in this research might reflect on studies reporting issue. The reporting quality might not reflect underlying methods or data and therefore, inadequate reporting of important aspects of methodological issues might affect the indexing, and adversely affect the PubMed related article searching performance for sibling retrieval. There should be a guideline to control sibling studies reporting such as

CONSORT guidelines. PubMed ATM should receive more attention – up to now there is no published large scale evaluation for PubMed ATM.

Moreover, sibling study identification and retrieval needs further consideration. It is assumed that siblings complement and supplement randomised controlled trials in an important way, however, in this research retrieval and assured identification of sibling studies proved to be problematic. Therefore it is really important for researchers, publishers and information professionals to collaborate to develop guidelines to group sibling studies. Trials register number can be considered as an option so that all siblings should have the same register number. However, not all trials have a trials register number at the moment, as it is still not obligatory to register trials, but this should be made a formal requirement for all trials and from that point it can be used in reports for the other siblings directly related to that trial. There is still the problem of different trial registries, of course. I argued that the sibling relationship should be made clear, explicitly stating which seed study RCT they are based on. Therefore a database administrator might use a tree like representation when indexing the seed study where a branch will be added for every sibling (published before, after or on the same year) even if not indexed in the database (its existence should be made clear).

At an early stage of this thesis, I did an experimental Google Scholar search for sibling retrieval, but at that point it did not perform well. Moreover, there is no download manager or save option for research purposes, and it is impractical to explore the retrieval set using Google Scholar interface, given that I might need to process the retrieval set more than once in the future. Therefore, Google Scholar was not used in the main search for siblings, although it was efficient in locating important, high quality literature for the literature review. However, Google Scholar is evolving and expanding, and therefore is now offering more options for analysis. In terms of sibling retrieval, Google Scholar might be useful since it is a huge information repository that links to several information resources with different coverage, interests and indexing processes. For social science research and the grey literature it may be particularly useful (for retrieval of material that is less well covered by the bibliographic databases). It is difficult to evaluate (as retrieval results change so quickly) but for a one-off search for siblings it may be promising. Therefore, more research studies are required on Google Scholar.

SCOPUS database has a wide coverage of social science subjects. It comprises many citations and abstracts containing peer-reviewed research literature (42.5 million records) – journals, books and conferences - and quality web sources. SCOPUS offers researchers a quick, easy, comprehensive and valuable information resource in the scientific, technical and medical and social sciences fields. From this aspect, it is an appropriate choice when considering evidence based practice with research designs other than the RCT. Its potential was shown in this thesis, as it was one of the best databases alongside WoK which offers a wide coverage of the scientific, technical, medical and social sciences. Early research on SCOPUS was less promising but this research suggests that it has more potential now, and should be reviewed more thoroughly.

Further work is needed to identify how grey literature, conference proceedings and thesis and dissertation material can be obtained efficiently, as Web of Knowledge found some of the direct sibling publications for the IDEATel studies that could not be obtained on any of the other databases used. (This is also where Google Scholar and SCOPUS might be useful in future). If a larger pool of retrieved items is obtained, other screening approaches (such as text mining) may help, however, at this point it seems that text mining tools still need to be thoroughly investigated, focusing on screening process, and applicability to outputs from several types of database.

The time period used in this research was large and therefore creating a gold standard using the conventional way was impractical. Moreover, the siblings have different nature form each other, different objectives and research design and accordingly it is difficult to decide on which databases or journal to look for these siblings. This suggests that the conventional gold standard is not always the perfect way for retrieval performance tests which therefore call for the need for another way that is capable of providing a good performance comparison base. Relative recall was used in this study to compare and calculate retrieval performance using a pooled list of siblings for each seed study. Relative recall proved to provide a good base for performance comparisons and calculations and therefore deserves more consideration. Researchers claim that relative recall is more realistic reflection of the real world, and this, after all, was one of the aims of this research study, to provide practical guidance, as well as exploring in more depth what the meaning – in retrieval terms – of a sibling study is.

Appendix One

Search Strings

This appendix presents the search strings that were used to search for sibling studies for each one of the seeds studies on different databases. The search strings are basically subject search based on the key terms extracted from the seed study which was combined with author names to refine the retrieval results. Search filters were added to subject search strings to test their retrieval performance of sibling studies.

PubMed on MEDLINE database

IDEATeL / Telemedicine seed study

***Authors**:** *(Shea OR Weinstock OR Starren OR Teresi OR Palmas OR Field OR Morin OR Goland OR Izquierdo OR Wolff OR Ashraf OR Hilliman OR Silver OR Meyer ORHolmes OR Petkova OR Capps OR Lantigua)*

- **Simple search strategy**

Diabetes AND telemedicine AND (Authors ORed)* AND ("1992"[PDat] : "2010"[PDat])

- ***Qualitative:***

(interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract])AND diabetes telemedicine AND (("1992"[PDat] : "2010"[PDat]))

- ***costs:***

"costs and cost analysis"[MeSH] OR costs[Title/Abstract] OR cost effective*[Title/Abstract] AND diabetes telemedicine AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND diabetes telemedicine AND (("1992"[PDat] : "2010"[PDat]))

IFV seed study

***Authors*[*]**: *(Diedrich OR Devroey OR Engels OR Quartarolo OR Hiller OR Rudolf OR Sterzik OR van der Ven OR Verhoeven OR Dirnfeld OR Dor OR Ron-El OR Laufer OR Levran OR Shalev OR Jansen OR Schmoutziguer OR Germound OR Haeberle OR Kingsland OR Johnson OR Klentzeris OR Murdoch OR Sathanandan OR Sharp)*

- **Simple search strategy**

((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (Authors ORed)[*]AND (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND ((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (("1992"[PDat] : "2010"[PDat]))

Note! *When I used the string ([(in vitro fertilization OR IVF) AND (highly purified menotropin OR hPM) AND (recombinant follicle-stimulating hormone OR rFSH)]) no records have been retrieved as search results.*

- ***Costs:***

costs[Title/Abstract] OR "costs and cost analysis"[MeSH] OR cost effective*[Title/Abstract] AND [(in vitro fertilization OR IVF) AND (highly purified menotropin OR hPM) AND (recombinant follicle-stimulating hormone OR rFSH)] AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND [(in vitro fertilization OR IVF) AND (highly purified menotropin OR hPM) AND (recombinant follicle-stimulating hormone OR rFSH)] AND (("1992"[PDat] : "2010"[PDat]))

Tamoxifen seed study

***Authors*[*]**: *(Fallowfield OR Fleissig OR Edwards OR West OR Powles OR Howell OR Cuzick)*

- **Simple search strategy**

Breast cancer prevention AND tamoxifen AND (authors ORed)[*] AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND breast cancer prevention and tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

- ***Costs:***

costs[Title/Abstract] OR "costs and cost analysis"[MeSH] OR cost effective*[Title/Abstract] AND breast cancer prevention AND tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND breast cancer prevention and tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

Dexamethasone seed study

Authors*: *(Anderson OR Bednarek OR Dreyer OR Magoon OR Mercier OR Soll OR Garland OR Havens OR McAuliffe OR Nelson OR (Vermont Oxford Network Steroid Study Group)*

- **Simple search strategy**

dexamethasone AND chronic lung disease AND (authors ORed)* AND ("1992"[PDat] : "2010"[PDat])

- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND dexamethasone AND chronic lung disease
AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND dexamethasone AND chronic lung disease AND (("1992"[PDat] : "2010"[PDat]))

- ***Costs:***

costs[Title/Abstract] OR "costs and cost analysis"[MeSH] OR cost effective*[Title/Abstract] AND dexamethasone AND chronic lung disease AND (("1992"[PDat] : "2010"[PDat]))

BMS seed study

Authors*: (Liu CJ OR Hsiung OR Chang OR Liu YF OR Wang OR Hsiao OR Ng OR Chan)*

- **Simple search strategy**

Breast cancer AND body mind spirit (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND breast cancer body mind spirit (("1992"[PDat] : "2010"[PDat]))

Neither costs nor economical studies were found.

MetaLib as database

MetaLib search was conducted in two different stages each stage compromise different choice of databases to be searched. The first search was conducted on February 2010.

Stage 1 search was a simple subject search as the following:

IDEATeL seed study

- Telemedicine AND diabetes

Tamoxifen seed study

- Tamoxifen AND breast cancer prevention

IVF seed study

- In vitro fertilization AND recombinant follicle stimulation

BMS seed study

- Breast cancer AND body mind spirit therapy
- Breast cancer AND mental health (this search string did not prove to be efficient to retrieve any siblings, therefore the terms mental health were not considered for further exploration on other databases)

CLD seed study

- Chronic lung disease AND dexamethasone AND early postnatal AND prevention

Web of Knowledge (WoK) database

IDEATeL / Telemedicine seed study

***Authors*[*]:** *(Shea OR Weinstock OR Starren OR Teresi OR Palmas OR Field OR Morin OR Goland OR Izquierdo OR Wolff OR Ashraf OR Hilliman OR Silver OR Meyer ORHolmes OR Petkova OR Capps OR Lantigua)*

- **Simple search strategy**

Diabetes AND telemedicine AND (Authors ORed)[*] AND ("1992"[PDat] : "2010"[PDat])

- ***Qualitative:***

(interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract])AND diabetes telemedicine AND (("1992"[PDat] : "2010"[PDat]))

- ***costs:***

"costs and cost analysis"[MeSH] OR costs[Title/Abstract] OR cost effective*[Title/Abstract] AND diabetes telemedicine AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND diabetes telemedicine AND (("1992"[PDat] : "2010"[PDat]))

IFV seed study

***Authors*[*]:** *(Diedrich OR Devroey OR Engels OR Quartarolo OR Hiller OR Rudolf OR Sterzik OR van der Ven OR Verhoeven OR Dirnfeld OR Dor OR Ron-El OR Laufer OR Levran OR Shalev OR Jansen OR Schmoutziguer OR Germound OR Haeberle OR Kingsland OR Johnson OR Klentzeris OR Murdoch OR Sathanandan OR Sharp)*

- **Simple search strategy**

((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (Authors ORed)[*]AND (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND ((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (("1992"[PDat] : "2010"[PDat]))

*****when I used the string ([(in vitro fertilization OR IVF) AND (highly purified menotropin OR hPM) AND (recombinant follicle-stimulating hormone OR rFSH)]) no records have been retrieved as search results.*

- ***Costs:***

costs[Title/Abstract] OR "costs and cost analysis"[MeSH] OR cost effective*[Title/Abstract] AND [(in vitro fertilization OR IVF) AND (highly purified menotropin OR hPM) AND (recombinant follicle-stimulating hormone OR rFSH)] AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND [(in vitro fertilization OR IVF) AND (highly purified menotropin OR hPM) AND (recombinant follicle-stimulating hormone OR rFSH)] AND (("1992"[PDat] : "2010"[PDat]))

Tamoxifen seed study

Authors*[*]: *(Fallowfield OR Fleissig OR Edwards OR West OR Powles OR Howell OR Cuzick)

- **Simple search strategy**

Breast cancer prevention AND tamoxifen AND (authors ORed)[*] AND (("1992"[PDat] : "2010"[PDat]))

- ***Economics:***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND breast cancer prevention and tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

- ***Costs:***

costs[Title/Abstract] OR "costs and cost analysis"[MeSH] OR cost effective*[Title/Abstract] AND breast cancer prevention AND tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND breast cancer prevention and tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

BMS seed study

Authors*[*]: *(Liu CJ OR Hsiung OR Chang OR Liu YF OR Wang OR Hsiao OR Ng OR Chan)

- **Simple search strategy**

Breast cancer AND body mind spirit AND(Authors ORed)[*] AND (("1992"[PDat] : "2010"[PDat]))

- **Simple search strategy with filters**
- ***Qualitative:***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND breast cancer body mind spirit (("1992"[PDat] : "2010"[PDat]))

<u>Dexamethasone seed study</u>

***Authors*[*]:** *(Anderson OR Bednarek OR Dreyer OR Magoon OR Mercier OR Soll OR Garland OR Havens OR McAuliffe OR Nelson OR (Vermont Oxford Network Steroid Study Group)*

- **Simple search strategy**

dexamethasone AND chronic lung disease AND (authors ORed)[*] AND (("1992"[PDat] : "2010"[PDat])

- ***<u>Qualitative:</u>***

interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract] AND dexamethasone AND chronic lung disease
AND (("1992"[PDat] : "2010"[PDat]))

- ***<u>Economics:</u>***

cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp] AND dexamethasone AND chronic lung disease AND (("1992"[PDat] : "2010"[PDat]))

- ***<u>Costs:</u>***

costs[Title/Abstract] OR "costs and cost analysis"[MeSH] OR cost effective*[Title/Abstract] AND dexamethasone AND chronic lung disease AND (("1992"[PDat] : "2010"[PDat]))

SCOPUS Database

IDEATeL / Telemedicine seed study

Authors*[*]*: *(Shea OR Weinstock OR Starren OR Teresi OR Palmas OR Field OR Morin OR Goland OR Izquierdo OR Wolff OR Ashraf OR Hilliman OR Silver OR Meyer ORHolmes OR Petkova OR Capps OR Lantigua)*

- **Simple search strategy**

Diabetes AND telemedicine AND (Authors ORed)[*] AND ("1992"[PDat] : "2010"[PDat])

- ***Qualitative:***

(TITLE-ABS-KEY(Diabetes AND telemedicine) AND TITLE-ABS-KEY(interview* OR experience* OR qualitative) AND AUTH((shea OR weinstock OR starren OR teresi OR palmas OR field OR morin OR goland OR izquierdo OR wolff Or ashraf OR hilliman OR silver OR meyer OR holmes OR petkova OR capps OR lantigua))) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Costs:***

(TITLE-ABS-KEY(Diabetes AND telemedicine) AND TITLE-ABS-KEY("costs and cost analysis" OR costs OR cost effective*) AND AUTH((shea OR weinstock OR starren OR teresi OR palmas OR field OR morin OR goland OR izquierdo OR wolff Or ashraf OR hilliman OR silver OR meyer OR holmes OR petkova OR capps OR lantigua))) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Economics:***

(TITLE-ABS-KEY(Diabetes AND telemedicine) AND TITLE-ABS-KEY (cost* OR "costs and cost analysis" OR cost benefit analys* OR cost-benefit analysis OR health care costs) AND AUTH((shea OR weinstock OR starren OR teresi OR palmas OR field OR morin OR goland OR izquierdo OR wolff Or ashraf OR hilliman OR silver OR meyer OR holmes OR petkova OR capps OR lantigua))) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

IFV seed study

Authors[*]: *(Diedrich OR Devroey OR Engels OR Quartarolo OR Hiller OR Rudolf OR Sterzik OR van der Ven OR Verhoeven OR Dirnfeld OR Dor OR Ron-El OR Laufer OR Levran OR Shalev OR Jansen OR Schmoutziguer OR Germound OR Haeberle OR Kingsland OR Johnson OR Klentzeris OR Murdoch OR Sathanandan OR Sharp)*

- **Simple search strategy**

((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (Authors ORed)[*]AND (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

(TITLE-ABS-KEY((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND TITLE-ABS-KEY(interview* OR experience* OR qualitative)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Costs:***

(TITLE-ABS-KEY((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND TITLE-ABS-KEY("costs and cost analysis" OR costs OR cost effective*)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Economics:***

TITLE-ABS-KEY((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND TITLE-ABS-KEY(cost* OR "costs and cost analysis" OR cost benefit analys* OR cost-benefit analysis OR health care costs) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

Tamoxifen seed study

Authors*[*]*: *(Fallowfield OR Fleissig OR Edwards OR West OR Powles OR Howell OR Cuzick)*

- **Simple search strategy**

Breast cancer prevention AND tamoxifen AND (authors ORed)[*] AND (("1992"[PDat] : "2010"[PDat]))

- ***Qualitative:***

TITLE-ABS-KEY(Breast cancer prevention AND tamoxifen) AND TITLE-ABS-KEY(interview* OR experience* OR qualitative)) AND AUTH (Fallowfield OR Fleissig OR Edwards OR West OR Powles OR Howell OR Cuzick) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Costs:***

TITLE-ABS-KEY(Breast cancer prevention AND tamoxifen) AND TITLE-ABS-KEY("costs and cost analysis" OR costs OR cost effective*)) AND (EXCLUDE(PUBYEAR,2012) OR EXCLUDE(PUBYEAR,2011)) AND (EXCLUDE(PUBYEAR,1991))

- ***Economics:***

(TITLE-ABS-KEY(Breast cancer prevention AND tamoxifen) AND TITLE-ABS-KEY(cost* OR "costs and cost analysis" OR cost benefit analys* OR cost-benefit analysis OR health care costs)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

Dexamethasone seed study

Authors*[*]*: *(Anderson OR Bednarek OR Dreyer OR Magoon OR Mercier OR Soll OR Garland OR Havens OR McAuliffe OR Nelson OR (Vermont Oxford Network Steroid Study Group))*

- **Simple search strategy**

dexamethasone AND chronic lung disease AND (authors ORed)[*] AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Qualitative:***

(TITLE-ABS-KEY(dexamethasone AND chronic lung disease) AND TITLE-ABS-KEY(interview* OR experience* OR qualitative)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Costs:***

(TITLE-ABS-KEY(dexamethasone AND chronic lung disease) AND TITLE-ABS-KEY("costs and cost analysis" OR costs OR cost effective*)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Economics:***

(TITLE-ABS-KEY(dexamethasone AND chronic lung disease) AND TITLE-ABS-KEY(cost* OR "costs and cost analysis" OR cost benefit analys* OR cost-benefit analysis OR health care costs)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

BMS seed study

Authors*[*]*: *(Liu CJ OR Hsiung OR Chang OR Liu YF OR Wang OR Hsiao OR Ng OR Chan)*

- **Simple search strategy**

Breast cancer AND body mind spirit (("1992"[PDat] : "2010"[PDat]))

- ***Costs:***

(TITLE-ABS-KEY(Breast cancer AND body mind spirit) AND TITLE-ABS-KEY("costs and cost analysis" OR costs OR cost effective*)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

- ***Economics:***

(TITLE-ABS-KEY(Breast cancer AND body mind spirit) AND TITLE-ABS-KEY(cost* OR "costs and cost analysis" OR cost benefit analys* OR cost-benefit analysis OR health care costs)) AND (EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,2011) OR EXCLUDE(PUBYEAR,1991))

CINAHL Database

IDEATeL seed study

***Authors*[*]*:** *(Shea OR Weinstock OR Starren OR Teresi OR Palmas OR Field OR Morin OR Goland OR Izquierdo OR Wolff OR Ashraf OR Hilliman OR Silver OR Meyer ORHolmes OR Petkova OR Capps OR Lantigua)*

- **Simple search strategy**

Diabetes AND telemedicine AND (Authors ORed)[*] AND ("1992"[PDat] : "2010"[PDat])

- ***Qualitative:***

TX ((Diabetes AND telemedicine)) AND AU (Authors ORed)[*] AND (MW interviews OR TX experience* OR AB (qualitative OR interview*) OR TI (qualitative OR interview*))

- ***Costs:***

(TX ((Diabetes AND telemedicine)) AND AU (Authors ORed)[*])) AND (MW (costs and cost analysis) OR TX (cost effective* OR costs))

- ***Economics:***

TX ((Diabetes AND telemedicine)) AND AU ((Authors ORed)[*]) AND (MW ("costs and cost analysis" OR cost-benefit analysis OR health care costs
) OR TX experience* OR AB (cost* OR cost benefit analys*) OR TI (cost* OR cost benefit analys*))

IFV seed study

***Authors*[*]*:** *(Diedrich OR Devroey OR Engels OR Quartarolo OR Hiller OR Rudolf OR Sterzik OR van der Ven OR Verhoeven OR Dirnfeld OR Dor OR Ron-El OR Laufer OR Levran OR Shalev OR Jansen OR Schmoutziguer OR Germound OR Haeberle OR Kingsland OR Johnson OR Klentzeris OR Murdoch OR Sathanandan OR Sharp)*

- **Simple search strategy**

((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (Authors ORed)*AND (("1992"[PDat] : "2010"[PDat]))

Tamoxifen seed study

***Authors*[*]**: *(Fallowfield OR Fleissig OR Edwards OR West OR Powles OR Howell OR Cuzick)*

- **Simple search strategy**

TX (Breast cancer prevention AND tamoxifen) AND AU (Authors ORed)[*]

- ***Qualitative:***

TX (Breast cancer prevention AND tamoxifen) AND AND (MW interviews OR TX experience* OR AB (qualitative OR interview*) OR TI (qualitative OR interview*))

- ***Costs:***

TX (Breast cancer prevention AND tamoxifen) AND AND (MW (costs and cost analysis) OR TX (cost effective* OR costs))

- ***Economics:***

TX (Breast cancer prevention AND tamoxifen) AND (MW ("costs and cost analysis" OR cost-benefit analysis OR health care costs) OR TX experience* OR AB (cost* OR cost benefit analys*) OR TI (cost* OR cost benefit analys*))

Dexamethasone seed study

Authors*: *(Anderson OR Bednarek OR Dreyer OR Magoon OR Mercier OR Soll OR Garland OR Havens OR McAuliffe OR Nelson OR (Vermont Oxford Network Steroid Study Group)*

- **Simple search strategy**

TX(dexamethasone AND chronic lung disease)AND AU(Authors ORed)* AND (("1992"[PDat] : "2010"[PDat])

- ***Qualitative:***

TX(dexamethasone AND chronic lung disease) AND (MW interviews OR TX experience* OR AB (qualitative OR interview*) OR TI (qualitative OR interview*)

- ***Costs:***

TX(dexamethasone AND chronic lung disease) AND AU AND (MW (costs and cost analysis) OR TX (cost effective* OR costs))

- ***Economics:***

TX(dexamethasone AND chronic lung disease) AND (MW ("costs and cost analysis" OR cost-benefit analysis OR health care costs) OR TX experience* OR AB (cost* OR cost benefit analys*) OR TI (cost* OR cost benefit analys*)

<u>Breast cancer body mind spirit:-</u>

Authors*: (Liu CJ OR Hsiung OR Chang OR Liu YF OR Wang OR Hsiao OR Ng OR Chan)

- **Simple search strategy**

Breast cancer AND body mind spirit (("1992"[PDat] : "2010"[PDat]))

- ***<u>Qualitative:</u>***

TX(Breast cancer AND body mind spirit) AND AND (MW interviews OR TX experience* OR AB (qualitative OR interview*) OR TI (qualitative OR interview*))

- ***<u>Costs:</u>***

TX(Breast cancer AND body mind spirit) AND (MW (costs and cost analysis) OR TX (cost effective* OR costs))

- ***<u>Economics:</u>***

TX(Breast cancer AND body mind spirit) AND (MW ("costs and cost analysis" OR cost-benefit analysis OR health care costs) OR TX experience* OR AB (cost* OR cost benefit analys*) OR TI (cost* OR cost benefit analys*))

Appendix Two

PubMed Sophisticated Search String and ATM

This appendix provides the full translated search query, authors were omitted from the query for a clearer view. In the end, most of the search terms used in the sophisticated search[♣], were used in the simple search strategy after PubMed did the mapping.

PubMed ATM *vs* Sophisticated Search String

Simple search string for IDEATeL seed study:

- **User's query terms:** 'diabetes AND telemedicine'

Translations:	
diabetes	"diabetes mellitus"[MeSH Terms] OR ("diabetes"[All Fields] AND "mellitus"[All Fields]) OR "diabetes mellitus"[All Fields] OR "diabetes"[All Fields] OR "diabetes insipidus"[MeSH Terms] OR ("diabetes"[All Fields] AND "insipidus"[All Fields]) OR "diabetes insipidus"[All Fields]
telemedicine	"telemedicine"[MeSH Terms] OR "telemedicine"[All Fields]

Sophisticated search string:

- **User's query terms**: (Diabetes Mellitus [MeSH] OR diabet* [ti] OR "Blood Glucose Self-Monitoring" [MeSH] OR "glucose self monitoring" [tw]) AND ("telemedicine"[MeSH Terms] OR "remote consultation"[MeSH Terms] OR ("telecommunications"[MeSH Terms] AND "Referral and Consultation"[MeSH]) OR Telemedicine[tw] OR telecare[tw] OR telehealth[tw] OR e-health[tw] OR telecommunications[tw] OR telecommunication[tw] OR mobile health[tw] OR m-health [tw])

Translations:	
Diabetes Mellitus[MeSH]	"diabetes mellitus"[MeSH Terms]

Simple search string for BMS seed study:

- **User's query:** breast cancer AND body mind spirit (("1992"[PDat] : "2010"[PDat]))

Translations:	
breast cancer	"breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR "breast cancer"[All Fields]
body	"human body"[MeSH Terms] OR ("human"[All Fields] AND "body"[All Fields]) OR "human body"[All Fields] OR "body"[All Fields]
mind	"Mind"[Journal] OR "mind"[All Fields]

[♣] Sophisticated search were amended with search filters for performance assessment. Therefore for each seed study there is four sophisticated search string.

Sophisticated search string:

- **User's query:** (Breast neoplasms [MeSH] OR Breast Tumor* [All Fields] OR breast tumour [All Fields] OR breast cancer [All Fields]) AND (holistic health [MeSH] OR holistic [All Fields]) OR body-mind relation [MeSH] OR body-soul relation [MeSH] OR body mind spirit [All Fields] OR mind-body therapies [MeSH]) AND (group therapy [All Fields] OR group psychotherapy [MeSH])

Translations:

Breast neoplasms[MeSH]	"breast neoplasms"[MeSH Terms]
breast tumour[All Fields]	"breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "tumour"[All Fields]) OR "breast tumour"[All Fields]
breast cancer[All Fields]	"breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR "breast cancer"[All Fields]
holistic health[MeSH]	"holistic health"[MeSH Terms]
body-mind relation[MeSH]	"mind-body relations, metaphysical"[MeSH Terms]
mind-body therapies[MeSH]	"mind-body therapies"[MeSH Terms]
group therapy[All Fields]	"psychotherapy, group"[MeSH Terms] OR ("psychotherapy"[All Fields] AND "group"[All Fields]) OR "group psychotherapy"[All Fields] OR ("group"[All Fields] AND "therapy"[All Fields]) OR "group therapy"[All Fields]
group psychotherapy[MeSH]	"psychotherapy, group"[MeSH Terms]

Simple search string for Tamoxifen seed study:

- **User's query:** breast cancer AND prevention AND tamoxifen AND (("1992"[PDat] : "2010"[PDat]))

Translations:

breast cancer	"breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR "breast cancer"[All Fields]
prevention	"prevention and control"[Subheading] OR ("prevention"[All Fields] AND "control"[All Fields]) OR "prevention and control"[All Fields] OR "prevention"[All Fields]
tamoxifen	"tamoxifen"[MeSH Terms] OR "tamoxifen"[All Fields]

Sophisticated search string:

- **User's query terms**:(Breast neoplasms [MeSH] OR Breast Tumor* [tw] OR breast tumour [tw] OR breast cancer [All Fields]) AND (Tamoxifen [MeSH] OR tamoxifen [All Fields]) AND (prevention [All Fields] OR prevention [MeSH])

Translations:

Breast neoplasms[MeSH]	"breast neoplasms"[MeSH Terms]
breast cancer[All Fields]	"breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR "breast cancer"[All Fields]
Tamoxifen[MeSH]	"tamoxifen"[MeSH Terms]
tamoxifen[All Fields]	"tamoxifen"[MeSH Terms] OR "tamoxifen"[All Fields]
prevention[All Fields]	"prevention and control"[Subheading] OR ("prevention"[All Fields] AND "control"[All Fields]) OR "prevention and control"[All Fields] OR "prevention"[All Fields]

Simple search string for CLD seed study:

- **User's query:** dexamethasone AND chronic lung disease AND (("1992"[PDat] : "2010"[PDat]))

Translations:	
dexamethasone	"dexamethasone"[MeSH Terms] OR "dexamethasone"[All Fields]
lung disease	"lung diseases"[MeSH Terms] OR ("lung"[All Fields] AND "diseases"[All Fields]) OR "lung diseases"[All Fields] OR ("lung"[All Fields] AND "disease"[All Fields]) OR "lung disease"[All Fields]

Sophisticated search string:

- **User's query:** ((Dexamethasone [MeSH] OR Dexamethasone [All Fields]) AND (postnatal [All Fields] OR early postnatal [All Fields]) AND (Lung Diseases, Obstructive [MeSH] OR Chronic Bronchitis [MeSH] OR chronic Lung Diseases [All Fields] OR chronic Lung Diseases [ti] OR CLD [All Fields]) AND (prevention [ti] OR prevention [All Fields])

Translations:	
Dexamethasone[MeSH]	"dexamethasone"[MeSH Terms]
Dexamethasone[All Fields]	"dexamethasone"[MeSH Terms] OR "dexamethasone"[All Fields]
Lung Diseases, Obstructive[MeSH]	"lung diseases, obstructive"[MeSH Terms]
Chronic Bronchitis[MeSH]	"bronchitis, chronic"[MeSH Terms]
prevention[All Fields]	"prevention and control"[Subheading] OR ("prevention"[All Fields] AND "control"[All Fields]) OR "prevention and control"[All Fields] OR "prevention"[All Fields]

Simple search string for IVF seed study:

- **User's query:** ((In vitro fertilisation OR IVF) AND (highly purified menotropin OR recombinant follicle-stimulating hormone)) AND (("1992"[PDat] : "2010"[PDat]))

Translations:	
In vitro fertilisation	"in vitro fertilisation"[All Fields] OR "fertilization in vitro"[MeSH Terms] OR ("fertilization"[All Fields] AND "vitro"[All Fields]) OR "fertilization in vitro"[All Fields] OR ("vitro"[All Fields] AND "fertilization"[All Fields]) OR "in vitro fertilization"[All Fields]
IVF	"J In Vitro Fert Embryo Transf"[Journal] OR "ivf"[All Fields]
menotropin	"menotropins"[MeSH Terms] OR "menotropins"[All Fields] OR "menotropin"[All Fields]
follicle-stimulating hormone	"follicle stimulating hormone"[MeSH Terms] OR ("follicle"[All Fields] AND "stimulating"[All Fields] AND "hormone"[All Fields]) OR "follicle stimulating hormone"[All Fields]

Sophisticated search string:

- **User's query:** (Fertilization in Vitro [MeSH] OR Fertilization in Vitro [All Fields] OR Sperm Injections, Intracytoplasmic [MeSH]) AND (((FSH [All Fields] OR Follicle Stimulating Hormone [MeSH] OR FSH [All Fields] OR Follicle Stimulating Hormone [All Fields]) AND recombinant [tw]) AND ((human menopausal gonadotropin [MeSH] OR menotropin [MeSH] OR human menopausal gonadotropin [All Fields] OR menotropin [All Fields] OR HP-HMG [All Fields] OR HP-hMG [All Fields]) AND (highly purified [All Fields])))))))

Translations:

Fertilization in Vitro[MeSH]	"fertilization in vitro"[MeSH Terms]
Fertilization in Vitro[All Fields]	"fertilisation in vitro"[All Fields] OR "fertilization in vitro"[MeSH Terms] OR ("fertilization"[All Fields] AND "vitro"[All Fields]) OR "fertilization in vitro"[All Fields]
Sperm Injections, Intracytoplasmic[MeSH]	"sperm injections, intracytoplasmic"[MeSH Terms]
Follicle Stimulating Hormone[MeSH]	"follicle stimulating hormone"[MeSH Terms]
Follicle Stimulating Hormone[All Fields]	"follicle stimulating hormone"[MeSH Terms] OR ("follicle"[All Fields] AND "stimulating"[All Fields] AND "hormone"[All Fields]) OR "follicle stimulating hormone"[All Fields]
human menopausal gonadotropin[MeSH]	"menotropins"[MeSH Terms]
menotropin[MeSH]	"menotropins"[MeSH Terms]
human menopausal gonadotropin[All Fields]	"menotropins"[MeSH Terms] OR "menotropins"[All Fields] OR ("human"[All Fields] AND "menopausal"[All Fields] AND "gonadotropin"[All Fields]) OR "human menopausal gonadotropin"[All Fields]
menotropin[All Fields]	"menotropins"[MeSH Terms] OR "menotropins"[All Fields] OR "menotropin"[All Fields]

PubMed automatic mapping for search filters

- **Qualitative filter**

- **User's query terms:** interview*[Title/Abstract] OR interviews[MeSH] OR experience*[Text Word] OR qualitative[Title/Abstract]

- **Query Translation:** (interview[Title/Abstract] OR interview/clinical[Title/Abstract] OR interview/criteria[Title/Abstract] OR interview/debriefing[Title/Abstract] OR interview/demoralization[Title/Abstract] OR interview/evaluation[Title/Abstract] OR interview/examination[Title/Abstract] OR interview/examine[Title/Abstract] OR interview/focus[Title/Abstract] OR interview/form[Title/Abstract] OR interview/history[Title/Abstract] OR interview/inspection[Title/Abstract] OR interview/intake[Title/Abstract] OR interview/observation[Title/Abstract] OR interview/profile[Title/Abstract] OR interview/questionnaire[Title/Abstract] OR interview/recall[Title/Abstract] OR interview/reporting[Title/Abstract] OR interview/residual[Title/Abstract] OR interview/s[Title/Abstract] OR interview/screening[Title/Abstract] OR interview/simulation[Title/Abstract] OR interview/survey[Title/Abstract] OR interview/test[Title/Abstract] OR interview/testing[Title/Abstract] OR interview/therapy[Title/Abstract] OR interview/weight[Title/Abstract] OR interview'[Title/Abstract] OR interview's[Title/Abstract] OR interviewability[Title/Abstract] OR interviewable[Title/Abstract] OR interviewable'[Title/Abstract] OR interviewd[Title/Abstract] OR interviewdata[Title/Abstract] OR interviewed[Title/Abstract] OR interviewed/examined[Title/Abstract] OR interviewed/serotested[Title/Abstract] OR interviewed'[Title/Abstract] OR intervieweds[Title/Abstract] OR interviewee[Title/Abstract] OR interviewee/interviewer[Title/Abstract] OR interviewee's[Title/Abstract] OR interviewee'speeches[Title/Abstract] OR interviewees[Title/Abstract] OR interviewees'[Title/Abstract] OR interviewer[Title/Abstract] OR interviewer/moderator[Title/Abstract] OR interviewer/observer[Title/Abstract] OR interviewer/participant[Title/Abstract] OR interviewer/researcher[Title/Abstract] OR interviewer/respondent[Title/Abstract] OR interviewer/volunteer[Title/Abstract] OR interviewer'[Title/Abstract] OR interviewer's[Title/Abstract] OR interviewer3[Title/Abstract] OR interviewers[Title/Abstract] OR interviewers/moderators[Title/Abstract] OR interviewers'[Title/Abstract] OR interviewes[Title/Abstract] OR interviewing[Title/Abstract] OR interviewing/brief[Title/Abstract] OR interviewing/cbt[Title/Abstract] OR interviewing/cognitive[Title/Abstract] OR interviewing/communicating[Title/Abstract] OR interviewing/communication[Title/Abstract] OR interviewing/history[Title/Abstract] OR interviewing/part[Title/Abstract] OR interviewing/physical[Title/Abstract] OR interviewing/skills[Title/Abstract] OR interviewing/supportive[Title/Abstract] OR interviewing'[Title/Abstract] OR interviewing's[Title/Abstract] OR interviews[Title/Abstract] OR interviews/6[Title/Abstract] OR interviews/assessments[Title/Abstract] OR interviews/blood[Title/Abstract] OR interviews/descriptive[Title/Abstract] OR interviews/discussions[Title/Abstract] OR interviews/doctor/day[Title/Abstract] OR interviews/drawings[Title/Abstract] OR interviews/enquiries[Title/Abstract] OR interviews/examinations[Title/Abstract] OR interviews/field[Title/Abstract] OR interviews/focus[Title/Abstract] OR interviews/group[Title/Abstract] OR interviews/inhabitant/year[Title/Abstract] OR interviews/investigations[Title/Abstract] OR interviews/label[Title/Abstract] OR interviews/narratives[Title/Abstract] OR interviews/patient[Title/Abstract] OR interviews/questionnaire[Title/Abstract] OR interviews/questionnaires[Title/Abstract] OR interviews/retrospective[Title/Abstract] OR interviews/storytelling[Title/Abstract] OR interviews/survey[Title/Abstract] OR interviews/surveys[Title/Abstract] OR interviews/surveys/focus[Title/Abstract] OR interviews/talks/discussions[Title/Abstract] OR interviews/targeted[Title/Abstract] OR interviews'[Title/Abstract] OR interviews1[Title/Abstract] OR interviews3d[Title/Abstract] OR interviewtechniken[Title/Abstract] OR interviewtext[Title/Abstract] OR interviewwas[Title/Abstract]) OR "interviews as topic"[MeSH Terms] OR (experience[Text Word] OR experience/activity[Text Word] OR experience/animals[Text Word] OR experience/anticipation[Text Word] OR experience/chronic[Text Word] OR experience/cognitive/limbic[Text Word] OR experience/competence[Text Word] OR experience/conditioning[Text Word] OR experience/demographic[Text Word] OR experience/event[Text Word] OR experience/exposure[Text Word] OR experience/expression[Text Word] OR experience/first[Text Word] OR experience/high[Text Word] OR experience/history[Text Word] OR experience/hostility[Text Word] OR experience/instruction[Text Word] OR experience/intellect[Text Word] OR experience/involvement[Text Word] OR experience/knowledge[Text Word] OR experience/learning[Text Word] OR experience/module[Text Word] OR experience/motivation[Text Word] OR experience/observation[Text Word] OR experience/opinion[Text Word] OR experience/pain[Text Word] OR experience/perception[Text Word] OR experience/performance[Text Word] OR experience/perpetrate[Text Word] OR experience/philosophy[Text Word] OR experience/practice[Text Word] OR experience/preferred[Text Word] OR experience/proficient[Text Word] OR experience/publication[Text Word] OR experience/quality[Text Word] OR experience/reporting[Text Word] OR experience/s[Text Word] OR experience/second[Text Word] OR experience/skill[Text Word] OR experience/skills[Text Word] OR experience/tolerate[Text Word] OR experience/training[Text Word] OR experience/visualize[Text Word] OR experience/wishes[Text Word] OR experience'[Text Word] OR experience"[Text Word] OR experience's[Text Word] OR experienceable[Text Word] OR experienced[Text Word] OR experienced/activated[Text Word] OR experienced/anticipated[Text Word] OR experienced/inexperienced[Text Word] OR experienced/memory[Text Word] OR experienced/naive[Text Word] OR experienced/unexperienced[Text Word] OR experienced/wished[Text Word] OR experienced/witnessed[Text Word] OR experienced'[Text Word] OR experiencedrecent[Text Word] OR experiencee[Text Word] OR experienceed[Text Word] OR experiencefirst[Text Word] OR experienceing[Text Word] OR experienceon[Text Word] OR experienceoof[Text Word] OR experiencer[Text Word] OR experiencers[Text Word] OR experiencers'[Text Word] OR experiences[Text Word] OR experiences/attitudes[Text Word] OR experiences/conditions[Text Word] OR experiences/coping[Text Word] OR experiences/feelings[Text Word] OR experiences/harassment[Text Word] OR experiences/hope/acceptance/ability[Text Word] OR experiences/ideas[Text Word] OR experiences/insights[Text Word] OR experiences/interactions[Text Word] OR experiences/interests[Text Word] OR experiences/knowledge[Text Word] OR experiences/orgasms[Text Word] OR experiences/perceptions[Text Word] OR experiences/processes[Text Word] OR experiences/recent[Text Word] OR experiences/sanctions[Text Word] OR experiences/satisfaction[Text Word] OR experiences/side[Text Word] OR experiences/training[Text Word] OR experiences/wishes[Text Word] OR experiences'[Text Word] OR experiencesbackground[Text Word]) OR qualitative[Title/Abstract]

- **Costs filter**

  - **User's query terms:** "costs and cost analysis"[MeSH] OR costs[Title/Abstract] OR cost effective*[Title/Abstract]

- **Query Translation:** "costs and cost analysis"[MeSH] OR costs[Title/Abstract] OR (cost effective[Title/Abstract] OR cost effective/effectiveness[Title/Abstract] OR cost effectively[Title/Abstract] OR cost effectivene[Title/Abstract] OR cost effectivenees[Title/Abstract] OR cost effectivenes[Title/Abstract] OR cost effectiveness[Title/Abstract] OR cost effectiveness/competitive[Title/Abstract] OR cost effectiveness/economics[Title/Abstract] OR cost effectiveness/financial[Title/Abstract] OR cost effectiveness/risk[Title/Abstract] OR cost effectiveness/safety[Title/Abstract] OR cost effectiveness/utility[Title/Abstract] OR cost effectivenesses[Title/Abstract] OR cost effectivenss[Title/Abstract] OR cost effectivensss[Title/Abstract] OR cost effectiveutilization[Title/Abstract])

- **Economics filter**

  - **User's query terms:** cost*[Title/Abstract] OR "costs and cost analysis"[MeSH:noexp] OR cost benefit analys*[Title/Abstract] OR cost-benefit analysis[MeSH] OR health care costs[MeSH:noexp]

  - **Query Translation:** (cost[Title/Abstract] OR cost/1[Title/Abstract] OR cost/100[Title/Abstract] OR cost/abnormality[Title/Abstract] OR cost/acceptability[Title/Abstract] OR cost/access[Title/Abstract] OR cost/accounting[Title/Abstract] OR cost/accuracy[Title/Abstract] OR cost/acquisition[Title/Abstract] OR cost/additional[Title/Abstract] OR cost/affordability[Title/Abstract] OR cost/and[Title/Abstract] OR cost/application[Title/Abstract] OR cost/attendance[Title/Abstract] OR cost/benefice[Title/Abstract] OR cost/beneficial[Title/Abstract] OR cost/benefit[Title/Abstract] OR cost/benefit/effectiveness[Title/Abstract] OR cost/benefit/risk[Title/Abstract] OR cost/benefit/security[Title/Abstract] OR cost/benefits[Title/Abstract] OR cost/bleeding[Title/Abstract] OR cost/burden[Title/Abstract] OR cost/cap[Title/Abstract] OR cost/capita[Title/Abstract] OR cost/case[Title/Abstract] OR cost/case/day/service[Title/Abstract] OR cost/ceftazidime[Title/Abstract] OR cost/charge[Title/Abstract] OR cost/child[Title/Abstract] OR cost/claim[Title/Abstract] OR cost/clinical[Title/Abstract] OR cost/closure[Title/Abstract] OR cost/cocaine[Title/Abstract] OR cost/controlled[Title/Abstract] OR cost/convenience[Title/Abstract] OR cost/convenient[Title/Abstract] OR cost/cost[Title/Abstract] OR cost/course[Title/Abstract] OR cost/coverage[Title/Abstract] OR cost/cure[Title/Abstract] OR cost/cycle[Title/Abstract] OR cost/cyp[Title/Abstract] OR cost/d[Title/Abstract] OR cost/daly[Title/Abstract] OR cost/day[Title/Abstract] OR cost/day/patient[Title/Abstract] OR cost/ddd[Title/Abstract] OR cost/ddd/1,000[Title/Abstract] OR cost/death[Title/Abstract] OR cost/defined[Title/Abstract] OR cost/delta[Title/Abstract] OR cost/dialysis[Title/Abstract] OR cost/disability[Title/Abstract] OR cost/discharge[Title/Abstract] OR cost/distance[Title/Abstract] OR cost/dose[Title/Abstract] OR cost/drug[Title/Abstract] OR cost/ds[Title/Abstract] OR cost/ecg[Title/Abstract] OR cost/effect[Title/Abstract] OR cost/effective[Title/Abstract] OR cost/effectiveness[Title/Abstract] OR cost/efficacity[Title/Abstract] OR cost/efficacy[Title/Abstract] OR cost/efficiency[Title/Abstract] OR cost/efficiency/quality[Title/Abstract] OR cost/efficient[Title/Abstract] OR cost/episode[Title/Abstract] OR cost/equity[Title/Abstract] OR cost/estimated[Title/Abstract] OR cost/evidence[Title/Abstract] OR cost/examination[Title/Abstract] OR cost/examination/year[Title/Abstract] OR cost/expectation[Title/Abstract] OR cost/expected[Title/Abstract] OR cost/extra[Title/Abstract] OR cost/fee[Title/Abstract] OR cost/finance[Title/Abstract] OR cost/financing[Title/Abstract] OR cost/formula[Title/Abstract] OR cost/fraction[Title/Abstract] OR cost/funding[Title/Abstract] OR cost/gain[Title/Abstract] OR cost/graft[Title/Abstract] OR cost/h[Title/Abstract] OR cost/harm/benefit[Title/Abstract] OR cost/hbeag[Title/Abstract] OR cost/hbv[Title/Abstract] OR cost/health[Title/Abstract] OR cost/healthcare[Title/Abstract] OR cost/hectare/year[Title/Abstract] OR cost/high[Title/Abstract] OR cost/higher[Title/Abstract] OR cost/hospital[Title/Abstract] OR cost/hour[Title/Abstract] OR cost/house[Title/Abstract] OR cost/human[Title/Abstract] OR cost/impact[Title/Abstract] OR cost/impregnated[Title/Abstract] OR cost/improvement[Title/Abstract] OR cost/improves[Title/Abstract] OR cost/income[Title/Abstract] OR cost/inefficiency[Title/Abstract] OR cost/infected[Title/Abstract] OR cost/information[Title/Abstract] OR cost/inhabitant[Title/Abstract] OR cost/injury[Title/Abstract] OR cost/inpatient[Title/Abstract] OR cost/insert[Title/Abstract] OR cost/instrumentation[Title/Abstract] OR cost/insurance[Title/Abstract] OR cost/item[Title/Abstract] OR cost/iteration[Title/Abstract] OR cost/itn[Title/Abstract] OR cost/iwa[Title/Abstract] OR cost/kg[Title/Abstract] OR cost/kilogram[Title/Abstract] OR cost/kilograms[Title/Abstract] OR cost/lack[Title/Abstract] OR cost/legal[Title/Abstract] OR cost/length[Title/Abstract] OR cost/life[Title/Abstract] OR cost/life/year[Title/Abstract] OR cost/low[Title/Abstract] OR cost/ly[Title/Abstract] OR cost/lyg[Title/Abstract] OR cost/lys[Title/Abstract] OR cost/management[Title/Abstract] OR cost/member/month[Title/Abstract] OR cost/mg[Title/Abstract] OR cost/mg/m2[Title/Abstract] OR cost/minimum[Title/Abstract] OR cost/mm[Title/Abstract] OR cost/month[Title/Abstract] OR cost/mortality[Title/Abstract] OR cost/mycologic[Title/Abstract] OR cost/need[Title/Abstract] OR cost/noninvasive[Title/Abstract] OR cost/outcome[Title/Abstract] OR cost/outpatient[Title/Abstract] OR cost/overall[Title/Abstract] OR cost/page[Title/Abstract] OR cost/participant[Title/Abstract] OR cost/participant/day[Title/Abstract] OR cost/patient[Title/Abstract] OR cost/patient/day[Title/Abstract] OR cost/patient/month[Title/Abstract] OR cost/patient/year[Title/Abstract] OR cost/payment[Title/Abstract] OR cost/pe[Title/Abstract] OR cost/percentage[Title/Abstract] OR cost/performance[Title/Abstract] OR cost/persistence[Title/Abstract] OR cost/person[Title/Abstract] OR cost/person/m2[Title/Abstract] OR cost/positive[Title/Abstract] OR cost/ppv[Title/Abstract] OR cost/pregnancy[Title/Abstract] OR cost/prescription[Title/Abstract] OR cost/price[Title/Abstract] OR cost/processing[Title/Abstract] OR cost/profit[Title/Abstract] OR cost/qaly[Title/Abstract] OR cost/qaly/daly/ly[Title/Abstract] OR cost/qaly's[Title/Abstract] OR cost/qalys[Title/Abstract] OR cost/qaty[Title/Abstract] OR cost/quality[Title/Abstract] OR cost/recurrence[Title/Abstract] OR cost/referral[Title/Abstract] OR cost/reimbursement[Title/Abstract] OR cost/remission[Title/Abstract] OR cost/remuneration[Title/Abstract] OR cost/resident/day[Title/Abstract] OR cost/resource[Title/Abstract] OR cost/resources[Title/Abstract] OR cost/result[Title/Abstract] OR cost/revenue[Title/Abstract] OR cost/reward[Title/Abstract] OR cost/risk[Title/Abstract] OR cost/risk/benefit[Title/Abstract] OR cost/risk/difficulty[Title/Abstract] OR cost/risks/information[Title/Abstract] OR cost/sacral[Title/Abstract] OR cost/safety[Title/Abstract] OR cost/sample[Title/Abstract] OR cost/savings[Title/Abstract] OR cost/savings/productivity[Title/Abstract] OR cost/senior[Title/Abstract] OR cost/service[Title/Abstract] OR cost/session[Title/Abstract] OR cost/smear[Title/Abstract] OR cost/staff[Title/Abstract] OR cost/staffing[Title/Abstract] OR cost/stay[Title/Abstract] OR cost/std[Title/Abstract] OR cost/student[Title/Abstract] OR cost/student/section[Title/Abstract] OR cost/subject[Title/Abstract] OR cost/success[Title/Abstract] OR cost/successful[Title/Abstract] OR cost/successfully[Title/Abstract] OR cost/supply[Title/Abstract] OR cost/survival[Title/Abstract] OR cost/survivor[Title/Abstract] OR cost/system[Title/Abstract] OR cost/tesla[Title/Abstract] OR cost/test[Title/Abstract] OR cost/the[Title/Abstract] OR cost/therapy[Title/Abstract] OR cost/time[Title/Abstract] OR cost/tny[Title/Abstract] OR cost/treatment[Title/Abstract] OR cost/true[Title/Abstract] OR cost/unit[Title/Abstract] OR cost/ur[Title/Abstract] OR cost/urgent[Title/Abstract] OR cost/usage[Title/Abstract] OR cost/use[Title/Abstract] OR cost/usefulness[Title/Abstract] OR cost/utility[Title/Abstract] OR cost/utilization[Title/Abstract] OR cost/value[Title/Abstract] OR cost/variance[Title/Abstract] OR cost/variceal[Title/Abstract] OR cost/vfa[Title/Abstract] OR cost/visit[Title/Abstract] OR cost/volume[Title/Abstract] OR cost/volume/profit[Title/Abstract] OR cost/watt[Title/Abstract] OR cost/worker[Title/Abstract] OR cost/y/person[Title/Abstract] OR cost/year[Title/Abstract] OR cost/yield[Title/Abstract] OR cost/yls[Title/Abstract] OR cost'[Title/Abstract] OR cost's[Title/Abstract] OR cost'successor'[Title/Abstract] OR cost1[Title/Abstract] OR cost1/effectiveness2[Title/Abstract] OR cost2[Title/Abstract] OR cost281[Title/Abstract] OR cost733[Title/Abstract] OR costa[Title/Abstract] OR costa/pereira[Title/Abstract] OR costa'[Title/Abstract] OR costa's[Title/Abstract] OR costabile[Title/Abstract] OR costabilization[Title/Abstract] OR costabilized[Title/Abstract] OR costabilizer[Title/Abstract] OR costabilizers[Title/Abstract] OR costable[Title/Abstract] OR costacea[Title/Abstract] OR costaceae[Title/Abstract] OR costacervical[Title/Abstract] OR costack[Title/Abstract] OR costacked[Title/Abstract] OR costacking[Title/Abstract] OR costaclavicular[Title/Abstract] OR costaclavine[Title/Abstract] OR costacou[Title/Abstract] OR costada[Title/Abstract] OR costae[Title/Abstract] OR costaeus[Title/Abstract] OR costai[Title/Abstract] OR costain[Title/Abstract] OR costained[Title/Abstract] OR costaining[Title/Abstract] OR costainings[Title/Abstract] OR costains[Title/Abstract] OR costakis[Title/Abstract] OR costal[Title/Abstract] OR costal/crural[Title/Abstract] OR costal1[Title/Abstract] OR costal2[Title/Abstract] OR costal2/fused[Title/Abstract] OR costalat[Title/Abstract] OR costale[Title/Abstract] OR costaleros[Title/Abstract] OR costaleros'[Title/Abstract] OR costales[Title/Abstract] OR costalgia[Title/Abstract] OR costalimai[Title/Abstract] OR costalis[Title/Abstract] OR costall[Title/Abstract] OR costalplasty[Title/Abstract] OR costals[Title/Abstract] OR costalsternal[Title/Abstract] OR costalupes[Title/Abstract] OR costam[Title/Abstract] OR costam/sfrr[Title/Abstract] OR costamagna[Title/Abstract] OR costamere[Title/Abstract] OR costamere'[Title/Abstract] OR costameres[Title/Abstract] OR costameric[Title/Abstract] OR costamerogenesis[Title/Abstract] OR costamers[Title/Abstract] OR costami[Title/Abstract] OR costamp[Title/Abstract] OR costan[Title/Abstract] OR costanchorage[Title/Abstract] OR costand[Title/Abstract] OR costandard[Title/Abstract] OR costane[Title/Abstract] OR costanera[Title/Abstract] OR costanero[Title/Abstract] OR costans[Title/Abstract] OR costant[Title/Abstract] OR costante[Title/Abstract] OR costantin[Title/Abstract] OR costantinescu[Title/Abstract] OR costantini[Title/Abstract] OR costantinii[Title/Abstract] OR costantino[Title/Abstract] OR costantinople[Title/Abstract] OR costantly[Title/Abstract] OR costanza[Title/Abstract] OR costanzi[Title/Abstract] OR costanzo[Title/Abstract] OR costaphrenic[Title/Abstract] OR costar[Title/Abstract] OR costar/corning[Title/Abstract] OR costar/mql[Title/Abstract] OR costar's[Title/Abstract] OR costaras[Title/Abstract] OR costarcuata[Title/Abstract] OR costarelis[Title/Abstract] OR costarelli[Title/Abstract] OR costaria[Title/Abstract] OR costaricae[Title/Abstract] OR costaricaensis[Title/Abstract] OR costarican[Title/Abstract] OR costaricana[Title/Abstract] OR costaricanus[Title/Abstract] OR costaricencis[Title/Abstract] OR costaricense[Title/Abstract] OR costaricensis[Title/Abstract] OR costaricensisis[Title/Abstract] OR costarician[Title/Abstract] OR costaricine[Title/Abstract] OR costarius[Title/Abstract] OR costarrican[Title/Abstract] OR costarricense[Title/Abstract] OR costarricensis[Title/Abstract] OR costarriquense[Title/Abstract] OR costars[Title/Abstract] OR costart[Title/Abstract] OR costarter[Title/Abstract] OR costarum[Title/Abstract] OR costarum'[Title/Abstract] OR costas[Title/Abstract] OR costasiella[Title/Abstract] OR costasis[Title/Abstract] OR costasistrade[Title/Abstract] OR costat[Title/Abstract] OR costata[Title/Abstract] OR costata2[Title/Abstract] OR costata2/133[Title/Abstract] OR costatae[Title/Abstract] OR costate[Title/Abstract] OR costatella[Title/Abstract] OR costational[Title/Abstract] OR costatolide[Title/Abstract] OR costatum[Title/Abstract] OR costatuma[Title/Abstract] OR costatus[Title/Abstract] OR costatus's[Title/Abstract] OR costbenefit[Title/Abstract] OR costbenefits[Title/Abstract] OR costbl[Title/Abstract] OR costc18[Title/Abstract] OR costcervical[Title/Abstract] OR costco[Title/Abstract] OR costconscious[Title/Abstract] OR costcontainment[Title/Abstract] OR costcontrol[Title/Abstract] OR costd17[Title/Abstract] OR coste[Title/Abstract] OR coste's[Title/Abstract] OR costebelle[Title/Abstract] OR costectasis[Title/Abstract] OR costectomy[Title/Abstract] OR costed[Title/Abstract] OR costedoat[Title/Abstract] OR costeff[Title/Abstract] OR costeff's[Title/Abstract] OR costeffective[Title/Abstract] OR costeffectiveness[Title/Abstract] OR costefficiency[Title/Abstract] OR costefficient[Title/Abstract] OR costegites[Title/Abstract] OR costeiros[Title/Abstract] OR costel[Title/Abstract] OR costelic[Title/Abstract] OR costell[Title/Abstract] OR costellatum[Title/Abstract] OR costelli[Title/Abstract] OR costello[Title/Abstract] OR costello's[Title/Abstract] OR costellofly2[Title/Abstract] OR costelytra[Title/Abstract] OR costelytrae[Title/Abstract] OR costen[Title/Abstract] OR costen's[Title/Abstract] OR costenbader[Title/Abstract] OR costenbander[Title/Abstract] OR costeno[Title/Abstract] OR costentin[Title/Abstract] OR costeochondritis[Title/Abstract] OR costeochrondritis[Title/Abstract] OR costep[Title/Abstract] OR coster[Title/Abstract] OR coster's[Title/Abstract] OR costera[Title/Abstract] OR costeras[Title/Abstract] OR costerase[Title/Abstract] OR costermansville[Title/Abstract] OR costero[Title/Abstract] OR costeroid[Title/Abstract] OR costeroidal[Title/Abstract] OR costeroidism[Title/Abstract] OR costeroids[Title/Abstract] OR costerol[Title/Abstract] OR costerols[Title/Abstract] OR costerton[Title/Abstract] OR costerton's[Title/Abstract] OR costertonia[Title/Abstract] OR costes[Title/Abstract] OR costesec[Title/Abstract] OR costesti[Title/Abstract] OR costestimating[Title/Abstract] OR costex[Title/Abstract] OR costful[Title/Abstract] OR costfunctions[Title/Abstract] OR costheta[Title/Abstract] OR costi[Title/Abstract] OR costia[Title/Abstract] OR costiasis[Title/Abstract] OR costiation[Title/Abstract] OR costic[Title/Abstract] OR costica[Title/Abstract] OR costicartilage[Title/Abstract] OR costich[Title/Abstract] OR costicola[Title/Abstract] OR costicolus[Title/Abstract] OR costicosteroid[Title/Abstract] OR costicosteroids[Title/Abstract] OR costicosterone[Title/Abstract] OR costid[Title/Abstract] OR costiform[Title/Abstract] OR costigan[Title/Abstract] OR costil[Title/Abstract] OR costill[Title/Abstract] OR costill/fox[Title/Abstract] OR costill's[Title/Abstract] OR costillas[Title/Abstract] OR costilow[Title/Abstract] OR costim[Title/Abstract] OR costimlulatory[Title/Abstract] OR costimmulating[Title/Abstract] OR costimmulatory[Title/Abstract] OR costimnulation[Title/Abstract] OR costimu[Title/Abstract] OR costimualtory[Title/Abstract] OR costimulant[Title/Abstract] OR costimulants[Title/Abstract] OR costimularory[Title/Abstract] OR costimulary[Title/Abstract] OR costimulatary[Title/Abstract] OR costimulate[Title/Abstract] OR costimulated[Title/Abstract] OR costimulates[Title/Abstract] OR costimulating[Title/Abstract] OR costimulation[Title/Abstract] OR costimulation/activation[Title/Abstract] OR costimulation/adhesion[Title/Abstract] OR costimulation/coinhibition[Title/Abstract] OR costimulation/coreceptor[Title/Abstract] OR costimulation'[Title/Abstract] OR costimulations[Title/Abstract] OR costimulative[Title/Abstract] OR costimulator[Title/Abstract] OR costimulator/b7[Title/Abstract] OR costimulator/b7h[Title/Abstract] OR costimulator/inducible[Title/Abstract] OR costimulator'[Title/Abstract] OR costimulators[Title/Abstract] OR costimulatory[Title/Abstract] OR costimulatory/accessory[Title/Abstract] OR costimulatory/activation[Title/Abstract] OR costimulatory/adhesion[Title/Abstract] OR costimulatory/adjuvant[Title/Abstract] OR costimulatory/coinhibitory[Title/Abstract] OR costimulatory/danger[Title/Abstract] OR costimulatory/induced[Title/Abstract] OR costimulatory/inhibitory[Title/Abstract] OR costimulatory/major[Title/Abstract] OR costimulatory/migratory[Title/Abstract] OR costimulatory/surface[Title/Abstract] OR costimulatory'[Title/Abstract] OR costimulatoty[Title/Abstract] OR costimuli[Title/Abstract] OR costimultory[Title/Abstract] OR costimulus[Title/Abstract] OR costimuratory[Title/Abstract] OR costin[Title/Abstract] OR costine[Title/Abstract] OR costinean[Title/Abstract] OR costineffective[Title/Abstract] OR costinescu[Title/Abstract] OR costinesti[Title/Abstract] OR costing[Title/Abstract] OR costing/productivity[Title/Abstract] OR costing/publication[Title/Abstract] OR costing'[Title/Abstract] OR costing's[Title/Abstract] OR costings[Title/Abstract] OR costiniana[Title/Abstract] OR costinones[Title/Abstract] OR costiosis[Title/Abstract] OR costipation[Title/Abstract] OR costipennis[Title/Abstract] OR costisol[Title/Abstract] OR costituent[Title/Abstract] OR costituents[Title/Abstract] OR costituing[Title/Abstract] OR costituisce[Title/Abstract] OR costituted[Title/Abstract] OR costitution[Title/Abstract] OR costitutional[Title/Abstract] OR costitutionel[Title/Abstract] OR costitutive[Title/Abstract] OR costitutively[Title/Abstract] OR costitutivity[Title/Abstract] OR costituzionale[Title/Abstract] OR costituzioni[Title/Abstract] OR costiveness[Title/Abstract] OR costivenia[Title/Abstract] OR costle[Title/Abstract] OR costleffectiveness[Title/Abstract] OR costless[Title/Abstract] OR costlessly[Title/Abstract] OR costlest[Title/Abstract] OR costlet[Title/Abstract] OR costlets[Title/Abstract] OR costley[Title/Abstract] OR costlier[Title/Abstract] OR costliest[Title/Abstract] OR costliness[Title/Abstract] OR costlow[Title/Abstract] OR costlowi[Title/Abstract] OR costly[Title/Abstract] OR costly/less[Title/Abstract] OR costly/risky/invasive[Title/Abstract] OR costly'[Title/Abstract] OR costmary[Title/Abstract] OR costmedia[Title/Abstract] OR costmetic[Title/Abstract] OR costminimisation[Title/Abstract] OR costminimization[Title/Abstract] OR costn[Title/Abstract] OR costner[Title/Abstract] OR costo[Title/Abstract] OR costoabdominal[Title/Abstract] OR costoapical[Title/Abstract] OR costobrachial[Title/Abstract] OR costobronchogenic[Title/Abstract] OR costocartilage[Title/Abstract] OR costocartilages[Title/Abstract] OR costocartilaginous[Title/Abstract] OR costocentral[Title/Abstract] OR costocervalis[Title/Abstract] OR costocervial[Title/Abstract] OR costocervical[Title/Abstract] OR costocervicalis[Title/Abstract] OR costochondral[Title/Abstract] OR costochondrectomy[Title/Abstract] OR costochondrial[Title/Abstract] OR costochondritic[Title/Abstract] OR costochondritis[Title/Abstract] OR costochondromyelitis[Title/Abstract] OR costochondroral[Title/Abstract] OR costochrondal[Title/Abstract] OR costoclavicular[Title/Abstract] OR costoclavicularis[Title/Abstract] OR costocondral[Title/Abstract] OR costocoracoid[Title/Abstract] OR costocostal[Title/Abstract] OR costocutanei[Title/Abstract] OR costocutaneous[Title/Abstract] OR costodesis[Title/Abstract] OR costodiafragmatic[Title/Abstract] OR costodiaphragamatic[Title/Abstract] OR costodiaphragmatic[Title/Abstract] OR costodiaphragmaticus[Title/Abstract] OR costodiasis[Title/Abstract] OR costodorsal[Title/Abstract]) OR "costs and cost analysis"[MeSH:noexp] OR (cost benefit analyses[Title/Abstract] OR cost benefit analysis[Title/Abstract] OR cost benefit analysts[Title/Abstract]) OR "cost-benefit analysis"[MeSH Terms] OR "health care costs"[MeSH Terms:noexp]

Appendix Three

Siblings Retrieval Performance

RCT siblings

Table 1: Odds Ratio

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Odds ratio
Related Search(PubMed)	1.825	0.11	33.14	2.33	0.31	7.54
Author-Subject(PubMed)	2.307	0.871	0.9	4.16	0	1.65
Citation(Web of Science)	3.285	0	8.8	0	0	2.42
Subject search(e-library)	2.837	0.3	101.4	13.27	40.47	31.66
SCOPUS Author-subject (Group)				11.62	22.68	17.15
SCOPUS Author-subject	129.641	35.632	9.92	0	797.6	194.56
SCOPUS citation	10.704	0.407	0	0	19.16	6.05
CINAHL Author-subject	708.889	26.474	0	302.5	797.6	367.09
CINAHL citation	47.259	0	0	0	0	9.45
WoK-Author-subject	291.692	76.309	15.72	16.33	1246.25	329.26
E-lib - Author-subject	102.325	0	4.35	3.24	7.64	23.51
Economics-Hedges filter	0.562	0	0	0	0	0.11
Costs-Hedges filter	0.639	0	0	0	0	0.13
Qualitative-Hedges filter	2.831	0.169	0	0	12..35	0.75
E-lib-Economics	1.098	0	1.94	0	0	0.61
E-lib-Costs	0.898	0	2.81	0	6.44	2.03
E-lib-Qualitative	0.671	0	1.17	0	2.96	0.96
WoK-Economics	0.765	0	0	0	0	0.15
WoK-Costs	1.528	0	0	0	0	0.31
WoK-Qualitative	3.342	0	0	0	7.21	2.11
Sophisticated2 – Author-Subject	29.876	87.763	40.13	15.82	8.79	36.48
Sophisticated2-Author-Subject-Qual	11.548	70.449	0	33.17	0	23.03
Sophisticated2-Author-Subject- Econ	47.259	0	0	0	0	9.45
Sophisticated2-Author-Subject-Costs	15.728	0	0	0	0	3.145
SCOPUS-Qualitative	98.154	0	0	75.25	3.54	35.39
SCOPUS-Costs	0	0	0	0	12.11	2.42
SCOPUS-Economics	0	0	60.52	0	10.27	14.16
CINAHL-Qualitative	47.259	0	0	24.75	0	14.40
CINAHL-Costs	0	0	0	0	0	0
CINAHL-Economics	47.259	0	0	0	41.46	17.74

Table 2: Precision

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Precision
Related Search(PubMed)	3%	1%	6%	1%	1%	4%
Author-Subject(PubMed)	4%	4%	0%	2%	0%	2%
Citation(Web of Science)	5%	0%	4%	0%	0%	2%
Subject search(e-library)	4%	1%	27%	4%	11%	11%
SCOPUS Author-subject (Group)	0%	0%	0%	4%	11%	27%
SCOPUS Author-subject	31%	25%	4%	0%	100%	32%
SCOPUS citation	13%	2%	0%	0%	9%	7%
CINAHL Author-subject	53%	42%	0%	50%	80%	45%
CINAHL citation	14%	0%	0%	0%	0%	3%
WoK-Author-subject	32%	38%	7%	5%	83%	33%
E-lib - Author-subject	31%	0%	2%	1%	5%	8%
Economics-Hedges filter	1%	0%	0%	0%	0%	0%
Costs-Hedges filter	1%	0%	0%	0%	0%	0%
Qualitative-Hedges filter	5%	1%	0%	0%	9%	5%
E-lib-Economics	2%	0%	1%	0%	0%	1%
E-lib-Costs	2%	0%	1%	0%	5%	2%
E-lib-Qualitative	2%	0%	1%	0%	2%	1%
WoK-Economics	2%	0%	0%	0%	0%	0%
WoK-Costs	3%	0%	0%	0%	0%	1%
WoK-Qualitative	5%	0%	0%	0%	6%	4%
Sophisticated2 – Author-Subject	5%	53%	11%	3%	7%	24%
Sophisticated2-Author-Subject-Qual	12%	50%	0%	9%	0%	24%
Sophisticated2-Author-Subject- Econ	17%	0%	0%	0%	0%	23%
Sophisticated2-Author-Subject-Costs	14%	0%	0%	0%	0%	23%
SCOPUS-Qualitative	14%	0%	0%	17%	6%	7%
SCOPUS-Costs	0%	0%	0%	0%	8%	2%
SCOPUS-Economics	0%	0%	25%	0%	8%	7%
CINAHL-Qualitative	17%	0%	0%	7%	0%	5%
CINAHL-Costs	0%	0%	0%	0%	0%	0%
CINAHL-Economics	17%	0%	0%	0%	25%	8%

Table 3: Recall

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Recall
Related Search(PubMed)	9%	1%	46%	13%	31%	20%
Author-Subject(PubMed)	9%	14%	23%	13%	0%	12%
Citation(Web of Science)	4%	0%	15%	0%	0%	4%
Subject search(e-library)	17%	4%	23%	13%	46%	21%
SCOPUS Author-subject (Group)				13%	31%	22%
SCOPUS Author-subject	23%	51%	15%	0%	31%	24%
SCOPUS citation	11%	1%	0%	0%	31%	9%
CINAHL Author-subject	14%	15%	0%	13%	31%	15%
CINAHL citation	1%	0%	0%	0%	0%	0%
WoK-Author-subject	29%	51%	8%	13%	39%	28%
E-lib - Author-subject	24%	0%	8%	13%	15%	12%
Economics-Hedges filter	1%	0%	0%	0%	0%	0%
Costs-Hedges filter	1%	0%	0%	0%	0%	0%
Qualitative-Hedges filter	6%	1%	0%	0%	8%	3%
E-lib-Economics	7%	0%	15%	0%	0%	4%
E-lib-Costs	6%	0%	15%	0%	8%	6%
E-lib-Qualitative	6%	0%	8%	0%	15%	6%
WoK-Economics	1%	0%	0%	0%	0%	0%
WoK-Costs	4%	0%	0%	0%	0%	1%
WoK-Qualitative	7%	0%	0%	0%	8%	3%
Sophisticated2 – Author-Subject	7%	23%	31%	25%	8%	19%
Sophisticated2-Author-Subject-Qual	4%	14%	0%	13%	0%	6%
Sophisticated2-Author-Subject- Econ	1%	0%	0%	0%	0%	0%
Sophisticated2-Author-Subject-Costs	1%	0%	0%	0%	0%	0%
SCOPUS-Qualitative	3%	0%	0%	13%	8%	5%
SCOPUS-Costs	0%	0%	0%	0%	15%	3%
SCOPUS-Economics	0%	0%	8%	0%	8%	3%
CINAHL-Qualitative	1%	0%	0%	13%	0%	3%
CINAHL-Costs	0%	0%	0%	0%	0%	0%
CINAHL-Economics	1%	0%	0%	0%	8%	2%

Qualitative siblings

Table 4: Odds Ratio

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Odds ratio
Related Search(PubMed)	2.51	0.325	0	0	0	0.57
Author-Subject(PubMed)	7.049	0.416	0	0	0	1.49
Citation(Web of Science)	8.052	1.347	0	0	0	1.88
Subject search(e-library)	3.152	0	0	13.27	0	3.28
SCOPUS Author-subject (Group)				0	0	0
SCOPUS Author-subject	140.444	3.959	0	0	0	28.88
SCOPUS citation	12.488	0	0	0	0	2.50
CINAHL Author-subject	283.556	0	0	0	0	56.71
CINAHL citation	60.762	0	0	0	0	12.15
WoK-Author-subject	208.352	5.451	0	65.33	0	55.83
E-lib - Author-subject	66.211	5.772	0	3.24	0	15.04
Economics-Hedges filter	1.516	0	0	0	0	0.30
Costs-Hedges filter	0.821	0	0	0	0	0.16
Qualitative-Hedges filter	6.37	0	0	50	0	11.27
E-lib-Economics	2.358	0	0	0	0	0.47
E-lib-Costs	1.197	0	0	0	0	0.24
E-lib-Qualitative	1.185	0	0	0	0	0.24
WoK-Economics	0	0	0	0	0	0
WoK-Costs	0	0	0	0	0	0
WoK-Qualitative	5.764	0	0	200	0	41.15
Sophisticated2 – Author-Subject	28.727	38.818	0	3.96	0	14.30
Sophisticated2-Author-Subject-Qual	21.385	33.623	0	33.17	0	17.64
Sophisticated2-Author-Subject- Econ	60.762	0	0	0	0	12.15
Sophisticated2-Author-Subject-Costs	0	0	0	0	0	0
SCOPUS-Qualitative	729.143	0	0	75.25	0	160.88
SCOPUS-Costs	0	0	0	0	0	0
SCOPUS-Economics	0	0	0	0	0	0
CINAHL-Qualitative	283.556	0	0	24.75	0	61.66
CINAHL-Costs	0	0	0	0	0	0
CINAHL-Economics	60.762	0	0	27.05	0	17.56

Table 5: Precision

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Precision
Related Search(PubMed)	3%	1%	0%	0%	0%	1%
Author-Subject(PubMed)	6%	1%	0%	0%	0%	1%
Citation(Web of Science)	9%	2%	0%	0%	0%	2%
Subject search(e-library)	3%	0%	0%	4%	0%	1%
SCOPUS Author-subject (Group)				0%	0%	0%
SCOPUS Author-subject	25%	3%	0%	0%	0%	6%
SCOPUS citation	11%	0%	0%	0%	0%	2%
CINAHL Author-subject	21%	0%	0%	0%	0%	4%
CINAHL citation	14%	0%	0%	0%	0%	3%
WoK-Author-subject	23%	3%	0%	10%	0%	7%
E-lib - Author-subject	20%	7%	0%	1%	0%	6%
Economics-Hedges filter	2%	0%	0%	0%	0%	0%
Costs-Hedges filter	1%	0%	0%	0%	0%	0%
Qualitative-Hedges filter	7%	0%	0%	14%	0%	4%
E-lib-Economics	3%	0%	0%	0%	0%	1%
E-lib-Costs	2%	0%	0%	0%	0%	0%
E-lib-Qualitative	2%	0%	0%	0%	0%	0%
WoK-Economics	0%	0%	0%	0%	0%	0%
WoK-Costs	0%	0%	0%	0%	0%	0%
WoK-Qualitative	6%	0%	0%	25%	0%	6%
Sophisticated2 – Author-Subject	2%	10%	0%	1%	0%	3%
Sophisticated2-Author-Subject-Qual	15%	9%	0%	9%	0%	7%
Sophisticated2-Author-Subject- Econ	17%	0%	0%	0%	0%	3%
Sophisticated2-Author-Subject-Costs	0%	0%	0%	0%	0%	0%
SCOPUS-Qualitative	57%	0%	0%	17%	0%	15%
SCOPUS-Costs	0%	0%	0%	0%	0%	0%
SCOPUS-Economics	0%	0%	0%	0%	0%	0%
CINAHL-Qualitative	67%	0%	0%	7%	0%	15%
CINAHL-Costs	0%	0%	0%	0%	0%	0%
CINAHL-Economics	17%	0%	0%	8%	0%	5%

Table 6: Recall

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Recall
Related Search(PubMed)	9%	1%	0%	0%	0%	2%
Author-Subject(PubMed)	14%	3%	0%	0%	0%	3%
Citation(Web of Science)	7%	1%	0%	0%	0%	2%
Subject search(e-library)	14%	0%	0%	13%	0%	5%
SCOPUS Author-subject (Group)				0%	0%	0%
SCOPUS Author-subject	19%	6%	0%	0%	0%	5%
SCOPUS citation	10%	0%	0%	0%	0%	2%
CINAHL Author-subject	6%	0%	0%	0%	0%	1%
CINAHL citation	1%	0%	0%	0%	0%	0%
WoK-Author-subject	20%	5%	0%	25%	0%	10%
E-lib - Author-subject	16%	3%	0%	13%	0%	6%
Economics-Hedges filter	3%	0%	0%	0%	0%	1%
Costs-Hedges filter	1%	0%	0%	0%	0%	0%
Qualitative-Hedges filter	9%	0%	0%	13%	0%	4%
E-lib-Economics	10%	0%	0%	0%	0%	2%
E-lib-Costs	6%	0%	0%	0%	0%	1%
E-lib-Qualitative	7%	0%	0%	0%	0%	1%
WoK-Economics	0%	0%	0%	0%	0%	0%
WoK-Costs	0%	0%	0%	0%	0%	0%
WoK-Qualitative	9%	0%	0%	25%	0%	7%
Sophisticated2 – Author-Subject	3%	5%	0%	13%	0%	4%
Sophisticated2-Author-Subject-Qual	6%	3%	0%	13%	0%	4%
Sophisticated2-Author-Subject- Econ	1%	0%	0%	0%	0%	0%
Sophisticated2-Author-Subject-Costs	0%	0%	0%	0%	0%	0%
SCOPUS-Qualitative	11%	0%	0%	13%	0%	5%
SCOPUS-Costs	0%	0%	0%	0%	0%	0%
SCOPUS-Economics	0%	0%	0%	0%	0%	0%
CINAHL-Qualitative	6%	0%	0%	13%	0%	4%
CINAHL-Costs	0%	0%	0%	0%	0%	0%
CINAHL-Economics	1%	0%	0%	13%	0%	3%

Economics siblings

Table 7: Odds Ratio

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Odds ratio
Related Search(PubMed)	2.23	0	0	0	0	0.45
Author-Subject(PubMed)	25.38	0	1.8	0	26.67	10.77
Citation(Web of Science)	27.38	0	0	0	0	5.48
Subject search(e-library)	3.78	0	0	0	0	0.76
SCOPUS Author-subject (Group)				0	0	0.00
SCOPUS Author-subject	97.23	0	0	0	0	19.45
SCOPUS citation	26.76	0	0	0	0	5.35
CINAHL Author-subject	425.33	0	0	0	0	85.07
CINAHL citation	0	0	0	0	0	0.00
WoK-Author-subject	291.69	0	0	0	0	58.34
E-lib - Author-subject	22.07	0	0	0	0	4.41
Economics-Hedges filter	5.06	0	0	0	0	1.01
Costs-Hedges filter	51.73	0	0	0	0	10.35
Qualitative-Hedges filter	0	0	0	0	0	0.00
E-lib-Economics	0	0	0	0	0	0.00
E-lib-Costs	5.39	0	0	0	0	1.08
E-lib-Qualitative	12.08	0	9.33	0	0	4.28
WoK-Economics	20.64	0	0	0	0	4.13
WoK-Costs	12.73	0	0	0	0	2.55
WoK-Qualitative	0	0	0	0	0	0.00
Sophisticated2 – Author-Subject	12.17	0	0	0	0	2.43
Sophisticated2-Author-Subject-Qual	32.08	0	0	0	0	6.42
Sophisticated2-Author-Subject- Econ	1276.00	0	0	0	0	255.20
Sophisticated2-Author-Subject-Costs	424.67	0	0	0	0	84.93
SCOPUS-Qualitative	0	0	0	0	0	0.00
SCOPUS-Costs	1276.00	0	0	0	0	255.20
SCOPUS-Economics	637.50	0	0	0	0	127.50
CINAHL-Qualitative	0	0	0	0	0	0.00
CINAHL-Costs	3828.00	0	0	0	0	765.60
CINAHL-Economics	3828.00	0	0	0	0	765.60

Table 8: Precision

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Precision
Related Search(PubMed)	1%	0%	0%	0%	0%	0%
Author-Subject(PubMed)	2%	0%	0%	0%	3%	1%
Citation(Web of Science)	4%	0%	0%	0%	0%	1%
Subject search(e-library)	1%	0%	0%	0%	0%	0%
SCOPUS Author-subject (Group)				0%	0%	0%
SCOPUS Author-subject	4%	0%	0%	0%	0%	1%
SCOPUS citation	3%	0%	0%	0%	0%	1%
CINAHL Author-subject	5%	0%	0%	0%	0%	1%
CINAHL citation	0%	0%	0%	0%	0%	0%
WoK-Author-subject	5%	0%	0%	0%	0%	1%
E-lib - Author-subject	2%	0%	0%	0%	0%	0%
Economics-Hedges filter	1%	0%	0%	0%	0%	0%
Costs-Hedges filter	4%	0%	0%	0%	0%	1%
Qualitative-Hedges filter	0%	0%	0%	0%	0%	0%
E-lib-Economics	0%	0%	0%	0%	0%	0%
E-lib-Costs	1%	0%	0%	0%	0%	0%
E-lib-Qualitative	1%	0%	1%	0%	0%	0%
WoK-Economics	3%	0%	0%	0%	0%	1%
WoK-Costs	2%	0%	0%	0%	0%	0%
WoK-Qualitative	0%	0%	0%	0%	0%	0%
Sophisticated2 – Author-Subject	2%	0%	0%	0%	0%	0%
Sophisticated2-Author-Subject-Qual	4%	0%	0%	0%	0%	1%
Sophisticated2-Author-Subject- Econ	33%	0%	0%	0%	0%	7%
Sophisticated2-Author-Subject-Costs	29%	0%	0%	0%	0%	6%
SCOPUS-Qualitative	0%	0%	0%	0%	0%	0%
SCOPUS-Costs	50%	0%	0%	0%	0%	10%
SCOPUS-Economics	50%	0%	0%	0%	0%	10%
CINAHL-Qualitative	0%	0%	0%	0%	0%	0%
CINAHL-Costs	75%	0%	0%	0%	0%	15%
CINAHL-Economics	50%	0%	0%	0%	0%	10%

Table 9: Recall

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Recall
Related Search(PubMed)	1%	0%	0%	0%	0%	0%
Author-Subject(PubMed)	4%	0%	8%	0%	8%	4%
Citation(Web of Science)	3%	0%	0%	0%	0%	1%
Subject search(e-library)	3%	0%	0%	0%	0%	1%
SCOPUS Author-subject (Group)				0%	0%	0%
SCOPUS Author-subject	3%	0%	0%	0%	0%	1%
SCOPUS citation	3%	0%	0%	0%	0%	1%
CINAHL Author-subject	1%	0%	0%	0%	0%	0%
CINAHL citation	0%	0%	0%	0%	0%	0%
WoK-Author-subject	4%	0%	0%	0%	0%	1%
E-lib - Author-subject	1%	0%	0%	0%	0%	0%
Economics-Hedges filter	1%	0%	0%	0%	0%	0%
Costs-Hedges filter	4%	0%	0%	0%	0%	1%
Qualitative-Hedges filter	0%	0%	0%	0%	0%	0%
E-lib-Economics	0%	0%	0%	0%	0%	0%
E-lib-Costs	3%	0%	0%	0%	0%	1%
E-lib-Qualitative	4%	0%	8%	0%	0%	2%
WoK-Economics	3%	0%	0%	0%	0%	1%
WoK-Costs	3%	0%	0%	0%	0%	1%
WoK-Qualitative	0%	0%	0%	0%	0%	0%
Sophisticated2 – Author-Subject	3%	0%	0%	0%	0%	1%
Sophisticated2-Author-Subject-Qual	1%	0%	0%	0%	0%	0%
Sophisticated2-Author-Subject- Econ	3%	0%	0%	0%	0%	1%
Sophisticated2-Author-Subject-Costs	3%	0%	0%	0%	0%	1%
SCOPUS-Qualitative	0%	0%	0%	0%	0%	0%
SCOPUS-Costs	3%	0%	0%	0%	0%	1%
SCOPUS-Economics	3%	0%	0%	0%	0%	1%
CINAHL-Qualitative	0%	0%	0%	0%	0%	0%
CINAHL-Costs	4%	0%	0%	0%	0%	1%
CINAHL-Economics	4%	0%	0%	0%	0%	1%

Process evaluation siblings

Table 10: Odds Ratio

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Odds ratio
Related Search(PubMed)	5.205	0.6	8.31	0	0.19	2.86
Author-Subject(PubMed)	1.952	1.22	0.9	8.32	53.4 4	13.17
Citation(Web of Science)	3.911	0	15.41	0	0	3.86
Subject search(e-library)	1.719	0	0	0	0	0.34
SCOPUS Author-subject (Group)				0	0	0.00
SCOPUS Author-subject	97.231	37.329	17.37	0	0	30.39
SCOPUS citation	1.784	0	0	0	0	0.36
CINAHL Author-subject	425.333	34.26	0	0	0	91.92
CINAHL citation	85.067	0	0	0	0	17.01
WoK-Author-subject	162.051	54.507	0	0	0	43.31
E-lib - Author-subject	51.497	0	0	0	0	10.30
Economics-Hedges filter	2.166	0	0	0	0	0.43
Costs-Hedges filter	2.463	0	0	0	0	0.49
Qualitative-Hedges filter	5.662	0	0	0	0	1.13
E-lib-Economics	3.031	0	0	0	0	0.61
E-lib-Costs	1.795	0	0	0	0	0.36
E-lib-Qualitative	2.417	0	4.66	0	0	1.42
WoK-Economics	0	0	0	0	0	0
WoK-Costs	0.849	0	0	0	0	0.17
WoK-Qualitative	6.987	0	0	100	0	21.40
Sophisticated2 – Author-Subject	2.807	155.273	25.08	0	0	36.63
Sophisticated2-Author-Subject-Qual	43.741	184.929	101.81	0	0	66.10
Sophisticated2-Author-Subject- Econ	85.067	0	0	0	0	17.01
Sophisticated2-Author-Subject-Costs	28.311	0	0	0	0	5.66
SCOPUS-Qualitative	425.333	0	0	0	0	85.07
SCOPUS-Costs	85.067	0	0	0	0	17.01
SCOPUS-Economics	0	0	0	0	0	0
CINAHL-Qualitative	85.067	0	0	0	0	17.01
CINAHL-Costs	85.067	0	0	0	0	17.01
CINAHL-Economics	85.067	0	0	0	0	17.01

Table 11: Precision

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Precision
Related Search(PubMed)	4%	1%	1%	0%	0%	1%
Author-Subject(PubMed)	2%	1%	0%	2%	5%	2%
Citation(Web of Science)	4%	0%	2%	0%	0%	1%
Subject search(e-library)	2%	0%	0%	0%	0%	0%
SCOPUS Author-subject (Group)				0%	0%	0
SCOPUS Author-subject	15%	5%	2%	0%	0%	4%
SCOPUS citation	2%	0%	0%	0%	0%	0%
CINAHL Author-subject	21%	10%	0%	0%	0%	6%
CINAHL citation	14%	0%	0%	0%	0%	3%
WoK-Author-subject	15%	7%	0%	0%	0%	4%
E-lib - Author-subject	13%	0%	0%	0%	0%	3%
Economics-Hedges filter	2%	0%	0%	0%	0%	0%
Costs-Hedges filter	3%	0%	0%	0%	0%	1%
Qualitative-Hedges filter	5%	0%	0%	0%	17%	1%
E-lib-Economics	3%	0%	0%	0%	0%	1%
E-lib-Costs	2%	0%	0%	0%	0%	0%
E-lib-Qualitative	2%	0%	1%	0%	2%	1%
WoK-Economics	0%	0%	0%	0%	0%	0%
WoK-Costs	1%	0%	0%	0%	0%	0%
WoK-Qualitative	5%	0%	0%	13%	17%	4%
Sophisticated2 – Author-Subject	3%	14%	3%	0%	0%	4%
Sophisticated2-Author-Subject-Qual	19%	19%	11%	0%	0%	10%
Sophisticated2-Author-Subject- Econ	17%	0%	0%	0%	0%	3%
Sophisticated2-Author-Subject-Costs	14%	0%	0%	0%	0%	3%
SCOPUS-Qualitative	29%	0%	0%	0%	0%	6%
SCOPUS-Costs	25%	0%	0%	0%	0%	5%
SCOPUS-Economics	0%	0%	0%	0%	0%	0%
CINAHL-Qualitative	17%	0%	0%	0%	0%	3%
CINAHL-Costs	25%	0%	0%	0%	0%	5%
CINAHL-Economics	17%	0%	0%	0%	0%	3%

Table 12: Recall

	IDEATel	Tamoxifen	IVF	BMS	CLD	Average Recall
Related Search(PubMed)	10%	1%	8%	0%	8%	5%
Author-Subject(PubMed)	4%	4%	8%	13%	15%	9%
Citation(Web of Science)	3%	0%	8%	0%	0%	2%
Subject search(e-library)	7%	0%	0%	0%	0%	1%
SCOPUS Author-subject (Group)				0%	0%	0%
SCOPUS Author-subject	11%	10%	8%	0%	0%	6%
SCOPUS citation	1%	0%	0%	0%	0%	0%
CINAHL Author-subject	6%	4%	0%	0%	0%	2%
CINAHL citation	1%	0%	0%	0%	0%	0%
WoK-Author-subject	13%	9%	0%	0%	0%	4%
E-lib - Author-subject	10%	0%	0%	0%	0%	2%
Economics-Hedges filter	3%	0%	0%	0%	0%	1%
Costs-Hedges filter	3%	0%	0%	0%	0%	1%
Qualitative-Hedges filter	6%	0%	0%	0%	0%	1%
E-lib-Economics	9%	0%	0%	0%	0%	2%
E-lib-Costs	6%	0%	0%	0%	0%	1%
E-lib-Qualitative	9%	0%	8%	0%	0%	3%
WoK-Economics	0%	0%	0%	0%	0%	0%
WoK-Costs	1%	0%	0%	0%	0%	0%
WoK-Qualitative	7%	0%	0%	13%	0%	4%
Sophisticated2 – Author-Subject	4%	6%	8%	0%	0%	4%
Sophisticated2-Author-Subject-Qual	7%	5%	8%	0%	0%	4%
Sophisticated2-Author-Subject- Econ	1%	0%	0%	0%	0%	0%
Sophisticated2-Author-Subject-Costs	1%	0%	0%	0%	0%	0%
SCOPUS-Qualitative	6%	0%	0%	0%	0%	1%
SCOPUS-Costs	1%	0%	0%	0%	0%	0%
SCOPUS-Economics	0%	0%	0%	0%	0%	0%
CINAHL-Qualitative	1%	0%	0%	0%	0%	0%
CINAHL-Costs	1%	0%	0%	0%	0%	0%
CINAHL-Economics	1%	0%	0%	0%	0%	0%

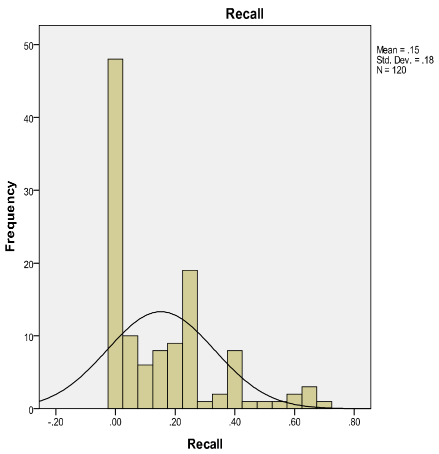
Appendix Four

Normality test

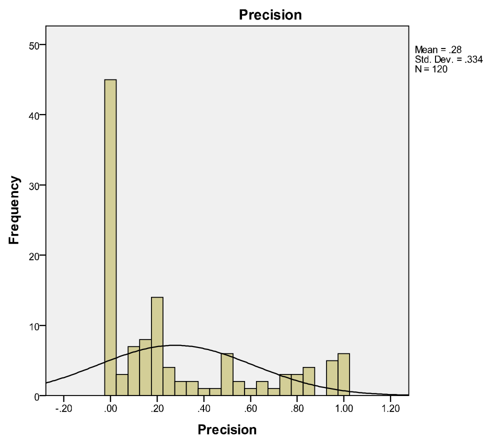
Normality tests were run to confirm whether the data does or do not follow normal distribution. It helps to decide on which statistical test will be appropriate for hypotheses testing. Normality was tested visually using histograms and again it was quantified using numbers.

Dataset1

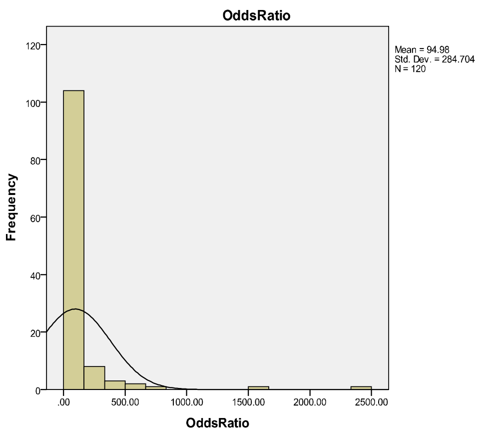
Recall



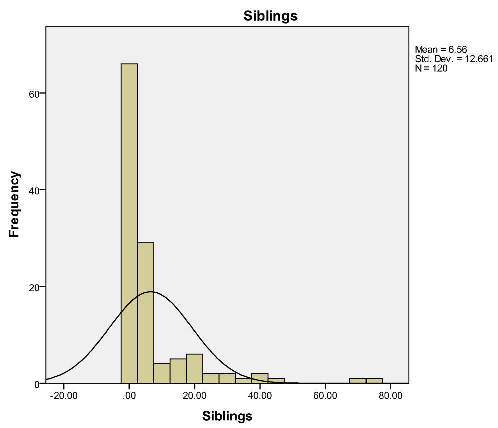
Precision



Odds Estimator



Number of Siblings Retrieved



The important measures here are skewness and kurtosis, where the value of both skewness and kurtosis should be zero in normal distribution, where the further the value from zero the more likely that the data are not normally distributed. According to the table the skewness is 1.238 and kurtosis is 0.961 which indicate the data in dataset 1 are not normally distributed, the same as was predicted from the visual test.

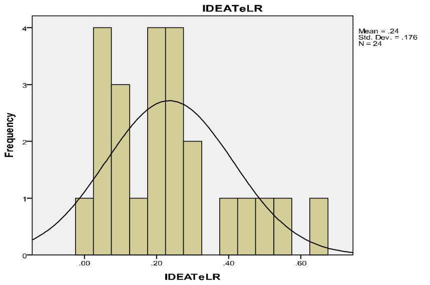
Table 1: Normality test

Statistics

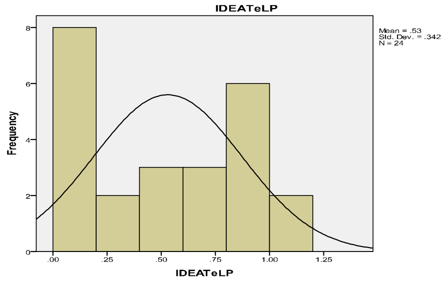
		Recall	Precision	OddsRatio	Siblings
N	Valid	120	120	120	120
	Missing	204	204	204	204
Mean		.1520	.2776	94.9798	6.5583
Std. Error of Mean		.01641	.03050	25.98981	1.15578
Median		.0800	.1500	5.6050	2.0000
Mode		.00	.00	.00	.00
Std. Deviation		.17980	.33406	284.70410	12.66093
Variance		.032	.112	81056.427	160.299
Skewness		1.238	1.030	5.840	3.223
Std. Error of Skewness		.221	.221	.221	.221
Kurtosis		.961	-.363	39.801	12.183
Std. Error of Kurtosis		.438	.438	.438	.438
Range		.69	1.00	2348.33	74.00
Minimum		.00	.00	.00	.00
Maximum		.69	1.00	2348.33	74.00
Percentiles	25	.0000	.0000	.0000	.0000
	50	.0800	.1500	5.6050	2.0000
	75	.2500	.5000	57.0000	5.7500

DATASET2

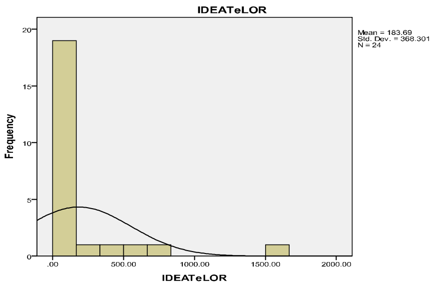
IDEATeL Recall



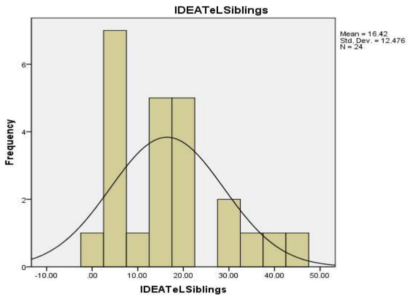
IDEATeL Precision



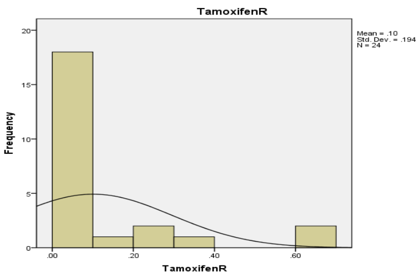
IDEATeLOdds Estimator



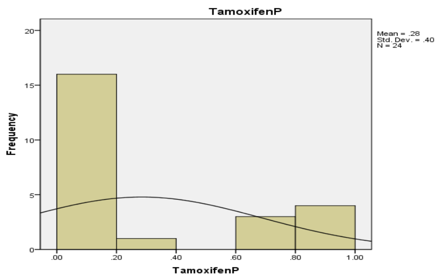
IDEATeL Siblings



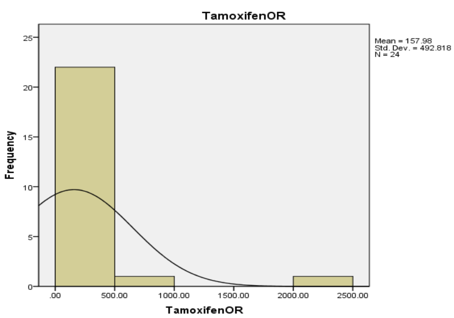
Tamoxifen Recall



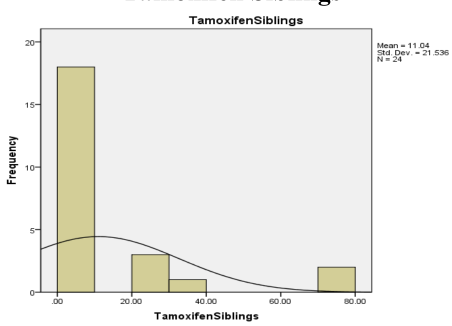
Tamoxifen Precision



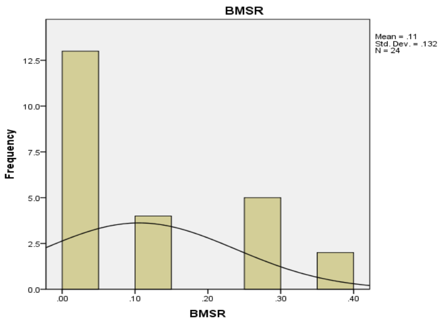
Tamoxifen Odds Estimator



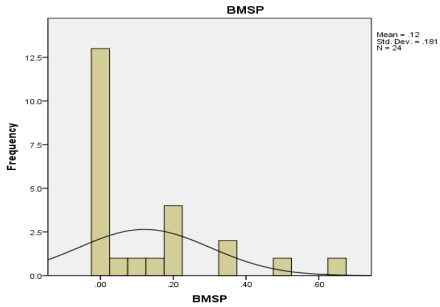
Tamoxifen Siblings



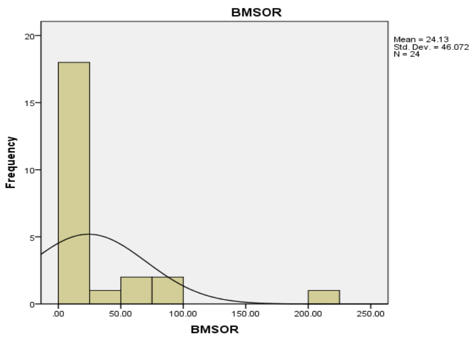
BMS Recall



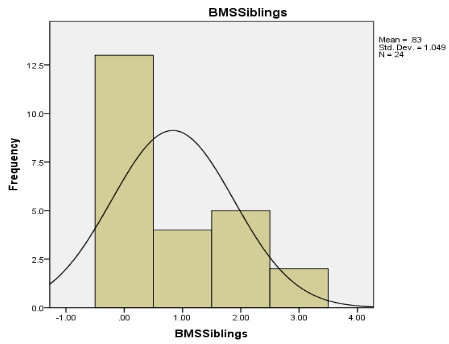
BMS Precision



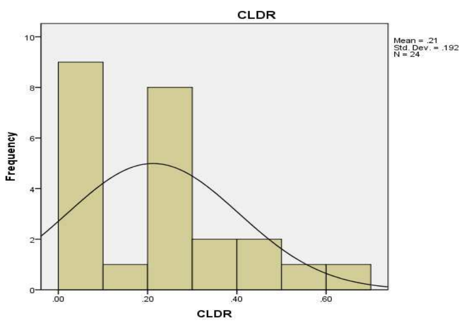
BMS Odds Estimator



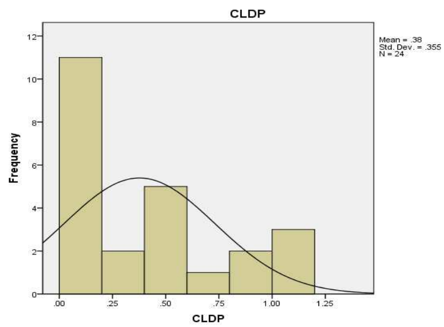
BMS Siblings



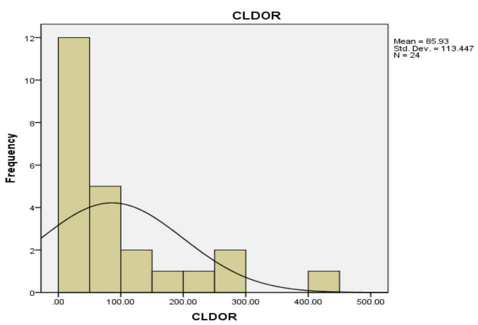
CLD Recall



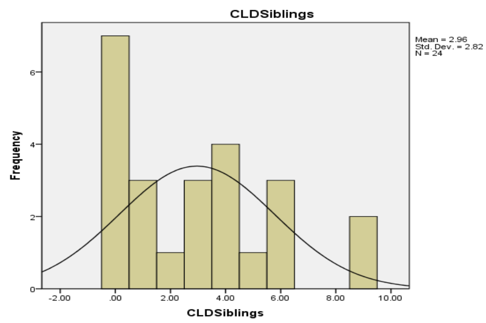
CLD Precision



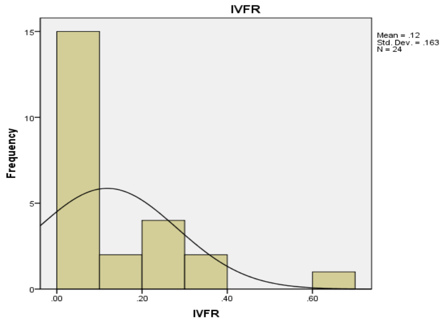
CLD Odds Estimator



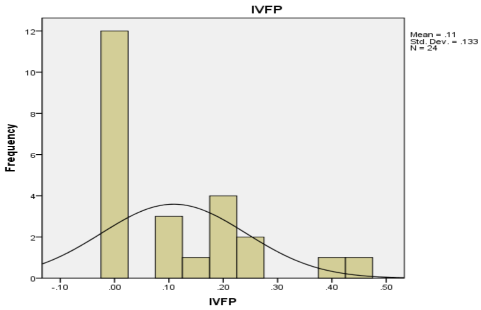
CLD Siblings



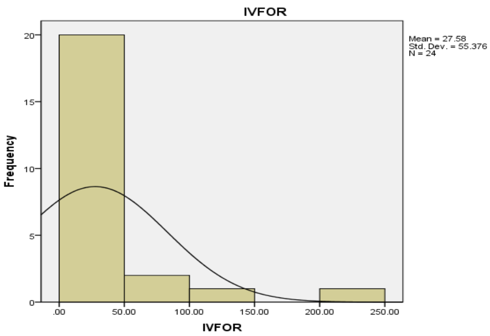
IVF Recall



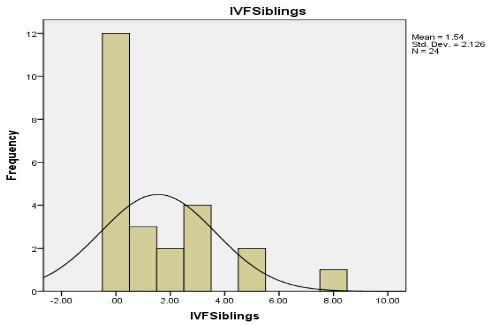
IVF Precision



IVF Odds Estimator



IVF Siblings



According to Tables 2 and 3 the skewness and kurtosis values were very much away from the zero indicating that the data in dataset 2 are not normally distributed as well, supporting the prediction from the visual test. And therefore, non-parametric test were used to test the research hypotheses.

Table 2: Normality Test for Dataset 2.

		IDEATeLR	IDEATeLP	IDEATeLOR	IDEATeL Siblings	TamoxifenR	TamoxifenP	TamoxifenOR	Tamoxifen Siblings
N	Valid	24	24	24	24	24	24	24	24
Mean		.2358	.5304	183.6921	16.4167	.1004	.2846	157.9817	11.0417
Std. Error of Mean		.03598	.06985	75.17904	2.54661	.03967	.08163	100.59607	4.39593
Median		.2000	.5650	29.9500	15.0000	.0050	.0150	.2250	.5000
Mode		.20	.19[a]	.00[a]	3.00	.00	.00	.00	.00
Std. Deviation		.17626	.34217	368.30058	12.47577	.19437	.39990	492.81807	21.53557
Variance		.031	.117	135645.315	155.645	.038	.160	242869.647	463.781
Skewness		.872	-.082	3.106	.817	2.226	.963	4.191	2.230
Std. Error of Skewness		.472	.472	.472	.472	.472	.472	.472	.472
Kurtosis		.167	-1.624	10.856	.045	4.291	-.978	18.672	4.296
Std. Error of Kurtosis		.918	.918	.918	.918	.918	.918	.918	.918
Range		.66	1.00	1643.03	46.00	.67	.99	2348.33	74.00
Minimum		.00	.00	.00	.00	.00	.00	.00	.00
Maximum		.66	1.00	1643.03	46.00	.67	.99	2348.33	74.00
Percentiles	25	.0900	.1825	4.9800	5.2500	.0000	.0000	.0000	.0000
	50	.2000	.5650	29.9500	15.0000	.0050	.0150	.2250	.5000
	75	.3050	.8375	129.9250	21.5000	.1525	.7725	31.4550	16.7500

(Continuous)

Table 2: (Continued)

		BMSR	BMSP	BMS OR	BMS Siblings	CLDR	CLDP	CLDOR	CLD Siblings	IVFR	IVFP	IVFOR	IVF Siblings
N	Valid	24	24	24	24	24	24	24	24	24	24	24	24
Mean		.1054	.1208	24.1312	.8333	.2121	.3767	85.9333	2.9583	.1183	.1088	27.5771	1.5417
Std. Error of Mean		.02699	.03700	9.40450	.21423	.03915	.07243	23.15719	.57571	.03332	.02722	11.30366	.43397
Median		.0000	.0000	.0000	.0000	.2000	.2250	43.5750	3.0000	.0400	.0500	.3750	.5000
Mode		.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Std. Deviation		.13224	.18125	46.07245	1.04950	.19179	.35481	113.44659	2.82041	.16325	.13333	55.37640	2.12601
Variance		.017	.033	2122.670	1.101	.037	.126	12870.129	7.955	.027	.018	3066.546	4.520
Skewness		.848	1.726	2.941	.853	.819	.570	1.536	.706	1.610	1.049	3.200	1.581
Std. Error of Skewness		.472	.472	.472	.472	.472	.472	.472	.472	.472	.472	.472	.472
Kurtosis		-.636	2.760	10.051	-.661	.394	-1.038	1.629	-.290	2.583	.339	11.557	2.402
Std. Error of Kurtosis		.918	.918	.918	.918	.918	.918	.918	.918	.918	.918	.918	.918
Range		.38	.67	204.00	3.00	.69	1.00	404.00	9.00	.62	.44	249.12	8.00
Minimum		.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Maximum		.38	.67	204.00	3.00	.69	1.00	404.00	9.00	.62	.44	249.12	8.00
Percentiles	25	.0000	.0000	.0000	.0000	.0175	.0225	.0000	.0000	.0000	.0000	.0000	.0000
	50	.0000	.0000	.0000	.0000	.2000	.2250	43.5750	3.0000	.0400	.0500	.3750	.5000
	75	.2500	.2000	28.5975	2.0000	.3150	.6225	130.9425	4.7500	.2300	.2075	40.8300	3.0000

Appendix Five

Pilot Study Retrieval Performance

Table 1: The IDEATeL Search and Retrieval Performance

Diabetes – Telemedicine								
search strategy	**Relevant/R (A)**	**Relevant N/R©**	**Not relevant (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Ratio**
PubMed-Related articles	25	97	161	0.21	21%	0.13	13%	0.8
PubMed-Author-subject	23	99	134	0.19	19%	0.15	15%	0.9
Citation(Web of Science)	14	108	43	0.12	12%	0.25	25%	1.6
Subject search (e-library)	39	83	257	0.38	38%	0.13	13%	1.2
SCOPUS Author-subject (Author(s))	39	83	13	0.32	32%	0.75	75%	18.0
SCOPUS citation	32	90	32	0.26	26%	0.50	50%	5.6
CINAHL Author-subject	17	105	1	0.14	14%	0.94	94%	84.4
CINAHL citation	4	118	3	0.03	3%	0.57	57%	6.0
WoK-Author-subject	55	67	7	0.45	45%	0.89	89%	56.6
E-lib – Author-subject	38	84	17	0.31	31%	0.69	69%	13.3
Total relevant R without duplicates						**122**		
Total retrieved without duplicate						**660**		
Total non-relevant R without duplicates (D)						**538**		

Table 2: IDEATeL Study Search Strategies Retrieval per Study Type and OR Calculations

search strategy	***RCTs***			***Qualitative***			***Economical evaluation***			***Process evaluation***		
	RR	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**
PubMed-Related articles	9	42	0.58	9	29	0.84	3	8	1.03	4	18	0.61
PubMed-Author-subject	8	43	0.60	9	29	1.00	2	9	0.73	4	18	0.72
Citation(Web of Science)	4	47	0.78	8	30	2.45	2	9	2.04	0	22	0
Subject search (e-library)	26	25	1.57	18	20	1.38	9	2	7.19	6	16	0.59
SCOPUS Author-subject (Author(s))	17	34	4.64	12	26	4.32	2	9	2.12	8	14	5.40
SCOPUS citation	19	28	4.52	8	23	2.08	2	9	1.75	3	14	1.24
CINAHL Author-subject	10	37	6.92	4	27	3.38	1	10	2.89	2	15	2.88
CINAHL citation	1	46	1.52	3	28	6.50	0	11	0.00	0	17	0
WoK-Author-subject	20	31	4.82	22	16	10.22	3	8	2.90	10	12	6.36
E-lib – Author-subject	18	33	4.78	13	25	4.60	1	10	0.91	6	16	3.37

Table 3: Tamoxifen: Search and Retrieval Performance

Breast Cancer (Tamoxifen)								
search strategy	**Relevant/R**	**Relevant**	**Not relevant**	**Recall**		**Precision**		**Odds**
	(A)	**N/R©**	**(B)**	**Value**	**%**	**Value**	**%**	**Ratio**
PubMed-Related articles	18	168	182	0.10	10%	0.09	9%	0.53
PubMed-Author-subject	17	169	434	0.09	9%	0.04	4%	0.21
Citation(Web of Science)	5	181	48	0.03	3%	0.09	9%	0.52
Subject search (e-library)	59	129	229	0.31	31%	0.20	20%	1.70
SCOPUS Author-subject (Author(s))	78	113	151	0.42	42%	0.52	52%	3.8
SCOPUS citation	4	182	55	0.02	2%	0.07	7%	0.36
CINAHL Author-subject	19	167	22	0.10	10%	0.46	46%	4.61
CINAHL citation	1	185	3	0.01	1%	0.25	25%	1.64
WoK-Author-subject	72	115	75	0.38	38%	0.48	48%	6.82
E-lib – Author-subject	4	182	39	0.02	2%	0.09	9%	0.51
Total relevant R without duplicates					**186**			
Total retrieved without duplicate					**1096**			
Total non-relevant R without duplicates (D)					**910**			

Table 4: Tamoxifen Study Search Strategies Retrieval per Study Type and OR Calculations

Search Strategy	***RCTs***			***Qualitative***			***Economical evaluation***			***Process evaluation***		
	RR	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**
PubMed-Related articles	12	109	0.48	4	33	0.54	1	0	4.46	1	26	0.17
PubMed-Author-subject	12	109	0.22	1	36	0.06	0	1	0	4	23	0.34
Citation(Web of Science)	2	119	0.29	2	35	0.98	0	1	0	1	26	0.66
Subject search (e-library)	26	95	0.79	18	19	2.74	0	1	0	13	14	2.71
SCOPUS Author-subject (Author(s))	52	69	2.58	11	26	1.53	0	1	0	10	17	2.12
SCOPUS citation	0	121	0	4	33	1.85	0	1	0	0	27	0
CINAHL Author-subject	15	106	3.02	1	36	0.60	0	1	0	3	24	2.71
CINAHL citation	1	120	1.89	0	37	0	0	1	0	0	27	0.00
WoK-Author-subject	57	64	4.74	7	30	1.32	0	1	0	7	20	1.98
E-lib – Author-subject	1	120	0.18	3	34	1.85	0	1	0	0	27	0

Table 5: BMS: Search and Retrieval Performance

Breast Cancer (BMS)								
search strategy	Relevant/R (A)	Relevant N/R©	Not relevant (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Ratio
PubMed-Related articles	12	57	96	0.17	17%	0.11	11%	1.10
PubMed-Author-subject	2	67	65	0.03	3%	0.03	3%	0.24
Citation(Web of Science)	2	67	1	0.03	3%	0.67	67%	15.28
Subject search (e-library)	12	52	11	0.17	17%	0.52	52%	10.53
SCOPUS Author-subject (Group)	5	64	21	0.07	7%	0.19	19%	1.85
SCOPUS Author-subject (Author(s))	1	68	0	0.01	1%	1.00	100%	7.54
SCOPUS citation	2	67	1	0.03	3%	0.67	67%	15.25
CINAHL Author-subject	2	67	0	0.03	3%	1.00	100%	15.28
CINAHL citation	2	67	0	0.03	3%	1.00	100%	15.28
WoK-Author-subject	15	54	6	0.22	22%	0.71	71%	23.10
E-lib – Author-subject	6	63	77	0.09	9%	0.07	7%	0.63
Total relevant R without duplicates						**69**		
Total retrieved without duplicate						**583**		
Total non-relevant R without duplicates (D)						**514**		

Table 6: BMS Study Search Strategies Retrieval per Study Type and OR Calculations

Search Strategy	***RCTs***			***Qualitative***			***Economical evaluation***			***Process evaluation***		
	RR	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**
PubMed-Related articles	5	12	1.92	6	38	0.73	0	0	0	1	7	0.66
PubMed-Author-subject	0	17	0	1	43	0.18	0	0	0	1	7	1.09
Citation(Web of Science)	0	17	0	0	44	0.00	0	0	0	2	6	56.78
Subject search (e-library)	7	10	14.29	3	41	1.51	0	0	0	7	1	142.92
SCOPUS Author-subject (Group)	2	15	2.60	1	43	0.45	0	0	0	2	6	6.50
SCOPUS Author-subject (Author(s))	1	16	32.00	0	44	0	0	0	0	0	8	0
SCOPUS citation	0	17	0	0	44	0	0	0	0	2	6	56.67
CINAHL Author-subject	2	15	34.00	0	44	0	0	0	0	0	8	0
CINAHL citation	0	17	0	0	44	0	0	6	6	2	7	72.86
WoK-Author-subject	4	13	7.25	10	33	7.06	0	0	0	1	8	2.96
E-lib – Author-subject	5	12	2.53	1	42	0.15	0	0	0	0	9	0

Table 7: In Vitro Fertilization (IVF): Search and Retrieval Performance

IVF								
search strategy	**Relevant/R (A)**	**Relevant N/R©**	**Not relevant (B)**	**Recall Value**	**Recall %**	**Precision Value**	**Precision %**	**Odds Ratio**
PubMed-Related articles	18	19	88	0.49	49%	0.17	17%	13.75
PubMed-Author-subject	5	32	909	0.14	14%	0.01	1%	0.22
Citation(Web of Science)	14	23	42	0.38	38%	0.25	25%	18.57
Subject search (e-library)	14	23	277	0.30	30%	0.05	5%	2.82
SCOPUS Author-subject (Author(s))	12	25	38	0.32	32%	0.24	24%	16.21
SCOPUS citation								
CINAHL Author-subject								
CINAHL citation								
WoK-Author-subject	5	32	9	0.14	14%	0.36	36%	22.40
E-lib-Author-subject	9	28	39	0.24	24%	0.19	19%	10.60
Total relevant R without duplicates					**37**			
Total retrieved without duplicate					**1332**			
Total non-relevant R without duplicates (D)					**1295**			

Table 8: IVF Study Search Strategies Retrieval per Study Type and OR Calculations

Search Strategy	***RCTs***			***Qualitative***			***Economical evaluation***			***Process evaluation***		
	RR	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**	**RR**	**R/N**	**OR**
Related Search(PubMed)	15	11	16.24	0	1	0	2	3	8.02	1	4	3.01
Author-Subject(PubMed)	4	22	0.26	0	1	0	1	4	0.35	0	5	0.00
Citation(Web of Science)	8	18	10.10	0	1	0	2	3	15.24	4	1	91.43
Subject search(e-library, hMP)	2	24	9.77	0	1	0	1	4	29.34	1	4	29.32
Subject search(e-library, rFSH)	8	18	1.96	0	1	0	2	3	2.95	1	4	1.11
SCOPUS Author-subject	8	18	11.33	1	0	25.64	2	3	17.08	1	4	6.41
SCOPUS citation												
CINAHL Author-subject												
CINAHL citation												
WoK-Author-subject	3	23	11.99	0	1	0	2	3	61.33	0	5	0
E-lib - Author-subject	6	20	8	0	1	0	3	2	40.09	0	5	0

Table 9: CLD Infant Infection (Dexamethasone): Search and Retrieval Performance

CLD Infant Infection (Dexamethasone)								
search strategy	Relevant/R (A)	Relevant N/R©	Not relevant (B)	Recall Value	Recall %	Precision Value	Precision %	Odds Ratio
PubMed-Related articles	46	46	681	0.5	50%	0.06	6%	1.02
PubMed-Author-subject	3	89	36	0.03	3%	0.08	8%	0.71
Citation(Web of Science)	5	87	30	0.05	5%	0.14	14%	1.41
Subject search (e-library)	33	59	20	0.36	36%	0.62	62%	20.28
SCOPUS Author-subject (Group)	8	84	30	0.09	9%	0.21	21%	2.38
SCOPUS Author-subject (Author(s))	3	89	1	0.03	3%	0.75	75%	25.45
SCOPUS citation	15	77	29	0.16	16%	0.34	34%	4.99
CINAHL Author-subject	5	87	0	0.05	5%	1.00	100%	43.28
CINAHL citation	1	91	1	0.01	1%	0.50	50%	8.32
WoK-Author-subject	5	87	0	0.05	5%	1.00	100%	43.28
E-lib-Author-subject	2	90	36	0.02	2%	0.05	5%	0.47
Total relevant R without duplicates						**92**		
Total retrieved without duplicate						**850**		
Total non-relevant R without duplicates (D)						**758**		

Table 10: CLD Infant Infection (Dexamethasone) Study Search Strategies Retrieval per Study Type and OR Calculations

Search Strategy	*RCTs*			*Qualitative*			*Economical evaluation*			*Process evaluation*		
	RR	R/N	OR	RR	R/N	OR	RR	R/N	OR	RR	R/N	OR
PubMed-Related articles	34	35	0.91	4	1	3.90	0	1	0	7	10	0.68
PubMed-Author-subject	0	69	0	0	5	0	1	0	19.33	2	15	2.57
Citation(Web of Science)	3	66	0.97	0	5	0	0	1	0	2	17	2.53
Subject search (e-library)	30	39	10.09	0	5	0	0	1	0	3	14	2.92
SCOPUS Author-subject (Group)	1	68	0.29	2	3	13.12	0	1	0	5	12	8.17
SCOPUS Author-subject (Author(s))	2	67	5.62	1	4	47.13	0	1	0	0	17	0
SCOPUS citation	14	55	4.22	1	4	4.22	0	1	0	0	17	0
CINAHL Author-subject	1	68	2.21	0	5	0	0	1	0	4	13	46.09
CINAHL citation	0	69	0	0	5	0	0	1	0	1	16	23.63
WoK-Author-subject	5	64	11.69	0	5	0	0	1	0	0	17	0
E-lib-Author-subject	2	67	0.59	0	5	0	0	1	0	0	17	0

Appendix Six

Sibling lists for all Seed Studies

IDEATeL Siblings List

Sibling Title	*Direct*	Indirect
2005b. JAMA -- Patient Participation in Electronic Medical Records, April 4, 2001, Tsai and Starren 285 (13): 1765.	√	
2005c. Second Interim Report on the Informatics for Diabetes Education and Telemedicine (IDEATel) Demonstration: Final Report on Phase.	√	
2006b. American TeleCare to provide telehealth technology for diabetes research by IDEATel Consortium and participate in Department of.	√	
2010p. JAMA -- Patient Participation in Electronic Medical Records, April 4, 2001, Tsai and Starren 285 (13): 1765.	√	
2010q. MCW Faculty Collaboration Database - Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel). 2010.	√	
BAKKEN, S., GRULLON-FIGUEBOA, L., IZQUIERDO, R., LEE, N. J., MORIN, P., PALMAS, W., TERESI, J., WEINSTOCK, R. S., SHEA, S., STARREN, J. & CONSORTIUM, I. 2006. Development, validation, and use of English and Spanish versions of the Telemedicine Satisfaction and Usefulness Questionnaire. *Journal of the American Medical Informatics Association,* 13**,** 660-667.	√	
BALAMURUGAN, A., HALL-BARROW, J., BLEVINS, M. A., BRECH, D., PHILLIPS, M., HOLLEY, E. & BITTLE, K. 2009. A pilot study of diabetes education via telemedicine in a rural underserved community--opportunities and challenges: a continuous quality improvement process. *Diabetes Educ,* 35**,** 147-54.		√
BASHSHUR, R. L., SHANNON, G. W., KRUPINSKI, E. A., GRIGSBY, J., KVEDAR, J. C., WEINSTEIN, R. S., SANDERS, J. H., RHEUBAN, K. S., NESBITT, T. S., ALVERSON, D. C., MERRELL, R. C., LINKOUS, J. D., FERGUSON, A. S., WATERS, R. J., STACHURA, M. E., ELLIS, D. G., ANTONIOTTI, N. M., JOHNSTON, B., DOARN, C. R., YELLOWLEES, P., NORMANDIN, S. & TRACY, J. 2009. National Telemedicine Initiatives: Essential to Healthcare Reform. *Telemedicine Journal and E-Health,* 15**,** 600-610.		√
BELLAZZI, R., LARIZZA, C., MONTANI, S., RIVA, A., STEFANELLI, M., D'ANNUNZIO, G., LORINI, R. & TUOMINEN, J. 2002. A telemedicine support for diabetes management: the T-IDDM project. *Computer Methods and Programs in Biomedicine,* 69**,** 147.		√
BELLAZZI, R. & MONTANI, S. 2000. Building telemedicine systems for supporting decisions in diabetes care: a report from a running experience. *Diabetes technology & therapeutics,* 2**,** 577-582.		√
BELLAZZI, R., MONTANI, S., RIVA, A. & STEFANELLI, M. 2001. Web-based telemedicine systems for home-care: technical issues and experiences. *Computer Methods and Programs in Biomedicine,* 64**,** 175-187.		√
BIERMANN, E. 2009a. Telemedicine and diabetes. *Telemedizin beim Diabetes,* 4**,** 384-389.		√
BIERMANN, E., DIETRICH, W., RIHL, J. & STANDL, E. 2002a. Are there time and cost savings by using telemanagement for patients on intensified insulin therapy? A randomised,		√

controlled trial. *Computer Methods and Programs in Biomedicine,* 69, 137-146.		
CHANG, K., DAVIS, R., BIRT, J., CASTELLUCCIO, P., WOODBRIDGE, P. & MARRERO, D. 2007. Nurse practitioner-based diabetes care management: Impact of telehealth or telephone intervention on glycemic control. *Disease Management and Health Outcomes,* 15, 377-385.		√
CHIEN, M. N., LIU, L., CHIANG, H. C., WU, C. C., LEE, Y. J., HUANG, C. L., LEE, C. C., LEUNG, C. H. & WANG, C. H. 2006. Building a certificate-based secure and quick response telemedical health-care system for diabetic patients. *Journal of Internal Medicine of Taiwan,* 17, 61-72.		√
DANG, S., MA, F., NEDD, N., FLOREZ, H., AGUILAR, E. & ROOS, B. A. 2007. Care coordination and telemedicine improves glycaemic control in ethnically diverse veterans with diabetes. *J Telemed Telecare,* 13, 263-7.		√
FISCHER, H., MACKENZIE, T., MCCULLEN, K., EVERHART, R. & ESTACIO, R. O. 2008. Design of a nurse-run, telephone-based intervention to improve lipids in diabetics. *Contemporary Clinical Trials,* 29, 809-816.		√
FONDA, S. J., PAULSEN, C. A., PERKINS, J., KEDZIORA, R. J., RODBARD, D. & BURSELL, S. E. 2008. Usability test of an internet-based informatics tool for diabetes care providers: The comprehensive diabetes management program. *Diabetes Technology and Therapeutics,* 10, 16-24.		√
GOMEZ, E. J., HERNANDO, M. E., GARCIA, A., DEL POZO, F., CERMENO, J., CORCOY, R., BRUGUES, E. & DE LEIVA, A. 2002. Telemedicine as a tool for intensive management of diabetes: the DIABTel experience. *Computer Methods and Programs in Biomedicine,* 69, 163-177.		√
GOMEZ, E. J. 1996. Telemedicine for diabetes care: The DIABTel approach towards diabetes telecare. *Informatics for Health and Social Care,* 21, 283-295.		√
GRANT, R. W., WALD, J. S., SCHNIPPER, J. L., GANDHI, T. K., POON, E. G., ORAV, E. J., WILLIAMS, D. H., VOLK, L. A. & MIDDLETON, B. 2008. Practice-linked online personal health records for type 2 diabetes mellitus: A randomized controlled trial. *Archives of Internal Medicine,* 168, 1776-1782.		√
HERSH, W. R., HELFAND, M., WALLACE, J., KRAEMER, D., PATTERSON, P., SHAPIRO, S. & GREENLICK, M. 2001. Clinical outcomes resulting from telemedicine interventions: a systematic review. *BMC Medical Informatics and Decision Making,* 1, 5.		√
HILLIMAN, C. A., CIMINO, J. J., LAI, A. M., KAUFMAN, D. R., STARREN, J. B. & SHEA, S. 2009. The Effects of Redesigning the IDEATel Architecture on Glucose Uploads. *Telemedicine Journal and E-Health,* 15, 248-254.	√	
HOMENKO, D. R., MORIN, P. C., EIMICKE, J. P., TERESI, J. A. & WEINSTOCK, R. S. 2010. Food Insecurity and Food Choices in Rural Older Adults with Diabetes Receiving Nutrition Education via Telemedicine. *Journal of Nutrition Education and Behavior,* 42, 404.	√	
IZQUIERDO, R. 2008. School centered telemedicine study for children with diabetes. *DIABETES,* 57, A501-A501.	√	
IZQUIERDO, R., KNUDSON, P., MEYER, S., KEARNS, J. & WEINSTOCK, R. S. 2000. A comparison of diabetes education administered through telemedicine versus in person. *Diabetes,* 49, 1441.	√	
IZQUIERDO, R., LAGUA, C. T., MEYER, S., PLOUTZ-SNYDER, R. J., PALMAS, W., EIMICKE, J. P., KONG, J., TERESI, J. A., SHEA, S. & WEINSTOCK, R. S. 2010. Telemedicine intervention effects on waist circumference and body mass index in the IDEATel project. *Diabetes Technol Ther,* 12, 213-20.	√	

IZQUIERDO, R., MEYER, S., STARREN, J., GOLAND, R., TERESI, J., SHEA, S. & WEINSTOCK, R. S. 2007. Detection and remediation of medically urgent situations using telemedicine case management for older patients with diabetes mellitus. *Therapeutics and Clinical Risk Management,* 3**,** 485-489.	√
IZQUIERDO, R., MORIN, P. C., BRATT, K., MOREAU, Z., MEYER, S., PLOUTZ-SNYDER, R., WADE, M. & WEINSTOCK, R. S. 2009. School-Centered Telemedicine for Children with Type 1 Diabetes Mellitus. *The Journal of Pediatrics,* 155**,** 374-379.	√
IZQUIERDO, R. E. 2003. A comparison of diabetes education administered through telemedicine versus in person. *DIABETES CARE,* 26**,** 1002-1007.	√
JAMESON, B. C., ZYGMONT, S. V., NEWMAN, N. & WEINSTOCK, R. S. 2007. Effect of telemedicine on glycemic control in incarcerated individuals. *Diabetes,* 56**,** A649-A649.	√
JAMESON, B. C., ZYGMONT, S. V., NEWMAN, N. & WEINSTOCK, R. S. 2008. Use of Telemedicine to Improve Glycemic Management in Correctional Institutions. *Journal of Correctional Health Care,* 14**,** 197-201.	√
KAUFMAN, D. R., PATEL, V. L., HILLIMAN, C., MORIN, P. C., PEVZNER, J., WEINSTOCK, R. S., GOLAND, R., SHEA, S. & STARREN, J. 2003a. Usability in the real world: assessing medical information technologies in patients' homes. *Journal of Biomedical Informatics,* 36**,** 45-60.	√
KAUFMAN, D. R., PEVZNER, J., HILLIMAN, C., WEINSTOCK, R. S., TERESI, J., SHEA, S. & STARREN, J. 2006b. Technology topics. Redesigning a telehealth diabetes management program for a digital divide seniors population. *Home Health Care Management & Practice,* 18**,** 223-234.	√
KAUFMAN, D. R., PEVZNER, J., RODRIGUEZ, M., CIMINO, J. J., EBNER, S., FIELDS, L., MORENO, V., MCGUINESS, C., WEINSTOCK, R. S., SHEA, S. & STARREN, J. 2009. Understanding workflow in telehealth video visits: Observations from the IDEATel project. *Journal of Biomedical Informatics,* 42**,** 581-592.	√
KAUFMAN, D. R., PEVZNER, J., RODRIGUEZ, M., MCGUINESS, C., MORENO, V. I., CIMINO, J. I., SHEA, S., WEINSTOCK, R. S. & STARREN, J. 2007. Understanding dimensions of complexity in nurse case managers' workflow in a telemedicine program. *AMIA Annu Symp Proc***,** 1006.	√
KAUFMAN, D. R., STARREN, J., PATEL, V. L., MORIN, P. C., HILLIMAN, C., PEVZNER, J., WEINSTOCK, R. S., GOLAND, R. & SHEA, S. 2003b. A cognitive framework for understanding barriers to the productive use of a diabetes home telemedicine system. *AMIA Annu Symp Proc***,** 356-60.	√
LAI, A. M., KAUFMAN, D. R. & STARREN, J. 2006. Training digital divide seniors to use a telehealth system: a remote training approach. *AMIA Annu Symp Proc***,** 459-63.	√
LAI, A. M., KAUFMAN, D. R., STARREN, J. & SHEA, S. 2009. Evaluation of a remote training approach for teaching seniors to use a telehealth system. *International Journal of Medical Informatics,* 78**,** 732-744.	√
LAI, A. M., STARREN, J. B., KAUFMAN, D. R., MENDONCA, E. A., PALMAS, W., NIEH, J. & SHEA, S. 2008. The REmote Patient Education in a Telemedicine Environment Architecture (REPETE). *Telemed J E Health,* 14**,** 355-61.	√
LAI, A. M., STARREN, J. B., SHEA, S. & CONSORTIUM, I. D. 2005. Architecture for remote training of home telemedicine patients. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium***,** 1015.	√
LARIZZA, C., BELLAZZI, R., STEFANELLI, M., FERRARI, P., DE CATA, P.,	√

GAZZARUSO, C., FRATINO, P., D'ANNUNZIO, G., HERNANDO, E. & GOMEZ, E. J. 2006. The (MDM)-D-2 project - The experience of two Italian clinical sites with clinical evaluation of a multi-access service for the management of diabetes mellitus patients. *Methods of Information in Medicine,* 45**,** 79-84.

LOPEZ, R., CHAGPAR, A., WHITE, R., HAMILL, M. H., TRUDEL, M., CAFAZZO, J. & LOGAN, A. G. 2009. Usability of a diabetes telemanagement system. *Journal of Clinical Engineering,* 34**,** 147-151. √

MONTANI, S., BELLAZZI, R., RIVA, A., LARIZZA, C., PORTINALE, L. & STEFANELLI, M. 2000. Artificial intelligence techniques for diabetes management: The T-IDDM project. *Ecai 2000: 14th European Conference on Artificial Intelligence, Proceedings,* 54**,** 716-720. √

MONTORI, V. M., HELGEMOE, P. K., GUYATT, G. H., DEAN, D. S., LEUNG, T. W., SMITH, S. A. & KUDVA, Y. C. 2004. Telecare for patients with type 1 diabetes and inadequate glycemic control - A randomized controlled trial and meta-analysis. *Diabetes Care,* 27**,** 1088-1094. √

MORAN, A., PALMAS, W., FIELD, L., BHATTARAI, J., SCHWARTZ, J. E., WEINSTOCK, R. S. & SHEA, S. 2004. Cardiovascular autonomic neuropathy is associated with microalbuminuria in older patients with type 2 diabetes. *Diabetes Care,* 27**,** 972-7. √

MORAN, A., PALMAS, W., PICKERING, T. G., SCHWARTZ, J. E., FIELD, L., WEINSTOCK, R. S. & SHEA, S. 2006. Office and ambulatory blood pressure are independently associated with albuminuria in older subjects with type 2 diabetes. *Hypertension,* 47**,** 955-961. √

MORENO, L., DALE, S. B., CHEN, A. Y. & MAGEE, C. A. 2009. Costs to Medicare of the Informatics for Diabetes Education and Telemedicine (IDEATel) Home Telemedicine Demonstration Findings from an independent evaluation. *Diabetes Care,* 32**,** 1202-1204. √

MORIN, P. C., WOLFF, L. T., EIMICKE, J. P., TERESI, J. A., SHEA, S. & WEINSTOCK, R. S. 2009. Record media used by primary care providers in medically underserved regions of upstate New York was not pivotal to clinical result in the Informatics for Diabetes Education and Telemedicine (IDEATel) project. *Inform Prim Care,* 17**,** 103-12. √

PALMAS, W., MORAN, A., PICKERING, T., EIMICKE, J. P., TERESI, J., SCHWARTZ, J. E., FIELD, L., WEINSTOCK, R. S. & SHEA, S. 2006a. Ambulatory pulse pressure and progression of urinary albumin excretion in older patients with type 2 diabetes mellitus. *Hypertension,* 48**,** 301-308. √

PALMAS, W., PICKERING, T., EIMICKE, J. P., MORAN, A., TERESI, J., SCHWARTZ, J. E., FIELD, L. & SHEA, S. 2007. Value of Ambulatory Arterial Stiffness Index and 24-h Pulse Pressure to Predict Progression of Albuminuria in Elderly. *American Journal of Hypertension,* 20**,** 493-500. √

Palmas W, Pickering TG, Moran A, Teresi J, Schwartz JE, Eguchi K, Field LS, Weinstock RS, Shea S. Nocturnal blood pressure elevation predicts progression of albuminuria in type-2 diabetes. J Clin Hypertension 2008;10:12-30. √

PALMAS, W., PICKERING, T. G., TERESI, J., SCHWARTZ, J. E., FIELD, L., WEINSTOCK, R. S. & SHEA, S. 2008a. Telemedicine home blood pressure measurements and progression of albuminuria in elderly people with diabetes. *Hypertension,* 51**,** 1282-1288. √

PALMAS, W., PICKERING, T. G., TERESI, J., SCHWARTZ, J. E., MORAN, A., WEINSTOCK, R. S. & SHEA, S. 2009. Ambulatory blood pressure monitoring and all-cause mortality in elderly people with diabetes mellitus. *Hypertension,* 53**,** 120-127. √

PALMAS, W., SHEA, S., STARREN, J., TERESI, J. A., GANZ, M. L., BURTON, T. M., PASHOS, C. L., BLUSTEIN, J., FIELD, L., MORTIN, P. C., IZQUIERDO, R. E., SILVER, S., ELMICKE, J. P., LANTIGUA, R. A., WEINSTOCK, R. S. & CONSORTIUM, I. D. 2010. √

Medicare payments, healthcare service use, and telemedicine implementation costs in a randomized trial comparing telemedicine case management with usual care in medically underserved participants with diabetes mellitus (IDEATel). *Journal of the American Medical Informatics Association,* 17**,** 196-202.		
PALMAS, W., TERESI, J., MORIN, P., WOLFF, L. T., FIELD, L., EIMICKE, J. P., CAPPS, L., PRIGOLLINI, A., ORBE, I., WEINSTOCK, R. S. & SHEA, S. 2006b. Recruitment and enrollment of rural and urban medically underserved elderly into a randomized trial of telemedicine case management for diabetes care. *Telemedicine Journal and e-Health,* 12**,** 601-607.		
PALMAS, W., TERESI, J., WEINSTOCT, R. S. & SHEA, S. 2008b. Acceptability to primary care providers of telemedicine in diabetes case management. *Journal of Telemedicine and Telecare,* 14**,** 306-308.	√	
PEVZNER, J., KAUFMANN, D. R., HILLIMAN, C., SHEA, S., WEINSTOCK, R. S. & STARREN, J. 2005. Developing computer skills and competencies in seniors. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium***,** 1078.	√	
RALSTON, J. D., HIRSCH, I. B., HOATH, J., MULLEN, M., CHEADLE, A. & GOLDBERG, H. I. 2009. Web-based collaborative care for type 2 diabetes a pilot randomized trial. *Diabetes Care,* 32**,** 234-239.		√
ROBINSON, K. S., MORIN, P. C., SHUPE, J. A. C., IZQUIERDO, R., PLOUTZ-SNYDER, R., MEYER, S., TERESI, J. A., STARREN, J., SHEA, S. & WEINSTOCK, R. S. 2010. Use of Three Computer Training Methods in Elderly Underserved Rural Patients Enrolled in a Diabetes Telemedicine Program. *Cin-Computers Informatics Nursing,* 28**,** 172-177.	√	
RODRÍGUEZ-IDÍGORAS, M. I., SEPÚLVEDA-MUÑOZ, J., SÁNCHEZ-GARRIDO-ESCUDERO, R., MARTÍNEZ-GONZÁLEZ, J. L., ESCOLAR-CASTELLÓ, J. L., PANIAGUA-GÓMEZ, I. M., BERNAL-LÓPEZ, R., FUENTES-SIMÓN, M. V. & GARÓFANO-SERRANO, D. 2009. Telemedicine influence on the follow-up of type 2 diabetes patients. *Diabetes Technology and Therapeutics,* 11**,** 431-437.		√
SANDBERG, J., TRIEF, P. M., IZQUIERDO, R., GOLAND, R., MORIN, P. C., PALMAS, W., LARSON, C. D., STRAIT, J. G., SHEA, S. & WEINSTOCK, R. S. 2009. A Qualitative Study of the Experiences and Satisfaction of Direct Telemedicine Providers in Diabetes Case Management. *Telemedicine Journal and E-Health,* 15**,** 742-750.	√	
SARKAR, I. N. & STARREN, J. 2002. Desiderata for personal electronic communication in clinical systems. *J Am Med Inform Assoc,* 9**,** 209-16.	√	
SHEA, S. 2006. Health delivery system changes required when integrating telemedicine into existing treatment flows of information and patients. *Journal of Telemedicine and Telecare,* 12.	√	
SHEA, S. 2007. The Informatics for Diabetes and Education Telemedicine (IDEATel) project. *Trans Am Clin Climatol Assoc,* 118**,** 289-304.	√	
SHEA, S., MAHMUD, K., KUN, L., NITZKIN, J., FITZMAURICE, M. & RIPPEN, H. Telemedicine - Homecare evaluation of diabetes elderly patients: A coordinated effort between government, industry and academia. 2000. 10.	√	
SHEA, S., STARREN, J., WEINSTOCK, R. S., KNUDSON, P. E., TERESI, J., HOLMES, D., PALMAS, W., FIELD, L., GOLAND, R., TUCK, C., HRIPCSAK, G., CAPPS, L. & LISS, D. 2002. Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel) project: Rationale and design. *Journal of the American Medical Informatics Association,* 9**,** 49-62.	√	
SHEA, S., WALTER, P., THOMAS, G. P., JEANNE TERESI JOSEPH, E. S. & ANDREW MORAN RUTH, S. W. 2009a. Ambulatory Blood Pressure Monitoring and All-Cause Mortality	√	

in Elderly People With Diabetes Mellitus -- Palmas et al. 53 (2).		
SHEA, S., WALTER, P., THOMAS, G. P., JEANNE TERESI JOSEPH, E. S. & LESLEY, F. 2009b. Telemedicine Home Blood Pressure Measurements and Progression of Albuminuria in Elderly People With Diabetes -- Palmas et al.	√	
SHEA, S., WALTER, P., THOMAS, G. P., JEANNE TERESI JOSEPH, E. S. & LESLEY, F. 2010. Telemedicine Home Blood Pressure Measurements and Progression of Albuminuria in Elderly People With Diabetes -- Palmas et al.	√	
SHEA, S., WEINSTOCK, R. S., STARREN, J., TERESI, J., PLAMAS, W., FIELD, L., MORIN, P., GOLAND, R., IZQUIERDO, R. E., WOLFF, I. T., ASHRAF, M., HILLIMAN, C., SILVER, S., MEYER, S., HOLMES, D., PETKOVA, E., CAPPS, L., LANTIGUA, R. A. & CONSORTIUM, I. D. 2006. A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus. *Journal of the American Medical Informatics Association,* 13**,** 40-51.	√	
SHEA, S., WEINSTOCK, R. S., TERESI, J. A., PALMAS, W., STARREN, J., CIMINO, J. J., LAI, A. M., FIELD, L., MORIN, P. C., GOLAND, R., IZQUIERDO, R. E., EBNER, S., SILVER, S., PETKOVA, E., KONG, J., EIMICKE, J. P. & CONSORTIUM, I. D. 2009c. A Randomized Trial Comparing Telemedicine Case Management with Usual Care in Older, Ethnically Diverse, Medically Underserved Patients with Diabetes Mellitus: 5 Year Results of the IDEATel Study. *Journal of the American Medical Informatics Association,* 16**,** 446-456.	√	
SHEA, S. J. 2004. A randomized trial comparing telemedicine case management with usual care in older patients with diabetes mellitus. *JOURNAL OF GENERAL INTERNAL MEDICINE,* 19**,** 112-112.	√	
SHOJANIA, K. G., RANJI, S. R., MCDONALD, K. M., GRIMSHAW, J. M., SUNDARAM, V., RUSHAKOFF, R. J. & OWENS, D. K. 2006. Effects of quality improvement strategies for type 2 diabetes on glycemic control: A meta-regression analysis. *Journal of the American Medical Association,* 296**,** 427-440.		√
STARREN, J. 2005. Federal demonstration project tackles diabetes with telemedicine. *Caring,* 24**,** 6-10.	√	
STARREN, J., HILLIMAN, C., WEINSTOCK, R. S. & SHEA, S. Theater Style Demonstration: The Informatics for Diabetes Education And Telemedicine (IDEATel) Project. *AMIA Annual Symposium Proceedings*.	√	
STARREN, J., HRIPCSAK, G., SENGUPTA, S., ABBRUSCATO, C. R., KNUDSON, P. E., WEINSTOCK, R. S. & SHEA, S. 2002. Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel) project: Technical implementation. *Journal of the American Medical Informatics Association,* 9**,** 25-36.	√	
STARREN, J., SENGUPTA, S., HRIPCSAK, G., RING, G., KLERER, R. & SHEA, S. 2001. Making grandma's data secure: A security architecture for home telemedicine. *Journal of the American Medical Informatics Association***,** 657-661.	√	
STARREN, J., TSAI, C., BAKKEN, S., AIDALA, A., MORIN, P. C., HILLIMAN, C., WEINSTOCK, R. S., GOLAND, R., TERESI, J., SHEA, S. & CONSORTIUM, I. D. 2005. The role of nurses in installing telehealth technology in the home. *Cin-Computers Informatics Nursing,* 23**,** 181-189.	√	
TRIEF, P. M. 2006a. Depression and glycemic control in elderly ethnically diverse patients with diabetes - The IDEATel project. *DIABETES CARE,* 29**,** 830-835.	√	
TRIEF, P. M. 2006b. The effects of telemedicine case management for elderly diabetes patients on psychosocial outcomes: The IDEATel project. *DIABETES,* 55**,** A432-A432.	√	

TRIEF, P. M. 2007. Depression in elderly diabetes patients. *Diabetes Spectrum,* 20, 71.	√
TRIEF, P. M., IZQUIERDO, R., MORIN, P. C., TERESI, J., SHEA, S., STARREN, T. & WEINSTOCK, R. S. 2005. Does depression predict intervention outcomes - The IDEATel Project. *Diabetes,* 54, A465-A465.	√
TRIEF, P. M., MORIN, P. C., IZQUIERDO, J., TERESI, J., STARREN, J., SHEA, S. & WEINSTOCK, R. S. 2006a. Marital quality and diabetes outcomes: the IDEATel Project. *Families, Systems & Health: The Journal of Collaborative Family HealthCare,* 24, 318-331.	√
TRIEF, P. M., MORIN, P. C., IZQUIERDO, R., TERESI, J., STARREN, J., SHEA, S. & WEINSTOCK, R. S. 2006b. Marital Quality and Diabetes Outcomes: The IDEATel Project. *Families, Systems & Health,* 24, 318.	√
TRIEF, P. M., SANDBERG, J., IZQUIERDO, R., MORIN, P. C., SHEA, S., BRITTAIN, R., FELDHOUSEN, E. B. & WEINSTOCK, R. S. 2008. Diabetes Management Assisted by Telemedicine: Patient Perspectives. *Telemedicine Journal and E-Health,* 14, 647-655.	√
TRIEF, P. M., SANDBERG, J. G., IZQUIERDO, R., MORIN, P. C., SHEA, S., BRITTAIN, R. A., FELDHOUSEN, E. B. & WEINSTOCK, R. S. 2007a. THE TELEMEDICINE EXPERIENCE OF ELDERLY DIABETES PATIENTS: A QUALITATIVE STUDY. *Annals of Behavioral Medicine,* 33, S96-S96.	√
TRIEF, P. M., TERESI, J. A., EIMICKE, J. P., SHEA, S. & WEINSTOCK, R. S. 2009. Improvement in diabetes self-efficacy and glycaemic control using telemedicine in a sample of older, ethnically diverse individuals who have diabetes: the IDEATel project. *Age Ageing,* 38, 219-25.	√
TRIEF, P. M., TERESI, J. A., IZQUIERDO, R., MORIN, P. C., GOLAND, R., FIELD, L., EIMICKE, J. P., BRITTAIN, R., STARREN, J., SHEA, S. & WEINSTOCK, R. S. 2007b. Psychosocial outcomes of telemedicine case management for elderly patients with diabetes: the randomized IDEATel trial. *Diabetes Care,* 30, 1266-1268.	√
TSAI, C. C. & STARREN, J. 2001. msJAMA. Patient participation in electronic medical records. *JAMA,* 285, 1765.	√
TUDIVER, F., WOLFF, L. T., MORIN, P. C., TERESI, J., PALMAS, W., STARREN, J., SHEA, S. & WEINSTOCK, R. S. 2007. Primary care providers' perceptions of home diabetes telemedicine care in the IDEATel project. *Journal of Rural Health,* 23, 55-61.	√
WEINSTOCK, R. S., IZQUIERDO, R., GOLAND, R., PALMAS, W., TERESI, J. A., EIMICKE, J. P. & SHEA, S. 2010. Lipid treatment in ethnically diverse underserved older adults with diabetes mellitus: statin use, goal attainment, and health disparities in the informatics for diabetes education and telemedicine project. *J Am Geriatr Soc,* 58, 401-2.	√
WEST, S. P., LAGUA, C., TRIEF, P. M., IZQUIERDO, R. & WEINSTOCK, R. S. 2010. Goal Setting Using Telemedicine in Rural Underserved Older Adults with Diabetes: Experiences from the Informatics for Diabetes Education and Telemedicine Project. *Telemedicine Journal and E-Health,* 16, 405-416.	√

Tamoxifen Siblings List

Sibling Title	*Direct*	Indirect
BOEHM, D. U., LEBRECHT, A., ECKHARDT, T., ALBRICH, S., SCHMIDT, M., SIGGELKOW, W., KANDELHARDT, E. & KOELBL, H. 2009. Quality of life and adjuvant tamoxifen treatment in breast cancer patients: Original article. *European Journal of Cancer Care,* 18**,** 500-506.		√
BONN, D. 1999. Doses of tamoxifen used to prevent breast cancer may be too high. *Lancet,* 354**,** 841-841.		√
CAMERON, L. D., LEVENTHAL, H. & LOVE, R. R. 1998. Trait anxiety, symptom perceptions, and illness-related responses among women with breast cancer in remission during a tamoxifen clinical trial. *Health Psychol,* 17**,** 459-69.		√
CELLA, D., LAND, S. R., CHANG, C. H., DAY, R., COSTANTINO, J. P., WOLMARK, N. & GANZ, P. A. 2008. Symptom measurement in the Breast Cancer Prevention Trial (BCPT) (P-1): psychometric properties of a new measure of symptoms for midlife women. *Breast Cancer Research and Treatment,* 109**,** 515-526.		√
CHANG, J., POWLES, T. J., ASHLEY, S. E., GREGORY, R. K., TIDY, V. A., TRELEAVEN, J. G. & SINGH, R. 1996. The effect of tamoxifen and hormone replacement therapy on serum cholesterol, bone mineral density and coagulation factors in healthy postmenopausal women participating in a randomised, controlled tamoxifen prevention study. *Annals of Oncology,* 7**,** 671-675.	√	
CLEMONS, M., DANSON, S. & HOWELL, A. 2002. Tamoxifen ('Nolvadex'): A review. *Cancer Treatment Reviews,* 28**,** 165-180.	√	
CUMMINGS, S. R., TICE, J. A., BAUER, S., BROWNER, W. S., CUZICK, J., ZIV, E., VOGEL, V., SHEPHERD, J., VACHON, C., SMITH-BINDMAN, R. & KERLIKOWSKE, K. 2009. Prevention of breast cancer in postmenopausal women: Approaches to estimating and reducing risk. *Journal of the National Cancer Institute,* 101**,** 384-398.	√	
CUZICK, J. 1992. METHODOLOGIC ISSUES IN THE CHEMOPREVENTION OF BREAST-CANCER. *Cancer Detection and Prevention,* 16**,** 81-85.	√	
CUZICK, J. 1995. Methodologic issues in the chemoprevention of breast cancer. *Hormone-Dependent Tumors,* 50**,** 111-120.	√	
CUZICK, J. 1996. Chemoprevention of breast cancer with tamoxifen. *Chemoprevention in Cancer Control***,** 95-109.	√	
CUZICK, J. 1998. Continuation of the International Breast Cancer Intervention Study (IBIS). *European Journal of Cancer,* 34**,** 1647-1648.	√	
CUZICK, J. 2000a. A brief review of the current breast cancer prevention trials and proposals for future trials. *European Journal of Cancer,* 36**,** 1298-1302.	√	
CUZICK, J. 2000b. Future possibilities in the prevention of breast cancer - Breast cancer prevention trials. *Breast Cancer Research,* 2**,** 258-263.	√	
CUZICK, J. 2001. A brief review of the International Breast Cancer Intervention Study (IBIS), the other current breast cancer prevention trials, and proposals for future trials. *Selective Estrogen Receptor Modulators (Serms),* 949**,** 123-133.	√	
CUZICK, J. 2003a. Aromatase inhibitors in prevention--data from the ATAC (arimidex, tamoxifen alone or in combination) trial and the design of IBIS-II (the second International Breast Cancer Intervention Study). *Recent results in cancer research. Fortschritte der Krebsforschung.*	√	

Progres dans les recherches sur le cancer, 163**,** 96-103; discussion 264.

CUZICK, J. 2003b. Epidemiology of breast cancer - Selected highlights. *Breast,* 12**,** 405-411. √

CUZICK, J. 2004a. Aromatase inhibitors in breast cancer prevention: The IBIS-II trial. *EJC Supplements,* 2**,** 19. √

CUZICK, J. 2004b. E9. Side-effects of tamoxifen and solutions with aromatase inhibitors. *European Journal of Cancer, Supplement,* 2**,** 33-34. √

CUZICK, J. 2005. Aromatase inhibitors for breast cancer prevention. *Journal of Clinical Oncology,* 23**,** 1636-1643. √

CUZICK, J. 2005. Aromatase inhibitors in breast cancer prevention: The IBIS-II trial. EJC Supplements, 2, 19. √

CUZICK, J. 2006. Chemoprevention of breast cancer. *Women's Health,* 2**,** 853-861. √

CUZICK, J. 2007. Evaluation of tamoxifen prophylaxis for breast cancer: 8-year follow-up of the IBIS-I study. *American Journal of Hematology/Oncology,* 6**,** 395. √

CUZICK, J. 2008a. Chemoprevention of breast cancer. *Breast cancer (Tokyo, Japan),* 15**,** 10-16. √

CUZICK, J. 2008b. Use of aromatase inhibitors in breast cancer prevention. *Ejc Supplements,* 6**,** S21. √

Cuzick, J. 2005. Breast cancer prevention - Beyond tamoxifen. Cancer Epidemiology Biomarkers & Prevention, 14, 2804S-2804S. √

CUZICK, J., BAUM, M., PATERSON, A. H. G., NEVEN, P., LOWE, D. G., SHEPHERD, J. H. & NICHOLSON, R. H. 1992. Prevention of breast cancer [22]. *Lancet,* 340**,** 1550-1552. √

Cuzick, J. Allen, D. Baum, M. Barrett, J. Clark, G. Kakkar, V. Melissari, E. Moniz, C. Moore, J. Parsons, V. Pemberton, K. Pitt, P. Richmond, W. Houghton, J. Riley, D. 1993. LONG-TERM EFFECTS OF TAMOXIFEN - BIOLOGICAL EFFECTS OF TAMOXIFEN WORKING PARTY. *European Journal of Cancer*, 29, 1521. √

CUZICK, J. & CHOUINARD, E. 2003. Tamoxifen reduced breast cancer risk but increased risks of thromboembolic events and all cause mortality in women. *Evidence-Based Medicine,* 8**,** 110. √

CUZICK, J. & EDWARDS, R. 1999. Drop-outs in tamoxifen prevention trials [7]. *Lancet,* 353**,** 930. √

CUZICK, J., FORBES, J., EDWARDS, R., BAUM, M., CAWTHORN, S., COATES, A., HAMED, H., HOWELL, A., POWLES, T., CLUNIE, G., COLLINS, R., DAY, N., NORTHOVER, J. & INVESTIGATORS, I. 2002. First results from the International Breast Cancer Intervention Study (IBIS-1): a randomised prevention trial. *Lancet,* 360**,** 817-824. √

CUZICK, J., FORBES, J., HOWELL, A. & INVESTIGATORS, I. 2003a. Tamoxifen for breast-cancer prevention - Reply. Lancet, 361, 178-178. √

CUZICK, J., FORBES, J. F., HOWELL, A., FISHER, B. & COSTANTINO, J. P. 2006a. Re: Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study... J Natl Cancer Inst. 2005 Nov 16;97(22):1652-62. JNCI: Journal of the National Cancer Institute, 98, 643-644. √

CUZICK, J., FORBES, J. F., SESTAK, I., CAWTHORN, S., HAMED, H., HOLLI, K., HOWELL, A. & IBIS 2007. Long-term results of tamoxifen prophylaxis for breast cancer-96-month follow-up of the randomized IBIS-I trial. Journal of the National Cancer Institute, 99, 272-282. √

CUZICK, J. & HOWELL, A. 2000. A brief review of the breast cancer prevention trials. European Journal of Cancer, 36, S51-S53. √

CUZICK, J. & INVESTIGATORS, I. 2006. Long term efficacy of tamoxifen for chemoprevention - results of the IBIS-I study. Breast Cancer Research and Treatment, 100, S16-S17. √

CUZICK, J., POWLES, T., VERONESI, U., FORBES, J., EDWARDS, R., ASHLEY, S. & BOYLE, P. 2003b. Overview of the main outcomes in breast-cancer prevention trials. Lancet, 361, 296-300. √

CUZICK, J., SASIENI, P. & HOWELL, A. 2006b. Should aromatase inhibitors be used as initial adjuvant treatment or sequenced after tamoxifen? British Journal of Cancer, 94, 460-464. √

CUZICK, J., WARWICK, J., PINNEY, E., WARREN, R. M. L. & DUFFY, S. W. 2004. Tamoxifen and breast density in women at increased risk of breast cancer. Journal of the National Cancer Institute, 96, 621-628. √

CYKERT, S., CAINE, G. J., BLANN, A. D., STONELAKE, P. S., KEHOE, S. T., LIP, G. Y. H., CUZICK, J., FORBES, J. & HOWELL, A. 2003. Tamoxifen for breast-cancer prevention [3] (multiple letters). *Lancet,* 361**,** 177-178. √

DAY, R. 2001. Quality of life and tamoxifen in a breast cancer prevention trial: a summary of findings from the NSABP P-1 study. National Surgical Adjuvant Breast and Bowel Project. *Ann N Y Acad Sci,* 949**,** 143-50. √

DAY, R., GANZ, P. A. & COSTANTINO, J. P. 2001. Tamoxifen and depression: more evidence from the National Surgical Adjuvant Breast and Bowel Project's Breast Cancer Prevention (P-1) Randomized Study. *J Natl Cancer Inst,* 93**,** 1615-23. √

DETRE, S., ASHLEY, S., TIDY, A., SMITH, I. E., POWLES, T. & DOWSETT, M. 2009. Immunohistochemical Phenotype after 20-Year Follow-Up of the Royal Marsden Tamoxifen Breast Cancer Prevention Trial (RMTBCPT). *Cancer Research,* 69**,** 553S-553S. √

DOWSETT, M., CUZICK, J., INGLE, J., COATES, A., FORBES, J., BLISS, J., BUYSE, M., BAUM, M., BUZDAR, A., COLLEONI, M., COOMBES, C., SNOWDON, C., GNANT, M., JAKESZ, R., KAUFMANN, M., BOCCARDO, F., GODWIN, J., DAVIES, C. & PETO, R. 2010. Meta-analysis of breast cancer outcomes in adjuvant trials of aromatase inhibitors versus tamoxifen. *Journal of Clinical Oncology,* 28**,** 509-518. √

DUGGAN, C., MARRIOTT, K., EDWARDS, R. & CUZICK, J. 2003. Inherited and acquired risk factors for venous thromboembolic disease among women taking tamoxifen to prevent breast cancer. *Journal of Clinical Oncology,* 21**,** 3588-3593. √

EELES, R. A., POWLES, T. P., ASHLEY, S., EASTON, D. F., ASSERSOHN, L., SODHA, N., DOWSETT, M., GUSTERSON, B., TIDY, A., MITCHELL, G. & KOTE-JARAI, Z. 1999. BRCA1, BRCA2 and pedigree genetic analysis to determine genetic risk in the UK Royal Marsden Hospital tamoxifen prevention trial. *American Journal of Human Genetics,* 65**,** 662. √

EELES, R. A., POWLES, T. P. 2000. Chemoprevention options for BRCA1 and BRCA2 mutation carriers. *Journal of Clinical Oncology,* 18, 93s-99. √

FALLOWFIELD, L. 2005. Acceptance of adjuvant therapy and quality of life issues. *Breast (Edinburgh, Scotland),* 14**,** 612-616. √

FALLOWFIELD, L., FLEISSIG, A., EDWARDS, R., WEST, A., POWLES, T. J., HOWELL, A. & CUZICK, J. 2001. Tamoxifen for the prevention of breast cancer: psychosocial impact on women participating in two randomized controlled trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 19**,** official j-1892. √

FALLOWFIELD, L., MCGURK, R. & DIXON, M. 2004. Same gain, less pain: Potential patient preferences for adjuvant treatment in premenopausal women with early breast cancer. *European Journal of Cancer,* 40**,** 2403-2410. √

FALLOWFIELD, L. J., BLISS, J. M., PORTER, L. S., PRICE, M. H., SNOWDON, C. F., JONES, S. E., COOMBES, R. C. & HALL, E. 2006. Quality of life in the intergroup exemestane study: a randomized trial of exemestane versus continued tamoxifen after 2 to 3 years of tamoxifen in postmenopausal women with primary breast cancer. *Journal of Clinical Oncology,* 24**,** 910-917. √

FISHER, B., COSTANTINO, J. P., WICKERHAM, D. L., CECCHINI, R. S., CRONIN, W. M., ROBIDOUX, A., BEVERS, T. B., KAVANAH, M. T., ATKINS, J. N., MARGOLESE, R. G., RUNOWICZ, C. D., JAMES, J. M., FORD, L. G. & WOLMARK, N. 2005. Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *J Natl Cancer Inst,* 97**,** 1652-62. √

FISHER, B., LAND, S., MAMOUNAS, E., DIGNAM, J., FISHER, E. R. & WOLMARK, N. 2001. Prevention of invasive breast cancer in women with ductal carcinoma in situ: An update of the national surgical adjuvant breast and bowel project experience. *Seminars in Oncology,* 28**,** 400-418. √

GANZ, P. A. & LAND, S. R. 2008. Risks, benefits, and effects on quality of life of selective estrogen-receptor modulator therapy in postmenopausal women at increased risk of breast cancer. *Menopause-the Journal of the North American Menopause Society,* 15**,** 797-803. √

GAO, J. N., FORBES, J. F., WARREN, R. M. L., CUZICK, J., HOWELL, A., D'ESTE, C. & WARREN-FORWARD, H. 2006. Change in marnmographic density after cessation of tamoxifen: results from international breast cancer intervention study I (IBIS I). *Breast Cancer Research and Treatment,* 100**,** S56-S56. √

HEISEY, R., PIMLOTT, N., CLEMONS, M., CUMMINGS, S. & DRUMMOND, N. 2006. Women's views on chemoprevention of breast cancer: Qualitative study. *Canadian Family Physician,* 52**,** 624-625. √

HOUSSAMI, N., CUZICK, J. & DIXON, J. M. 2006. The prevention, detection, and management of breast cancer. *Medical Journal of Australia,* 184**,** 230-234. √

HOWELL, A. 1998. Endocrine prevention of breast cancer: The jury is half in. *Endocrine-Related Cancer,* 5**,** 249-251. √

HOWELL, A. 2000. 2nd European Breast Cancer Conference, Bruxelles, September 26-30, 2000. Interview with Anthony Howell. *Tumori,* 86**,** A1-4. √

HOWELL, A. 2008. The endocrine prevention of breast cancer. *Best Practice and Research: Clinical Endocrinology and Metabolism,* 22**,** 615-623. √

HOWELL, A., BRAMLEY, M. D., BAILDAM, A. D., CLARKE, R. B. & ANDERSON, E. 1999. Endocrine approaches to breast cancer prevention. *Familial Cancer and Prevention***,** 431-439. √

HOWELL, A., HOWELL, S. J. & EVANS, D. G. 2003. New approaches to the endocrine prevention and treatment of breast cancer. *Cancer Chemotherapy and Pharmacology, Supplement,* 52. √

HOWELL, A., LONNING, P., OSBORNE, K., PRITCHARD, K. & WAKELING, A. 2001. Preliminary experience with pure antiestrogens. *Clinical Cancer Research,* 7. √

Howell, A., & Buzdar, A. 2005. Are aromatase inhibitors superior to antiestrogens? *Journal of Steroid Biochemistry and Molecular Biology*, 93, 237-247. √

HOWELL, S. J., JOHNSTON, S. R. D. & HOWELL, A. 2004. The use of selective estrogen receptor modulators and selective estrogen receptor down-regulators in breast cancer. *Best* √

Practice & Research Clinical Endocrinology & Metabolism, 18**,** 47-66.

HUTCHINGS, O., EVANS, G., FALLOWFIELD, L., CUZICK, J. & HOWELL, A. 1998. Effect of early American results on patients in a tamoxifen prevention trial (IBIS) [3]. *Lancet,* 352**,** 1222. √

INVESTIGATORS, I. & CUZICK, J. 2002. First results from the International Breast Cancer Intervention Study (IBIS-I): A randomised prevention trial. *Lancet (North American Edition),* 360**,** 817-824. √

JENKINS, V. A., AMBROISINE, L. M., ATKINS, L., CUZICK, J., HOWELL, A. & FALLOWPELD, L. 2008. Effects of anastrozole on cognitive performance in postmenopausal women: a randomised, double-blind chemoprevention trial (IBIS II). *Lancet Oncology,* 9**,** 953-961. √

KEDAR, R. P., BOURNE, T. H., POWLES, T. J., COLLINS, W. P., ASHLEY, S. E., COSGROVE, D. O. & CAMPBELL, S. 1994. Effects of tamoxifen on uterus and ovaries of postmenopausal women in a randomised breast cancer prevention trial. *Lancet,* 343**,** 1318-1321. √

KOTE-JARAI, Z., POWLES, T. J., MITCHELL, G., TIDY, A., ASHLEY, S., EASTON, D., ASSERSOHN, L., SODHA, N., SALTER, J., GUSTERSON, B., DOWSETT, M. & EELES, R. 2007. BRCA1/BRCA2 mutation status and analysis of cancer family history in participants of the Royal Marsden Hospital tamoxifen chemoprevention trial. *Cancer Letters,* 247**,** 259-265. √

LAND, S. R., WICKERHAM, D. L., COSTANTINO, J. P., RITTER, M. W., VOGEL, V. G., LEE, M., PAJON, E. R., WADE, J. L., DAKHIL, S., LOCKHART, J. B., WOLMARK, N. & GANZ, P. A. 2006. Patient-reported symptoms and quality of life during treatment with tamoxifen or raloxifene for breast cancer prevention - The NSABP study of tamoxifen and raloxifene (STAR) P-2 trial. *Jama-Journal of the American Medical Association,* 295**,** 2742-2751. √

MARTINO, S., COSTANTINO, J., MCNABB, M., MERSHON, J., BRYANT, K., POWLES, T. & SECREST, R. J. 2004. The Role of Selective Estrogen Receptor Modulators in the Prevention of Breast Cancer: Comparison of the Clinical Trials. *Oncologist,* 9**,** 116-125. √

MAURICE, A., HOWELL, A., EVANS, D. G., O'NEIL, A. C. & SCOBIE, S. 2006. Predicting compliance in a breast cancer prevention trial. *Breast Journal,* 12**,** 446-450. √

MCGURK, R., FALLOWFIELD, L. & WINTERS, Z. 2006. Information provision for patients by breast cancer teams about the side-effects of hormone treatments. *European Journal of Cancer,* 42**,** 1760-1767. √

METCALFE, K. A., SNYDER, C., SEIDEL, J., HANNA, D., LYNCH, H. T. & NAROD, S. 2005. The use of preventive measures among healthy women who carry a BRCA1 or BRCA2 mutation. *Fam Cancer,* 4**,** 97-103. √

MORTIMER, J. E., BOUCHER, L., BATY, J., KNAPP, D. L., RYAN, E. & ROWLAND, J. H. 1999. Effect of tamoxifen on sexual functioning in patients with breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 17**,** official j-1492. √

O'BRIEN, M. & POWLES, T. J. 1994. Tamoxifen in the prevention of breast cancer: Are the risks likely to outweigh the benefits? *Drug Safety,* 10**,** 1-4. √

POWLES, T. 1998a. Tamoxifen and bone: Data from breast cancer prevention studies. *Obstetrical and Gynecological Survey,* 53. √

POWLES, T. 1998b. Use of tamoxifen plus estrogen. *Obstetrical and Gynecological Survey,* 53. √

POWLES, T. 2000. Breast Cancer Prevention. Breast Cancer Research, √

POWLES, T. 2001a. Breast cancer chemoprevention - Quo vadis? *Breast,* 10**,** 58-61. √

POWLES, T. 2004a. Breast cancer chemoprevention: Quo vadis? *EJC Supplements,* 2**,** 20. √

POWLES, T. 2006a. Endocrine prevention of breast cancer. *EJC Supplements,* 4**,** 34.	√
POWLES, T. 2006b. The Jordan article reviewed. *ONCOLOGY,* 20.	√
POWLES, T. 2010. Prevention of breast cancer by newer SERMs and the future. *Ejc Supplements,* 8**,** 6-6.	√
POWLES, T., EELES, R., ASHLEY, S., EASTON, D., CHANG, J., DOWSETT, M., TIDY, A., VIGGERS, J. & DAVEY, J. 1998a. Interim analysis of the incidence of breast cancer in the Royal Marsden Hospital tamoxifen randomised chemoprevention trial. *Lancet,* 352**,** 98-101.	√
POWLES, T., HICKISH, T., KANIS, J. A., TIDY, A. & ASHLEY, S. 1996. Effect of tamoxifen on bone mineral density measured by dual-energy x-ray absorptiometry in healthy premenopausal and postmenopausal women. *Journal of Clinical Oncology,* 14**,** 78-84.	√
POWLES, T., PATERSON, A., MCCLOSKEY, E., SCHEIN, P., SCHEFFLER, B., TIDY, A., ASHLEY, S., SMITH, I., OTTESTAD, L. & KANIS, J. 2006. Reduction in bone relapse and improved survival with oral clodronate for adjuvant treatment of operable breast cancer ISRCTN83688026. *Breast Cancer Research,* 8.	√
POWLES, T., PATERSON, S., KANIS, J. A., MCCLOSKEY, E., ASHLEY, S., TIDY, A., ROSENQVIST, K., SMITH, I., OTTESTAD, L., LEGAULT, S., PAJUNEN, M., NEVANTAUS, A., MÄNNISTÖ, E., SUOVUORI, A., ATULA, S., NEVALAINEN, J. & PYLKKÄNEN, L. 2002. Randomized, placebo-controlled trial of clodronate in patients with primary operable breast cancer. *Journal of Clinical Oncology,* 20**,** 3219-3224.	√
POWLES, T. J. 1992. The case for clinical trials of tamoxifen for prevention of breast cancer. *Lancet,* 340**,** 1145-7.	√
POWLES, T. J. 1994. CLINICAL-TRIALS OF TAMOXIFEN FOR PREVENTION OF BREAST-CANCER. *Journal of Cellular Biochemistry***,** 249-249.	√
POWLES, T. J. 1997a. Chemoprevention of breast cancer using tamoxifen. *Endocrine-Related Cancer,* 4**,** 255-260.	√
POWLES, T. J. 1997b. Efficacy of tamoxifen as treatment of breast cancer. *Semin Oncol,* 24**,** S1-48-S1-54.	√
POWLES, T. J. 1998c. Ii.1 Tamoxifen's oestrogen-like effects in a breast cancer chemoprevention trial. *European Journal of Cancer,* 34.	√
POWLES, T. J. 1998d. Status of antiestrogen breast cancer prevention trials. *Oncology-New York,* 12**,** 28-31.	√
POWLES, T. J. 1999a. Re: Tamoxifen for prevention of breast cancer: Report of the national surgical adjuvant breast and bowel project P-1 study [7]. *Journal of the National Cancer Institute,* 91**,** 730.	√
POWLES, T. J. 1999b. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *Journal of the National Cancer Institute,* 91**,** 730-730.	√
POWLES, T. J. 2000a. Re: Combined treatment with buserelin and tamoxifen in premenopausal metastatic breast cancer: a randomized study. *Journal of the National Cancer Institute,* 92**,** 2040-2041.	√
POWLES, T. J. 2000b. Tamoxifen for the prevention of breast cancer - Contra. *European Journal of Cancer,* 36**,** 146-147.	√
POWLES, T. J. 2001b. Role of antioestrogens and SERMs in breast cancer prevention. *Clinical Cancer Research,* 7**,** P815.	√

POWLES, T. J. 2001c. The Royal Marsden Hospital (RMH) trial: Key points and remaining questions. √

POWLES, T. J. 2002. Breast cancer prevention. *Oncologist,* 7**,** 60-64. √

POWLES, T. J. 2003. Anti-oestrogenic chemoprevention of breast cancer-the need to progress. *Eur J Cancer,* 39**,** 572-9. √

POWLES, T. J. 2004b. E1. Endocrine prevention of breast cancer. *European Journal of Cancer, Supplement,* 2**,** 15-16. √

POWLES, T. J. 2006c. Do we have a clinically useful breast cancer risk reduction agent for healthy women?: Commentary. *Nature Clinical Practice Oncology,* 3**,** 650-651. √

POWLES, T. J. 2006d. Prevention of breast cancer using selective oestrogen receptor modulators (SERMs). *Breast Cancer Research,* 8. √

POWLES, T. J. 2008. The Vogel article reviewed. Addressing concerns about breast cancer prevention. *Oncology (08909091),* 22**,** 679. √

POWLES, T. J. & ASHLEY, S. 1994. Endometrial cancer during tamoxifen treatment. *Lancet,* 343**,** 978. √

POWLES, T. J., ASHLEY, S., SMITH, I. E. & DOWSETT, M. 2007a. Responses. *Journal of the National Cancer Institute,* 99**,** 1645-1646. √

POWLES, T. J., ASHLEY, S., TIDY, A. & DOWSETT, M. 2007b. Twenty-year follow-up of the royal Marsden randomized, double-blinded tamoxifen breast cancer prevention trial. *Journal of the National Cancer Institute,* 99**,** 283-290. √

POWLES, T. J., BOURNE, T., ATHANASIOU, S., CHANG, J., GRUBOCK, K., ASHLEY, S., OAKES, L., TIDY, A., DAVEY, J., VIGGERS, J., HUMPHRIES, S. & COLLINS, W. 1998b. The effects of norethisterone on endometrial abnormalities identified by transvaginal ultrasound screening of healthy post menopausal women on tamoxifen or placebo. *British Journal of Cancer,* 78**,** 272-275. √

POWLES, T. J. & CHANG, J. 1997. Chemoprevention of breast cancer: Why is tamoxifen not the answer? *Endocrine-Related Cancer,* 4**,** 135-139. √

POWLES, T. J. & HICKISH, T. 1995. Breast cancer response to hormone replacement therapy withdrawal. *Lancet,* 345**,** 1442. √

POWLES, T. J., HICKISH, T., CASEY, S. & O'BRIEN, M. 1993. Hormone replacement after breast cancer. *Lancet,* 342**,** 60-1. √

POWLES, T. J., HOWELL, A., EVANS, D. G., MCCLOSKEY, E. V., ASHLEY, S., GREENHALGH, R., AFFEN, J., FLOOK, L. A. & TIDY, A. 2008. Red clover isoflavones are safe and well tolerated in women with a family history of breast cancer. *Menopause International,* 14**,** 6-12. √

POWLES, T. J., JONES, A. L., ASHLEY, S. E., O'BRIEN, M. E., TIDY, V. A., TRELEAVAN, J., COSGROVE, D., NASH, A. G., SACKS, N., BAUM, M. & ET AL. 1994. The Royal Marsden Hospital pilot tamoxifen chemoprevention trial. *Breast Cancer Res Treat,* 31**,** 73-82. √

ROSS, P. J. & POWLES, T. J. 2001. Results and implications of the Royal Marsden and other tamoxifen chemoprevention trials... including commentary by Sparano JA and Fabian C. *Clinical Breast Cancer,* 2**,** 33-40. √

SANTEN, R. J., BOYD, N. F., CHLEBOWSKI, R. T., CUMMINGS, S., CUZICK, J., DOWSETT, M., EASTON, D., FORBES, J. F., KEY, T., HANKINSON, S. E., HOWELL, A. & INGLE, J. 2007. Critical assessment of new risk factors for breast cancer: considerations for √

development of an improved risk prediction model. *Endocrine-Related Cancer,* 14**,** 169-187.

SESTAK, I., FORBES, J., ROSEANN, K., EDWARDS, R., HOWELL, A. & CUZICK, J. 2006a. Comparison of side-effect profiles during active treatment versus follow-up in the IBIS-1 tamoxifen prevention study. *Breast Cancer Research and Treatment,* 100**,** S55. √

SESTAK, I., KEALY, R., EDWARDS, R., FORBES, J. & CUZICK, J. 2006b. Influence of hormone replacement therapy on tamoxifen-induced vasomotor symptoms. *Journal of Clinical Oncology,* 24**,** 3991-3996. √

THIRLAWAY, K. 1995. Entry into the tamoxifen prevention trial depends on women's estimates of the population risk of breast cancer. *The Breast,* 4**,** 203-204. √

THIRLAWAY, K., FALLOWFIELD, L. & CUZICK, J. 1996a. The sexual activity questionnaire: A measure of women's sexual functioning. *Quality of Life Research,* 5**,** 81-90. √

THIRLAWAY, K., FALLOWFIELD, L., EVANS, G. & HOWELL, A. 1995. Entry into the tamoxifen prevention trial depends on women's estimates of the population risk of breast cancer. *Breast,* 4**,** 203-204. √

THIRLAWAY, K., FALLOWFIELD, L., NUNNERLEY, H. & POWLES, T. 1996b. Anxiety in women 'at risk' of developing breast cancer. *British Journal of Cancer,* 73**,** 1422-1424. √

VERONESI, A., PIZZICHETTA, M. A., FERLANTE, M. A., ZOTTAR, M., MAGRI, M. D., CRIVELLARI, D. & FOLADORE, S. 1998. Tamoxifen as adjuvant after surgery for breast cancer and tamoxifen or placebo as chemoprevention in healthy women: different compliance with treatment. *Tumori,* 84**,** 372-375. √

VERONESI, U. 1995a. Prevention of breast cancer with tamoxifen: the Italian study in hysterectomized women. *The Breast,* 4**,** 267. √

VERONESI, U. 2007. Tamoxifen: An enduring star. *Journal of the National Cancer Institute,* 99**,** 258-260. √

VERONESI, U., MAISONNEUVE, P., COSTA, A., ROTMENSZ, N. & BOYLE, P. 1999. Drop-outs in tamoxifen prevention trials. *The Lancet,* 353**,** 244. √

VERONESI, U., MAISONNEUVE, P., ROTMENSZ, N., BONANNI, B., BOYLE, P., VIALE, G., COSTA, A. & ITALIAN TAMOXIFEN STUDY, G. 2007. Tamoxifen for the prevention of breast cancer: late results of the Italian Randomized Tamoxifen Prevention Trial among women with hysterectomy. *Journal of the National Cancer Institute,* 99**,** 727-737. √

VERONESI, U., MAISONNEUVE, P., ROTMENSZ, N., COSTA, A., SACCHINI, V., TRAVAGLINI, R., D'AIUTO, G., LOVISON, F., GUCCIARDO, G., MURACA, M. G., PIZZICHETTA, M. A., CONFORTI, S., DECENSI, A., ROBERTSON, C. & BOYLE, P. 2003. Italian randomized trial among women with hysterectomy: tamoxifen and hormone-dependent breast cancer in high-risk women. *J Natl Cancer Inst,* 95**,** 160-5. √

VISVANATHAN, K., CHLEBOWSKI, R. T., HURLEY, P., COL, N. F., ROPKA, M., COLLYAR, D., MORROW, M., RUNOWICZ, C., PRITCHARD, K. I., HAGERTY, K., ARUN, B., GARBER, J., VOGEL, V. G., WADE, J. L., BROWN, P., CUZICK, J., KRAMER, B. S. & LIPPMAN, S. M. 2009. American Society of Clinical Oncology Clinical Practice Guideline Update on the Use of Pharmacologic Interventions Including Tamoxifen, Raloxifene, and Aromatase Inhibition for Breast Cancer Risk Reduction. *Journal of Clinical Oncology,* 27**,** 3235-3258. √

WARWICK, J., PINNEY, E., WARREN, R. M. L., DUFFY, S. W., HOWELL, A., WILSON, M. & CUZICK, J. 2003. Breast density and breast cancer risk factors in a high-risk population. *Breast,* 12**,** 10-16. √

WEST, A. & FALLOWFIELD, L. J. 1998. Psychosocial compliance and response rates in a chemoprevention randomised control trial of women at high risk to breast. *European Journal of Cancer*, 1998.	√

BMS Siblings List

Sibling Title	***Direct***	**Indirect**
CHAN, C. L. W., HO, R. T. H., FU, W. & CHOW, A. Y. M. 2006a. Turning curses into blessings: An eastern approach to psychosocial oncology. *Journal of Psychosocial Oncology,* 24, 15-32.	√	
CHAN, C. L. W., HO, R. T. H., LEE, P. W. H., CHENG, J. Y. Y., LEUNG, P. P. Y., FOO, W., CHOW, L. W. H., SHAM, J. S. T. & SPIEGEL, D. 2006b. A randomized controlled trial of psychosocial interventions using the psychophysiological framework for Chinese breast cancer patients. *Journal of Psychosocial Oncology,* 24, 3-26.	√	
CHAN, T. H., HO, R. T. & CHAN, C. L. 2007. Developing an outcome measurement for meaning-making intervention with Chinese cancer patients. *Psychooncology,* 16, 843-50.	√	
ERCEG-HURN, D. 2009. Commentary on liu et al. (2008) A study on the efficacy of body-mindspirit group therapy for patients with breast cancer. Journal of clinical nursing 17, 2539-2549. *Journal of Clinical Nursing,* 18, 1660-1661.		√
HO, R. T., POTASH, J. S., FU, W., WONG, K. P. & CHAN, C. L. 2009. Changes in breast cancer patients after psychosocial intervention as indicated in drawings. *Psychooncology*.	√	
HO, R. T. H., CHAN, C. L. W., HO, S. M. Y., LEE, P. W. H., SHAM, J. S. T. & CHOW, L. W. C. 2005. #57 Psychoneuroendocrine outcomes of psychosocial interventions for breast cancer patients in Hong Kong: Phase I results. *Brain Behavior and Immunity,* 19.	√	
HO, R. T. H., POTASH, J. S., FU, W., WONG, K. P. L. & CHAN, C. L. W. 2010. Changes in breast cancer patients after psychosocial intervention as indicated in drawings. *Psycho-Oncology,* 19, 353-360.	√	
LIU, C. J., HSIUNG, P. C., CHANG, K. J., LIU, Y. F., WANG, K. C., HSIAO, F. H., NG, S. M. & CHAN, C. L. W. 2008. A study on the efficacy of body-mind-spirit group therapy for patients with breast cancer. *Journal of Clinical Nursing,* 17, 2539-2549.		
LIU, Y. 2005. Personality, Self-evaluation and Emotional Adaptation of 63 Patients with Breast Cancer. *Zhongguo xin li wei sheng za zhi = Chinese mental health journal.,* 19, 820.	√	
SMITH, S. K. 2009. Health Status and Quality of Life Among Non-Hodgkin Lymphoma Survivors. *CANCER,* 115, 3312-3323.		√

Dexamethasone Siblings List

Sibling Title	*Direct*	Indirect
2001. Early postnatal dexamethasone therapy for the prevention of chronic lung disease. *Pediatrics,* 108**,** 741-8.		√
2007a. Early treatment of premature infants with steroids: Neurological sequelae. *Prescrire International,* 16**,** 108-109.		√
2009a. Late (>7 days) postnatal corticosteroids for chronic lung disease in preterm infants.		√
2010s. Early Postnatal Dexamethasone Therapy for the Prevention of Chronic Lung Disease -- The Vermont Oxford Network Steroid Study.	√	√
ANDAY, E. K. & CONWAY, D. 2003. Steroid therapy in the high-risk neonate: Benefits and risks. *Clinical Obstetrics and Gynecology,* 46**,** 190-210.		√
ANDERSON, M. E., BEDNAREK, F. J., DREYER, G., MAGOON, M. W., MERCIER, C. E., SOLL, R. F., GARLAND, J., HAVENS, P., MCAULIFFE, T., NELSON, R. & VERMONT OXFORD NETWORK STEROID, S. 2001. Early postnatal dexamethasone therapy for the prevention of chronic lung disease. *Pediatrics,* 108**,** 741-748.	√	
ANTTILA, E., PELTONIEMI, O., HAUMONT, D., HERTING, E., TER HORST, H., HEINONEN, K., KERO, P., NYKÄNEN, P., OETOMO, S. B. & HALLMAN, M. 2005. Early neonatal dexamethasone treatment for prevention of bronchopulmonary dysplasia. Randomised trial and meta-analysis evaluating the duration of dexamethasone therapy. *European Journal of Pediatrics,* 164**,** 472-481.		√
BARRINGTON, K. J., OHLSSON, A., AZIZ, K., DAVIS, D., LEE, S., SANKARAN, K., VAN AERDE, J., BLACKMON, L., BOULTON, J., CRANE, J., MCCOURT, C., REYNOLDS, L., LEMONS, J., SHAH, V., BELL, E. F., ENGLE, W. A., KANTO JR, W. P., MARTIN, G. I., MILLER, C. A., ROSENFELD, W., SPEER, M. E., STARK, A. R., ECORD, J., IYASU, S., LOCKWOOD, C. J. & WRIGHT, L. L. 2002. Postnatal corticosteroids to treat or prevent chronic lung disease in preterm infants. *L'administration postnatale de corticoïdes pour traiter ou prévenir les affections pulmonaires chroniques chez les prématurés,* 7**,** 20-28+37.		√
BHANDARI, A. & BHANDARI, V. 2009. Pitfalls, problems, and progress in bronchopulmonary dysplasia. *Pediatrics,* 123**,** 1562-1573.		√
BLACKMON, L. R., BELL, E. F., ENGLE, W. A., KANTO, W. P., MARTIN, G. I., MILLER, C. A., ROSENFELD, W., SPEER, M. E., STARK, A. R., BARRINGTON, K. J., OHLSSON, A., AZIZ, K., DAVIS, D., LEE, S., SANKARAN, K., VAN AERDE, J. & AMER ACAD, P. 2002. Postnatal corticosteroids to treat or prevent chronic lung disease in preterm infants. *Pediatrics,* 109**,** 330-338.		√
BOS, A. F., MARTIJN, A., VAN ASPEREN, R. M., HADDERS-ALGRA, M., OKKEN, A. & PRECHTL, H. F. 1998. Qualitative assessment of general movements in high-risk preterm infants with chronic lung disease requiring dexamethasone therapy. *J Pediatr,* 132**,** 300-6.		√
BOS, A. F., DIBIASI, J., TIESSEN, A. H. & BERGMAN, K. A. 2002. Treating preterm infants at risk for chronic lung disease with dexamethasone leads to an impaired quality of general movements. *Biol Neonate,* 82**,** 155-8.		√
BOSE, C., VAN MARTER, L. J., LAUGHON, M., O'SHEA, T. M., ALLRED, E. N., KARNA, P., EHRENKRANZ, R. A., BOGGESS, K. & LEVITON, A. 2009. Fetal Growth Restriction and Chronic Lung Disease Among Infants Born Before the 28th Week of Gestation. *Pediatrics*.		√
CHEDID, F., SHANTEER, S., HADDAD, H., MUSHARRAF, I., SHIHAB, Z., IMRAN, A., ADMA, H. A., SALMAN, N. & RAHMANI, A. 2009. Short-term outcome of very low birth weight infants in a developing country: comparison with the Vermont Oxford Network. *J Trop Pediatr,* 55**,** 15-9.	√	√

COLE, C. H. 2000. Postnatal glucocorticosteroid therapy for treatment and prevention of neonatal chronic lung disease. *EXPERT OPINION ON INVESTIGATIONAL DRUGS,* 9, 53-67.		√
GARLAND, J. S. 1999. A three-day course of dexamethasone therapy to prevent chronic lung disease in ventilated neonates: A randomized trial. *PEDIATRICS,* 104, 91-99.	√	
GRIER, D. G. & HALLIDAY, H. L. 2003. Corticosteroids in the prevention and management of bronchopulmonary dysplasia. *Seminars in Neonatology,* 8, 83-91.		√
GRONECK, P., OPPERMANN, M. & SPEER, C. P. 1993a. Levels of complement anaphylatoxin C5a in pulmonary effluent fluid of infants at risk for chronic lung disease and effects of dexamethasone treatment. *Pediatr Res,* 34, 586-90.		√
GRONECK, P., REUSS, D., GOTZE-SPEER, B. & SPEER, C. P. 1993b. Effects of dexamethasone on chemotactic activity and inflammatory mediators in tracheobronchial aspirates of preterm infants at risk for chronic lung disease. *J Pediatr,* 122, 938-44.		√
HALLIDAY, H. L. 1999. Clinical trials of postnatal corticosteroids: Inhaled and systemic. *BIOLOGY OF THE NEONATE,* 76, 29-40.		√
HALLIDAY, H. L. 2001a. A multicenter, randomized open study of early corticosteroid treatment (OSECT) in preterm infants with respiratory illness: Comparison of early and late treatment and of dexamethasone and inhaled budesonide. *PEDIATRICS,* 107, 232-240.		√
HALLIDAY, H. L. 2001b. Postnatal steroids: a dilemma for the neonatologist. *Acta Paediatr,* 90, 116-8.		√
HALLIDAY, H. L. 2002. Postnatal steroids are not recommended. *ACTA PHARMACOLOGICA SINICA,* 23, 8-10.		√
HALLIDAY, H. L. 2003. Postnatal dexamethasone: what is the real cost-benefit ratio? *Acta Paediatr,* 92, 888-9.		√
HALLIDAY, H. L. 2004a. Postnatal steroids and chronic lung disease in the newborn. *Paediatr Respir Rev,* 5 Suppl A, S245-8.		√
HALLIDAY, H. L. 2004b. Use of steroids in the perinatal period. *Paediatr Respir Rev,* 5 Suppl A, S321-7.		√
HALLIDAY, H. L. & EHRENKRANZ, R. A. 2000. Moderately early (7-14 days) postnatal corticosteroids for preventing chronic lung disease in preterm infants. *Cochrane Database Syst Rev*, CD001144.		√
HALLIDAY, H. L., EHRENKRANZ, R. A. & DOYLE, L. W. 2003a. Delayed (>3 weeks) postnatal corticosteroids for chronic lung disease in preterm infants. *Cochrane Database Syst Rev*, CD001145.		√
HALLIDAY, H. L., EHRENKRANZ, R. A. & DOYLE, L. W. 2003b. Early postnatal (<96 hours) corticosteroids for preventing chronic lung disease in preterm infants. *Cochrane Database Syst Rev*, CD001146.		√
HALLIDAY, H. L., EHRENKRANZ, R. A. & DOYLE, L. W. 2003c. Moderately early (7-14 days) postnatal corticosteroids for preventing chronic lung disease in preterm infants. *Cochrane Database Syst Rev*, CD001144.		√
HALLIDAY, H. L., EHRENKRANZ, R. A. & DOYLE, L. W. 2009a. Early (< 8 days) postnatal corticosteroids for preventing chronic lung disease in preterm infants. *Cochrane Database Syst Rev*, CD001146.		√
HALLIDAY, H. L., EHRENKRANZ, R. A. & DOYLE, L. W. 2009b. Late (>7 days) postnatal corticosteroids for chronic lung disease in preterm infants. *Cochrane Database Syst Rev*, CD001145.		√
HORBAR, J. D., ROGOWSKI, J., PLSEK, P. E., DELMORE, P., EDWARDS, W. H., HOCKER, J., KANTAK, A. D., LEWALLEN, P., LEWIS, W., LEWIT, E., MCCARROLL, C. J., MUJSCE, D., PAYNE, N. R., SHIONO, P., SOLL, R. F., LEAHY, K. & CARPENTER, J. H.	√	

2001. Collaborative quality improvement for neonatal intensive care. NIC/Q Project Investigators of the Vermont Oxford Network. *Pediatrics,* 107**,** 14-22.		
KAEMPF, J. W., CAMPBELL, B., SKLAR, R. S., ARDUZA, C., GALLEGOS, R., ZABARI, M., BROWN, A. & MCDONALD, J. V. 2003. Implementing potentially better practices to improve neonatal outcomes after reducing postnatal dexamethasone use in infants born between 501 and 1250 grams. *Pediatrics,* 111**,** e534-41.		√
KANTAK, A. D., GROW, J. L., OHLINGER, J., ADAMS, H. J., KNUPP, A. M. & LAVIN, J. P., JR. 2006. Management of high-order multiple births: application of lessons learned because of participation in Vermont Oxford Network collaboratives. *Pediatrics,* 118 Suppl 2**,** S159-68.	√	
KARI, M. A., HALLMAN, M., ERONEN, M., TERAMO, K., VIRTANEN, M., KOIVISTO, M. & IKONEN, R. S. 1994. Prenatal dexamethasone treatment in conjunction with rescue therapy of human surfactant: a randomized placebo-controlled multicenter study. *Pediatrics,* 93**,** 730-6.		√
KIEFER, A. S., WICKREMASINGHE, A. C., JOHNSON, J. N., HARTMAN, T. K., HINTZ, S. R., CAREY, W. A. & COLBY, C. E. 2009. Medical management of extremely low-birth-weight infants in the first week of life: A survey of practices in the United States. *American Journal of Perinatology,* 26**,** 407-418.		√
KUMAR, P. 2005. Effect of decreased use of postnatal corticosteroids on morbidity in extremely low birthweight infants. *American Journal of Perinatology,* 22**,** 77-81.		√
LIN, Y. J. 1998. Effects of early postnatal dexamethasone therapy on calcium homeostasis and bone growth in preterm infants with respiratory distress syndrome. *ACTA PAEDIATRICA,* 87**,** 1061-1065.		√
LIN, Y. J. 1999. Prevention of chronic lung disease in preterm infants by early postnatal dexamethasone therapy. *PEDIATRIC PULMONOLOGY,* 27**,** 21-26.		√
LIN, Y. J. 2005. The effects of early postnatal dexamethasone therapy on pulmonary outcome in premature infants with respiratory distress syndrome: A two-year follow-up study. *ACTA PAEDIATRICA,* 94**,** 310-316.		√
MCEVOY, C., BOWLING, S., WILLIAMSON, K., MCGAW, P. & DURAND, M. 2004. Randomized, double-blinded trial of low-dose dexamethasone: II. Functional residual capacity and pulmonary outcome in very low birth weight infants at risk for bronchopulmonary dysplasia. *Pediatric Pulmonology,* 38**,** 55-63.		√
MCGUIRE, W., FOWLIE, P. W. & SOLL, R. F. 2010. What has the Cochrane Collaboration ever done for newborn infants? *Archives of Disease in Childhood: Fetal and Neonatal Edition,* 95.	√	
NIXON, P. A., WASHBURN, L. K., MUDD, L. M., WEBB, H. H. & O'SHEA, T. M. 2010. Aerobic Fitness and Physical Activity Levels of Children Born Prematurely following Randomization to Postnatal Dexamethasone. *The Journal of Pediatrics*.		√
NIXON, P. A., WASHBURN, L. K., SCHECHTER, M. S. & OSHEA, T. M. 2007. Follow-up Study of a Randomized Controlled Trial of Postnatal Dexamethasone Therapy in Very Low Birth Weight Infants: Effects on Pulmonary Outcomes at Age 8 to 11 Years. *The Journal of Pediatrics,* 150**,** 345.		√
NYE, J. 2007. Postnatal corticosteroids in the treatment of chronic lung disease in the preterm infant: past, present, and future. *Neonatal Network,* 26**,** 293.		√
O'SHEA, T. M., WASHBURN, L. K., NIXON, P. A. & GOLDSTEIN, D. J. 2007. Follow-up of a randomized, placebo-controlled trial of dexamethasone to decrease the duration of ventilator dependency in very low birth weight infants: Neurodevelopmental outcomes at 4 to 11 years of age. *Pediatrics,* 120**,** 594-602.		√
PLADYS, P., MAGAGI, J., BEUCHÉE, A., LE BERRE, A., BRANGER, B. & BÉTRÉMIEUX, P. 2002. Respiratory distress syndrome, early postnatal dexamethasone and cavitating periventricular leukomalacia. *Détresse respiratoire, corticothérapie postnatale précoce et leucomalacie périventriculaire,* 9**,** 898-902.		√

ROGOWSKI, J. A., HORBAR, J. D., PLSEK, P. E., BAKER, L. S., DETERDING, J., EDWARDS, W. H., HOCKER, J., KANTAK, A. D., LEWALLEN, P., LEWIS, W., LEWIT, E., MCCARROLL, C. J., MUJSCE, D., PAYNE, N. R., SHIONO, P., SOLL, R. F. & LEAHY, K. 2001. Economic implications of neonatal intensive care unit collaborative quality improvement. *Pediatrics,* 107**,** 23-9.	√	
ROMAGNOLI, C. 1999a. Early postnatal dexamethasone for the prevention of chronic lung disease in high-risk preterm infants. *Intensive care medicine.,* 25**,** 717.		√
ROMAGNOLI, C. 1999b. Effect on growth of two different dexamethasone courses for preterm infants at risk of chronic lung disease - A randomized trial. *PHARMACOLOGY,* 59**,** 266-274.		√
ROMAGNOLI, C. 2002. Controlled trial of early dexamethasone treatment for the prevention of chronic lung disease in preterm infants: A 3-year follow-up. *PEDIATRICS,* 109.		√
ROMAGNOLI, C., ZECCA, E., LUCIANO, R., TORRIOLI, G. & TORTOROLO, G. 2002. A three year follow up of preterm infants after moderately early treatment with dexamethasone. *Arch Dis Child Fetal Neonatal Ed,* 87**,** F55-8.		√
SHINWELL, E. S. 2000. Early postnatal dexamethasone treatment and increased incidence of cerebral palsy. *ARCHIVES OF DISEASE IN CHILDHOOD,* 83**,** F177-F181.		√
SHINWELL, E. S., KARPLUS, M., ZMORA, E., REICH, D., ROTHSCHILD, A., BLAZER, S., BADER, D., YURMAN, S., DOLFIN, T., KUINT, J., MILBAUER, B., KOHELET, D., GOLDBERG, M., ARMON, Y., DAVIDSON, S., SIROTA, L., AMITAI, M., ZARETSKY, A., BARAK, M. & GOTTFRIED, S. 1996. Failure of early postnatal dexamethasone to prevent chronic lung disease in infants with respiratory distress syndrome. *Arch Dis Child Fetal Neonatal Ed,* 74**,** F33-7.		√
SINKIN, R. A., DWECK, H. S., HORGAN, M. J., GALLAHER, K. J., COX, C., MANISCALCO, W. M., CHESS, P. R., D'ANGIO, C. T., GUILLET, R., KENDIG, J. W., RYAN, R. M. & PHELPS, D. L. 2000. Early dexamethasone-attempting to prevent chronic lung disease. *Pediatrics,* 105**,** 542-8.		√
SOLL, R. F. 1999. Early postnatal dexamethasone therapy for the prevention of chronic lung disease. *PEDIATRIC RESEARCH,* 45**,** 226A-226A.	√	
SOLL, R. F. 2001. Early postnatal dexamethasone therapy for the prevention of chronic lung disease. *Pediatrics,* 108**,** 741-748.	√	
SOLL, R. F. 2003. Corticosteroids for the treatment and prevention of chronic lung disease. *Acta Paediatrica, International Journal of Paediatrics,* 92**,** 886-888.	√	
TAPIA I, J. L. & KATTAN I, J. 2003. Postnatal steroids and bronchopulmonar dysplasia: Risks and benefits. *Corticoides postnatal y displasia broncopulmonar: Beneficios versus riesgos,* 74**,** 70-80.		√
TSU, F., YEH *, Y. J. L., CHAO, C. H., YUNG, J. C. & CHYI, H. L. 2010. Early Dexamethasone Therapy in Preterm Infants: A Follow-up Study -- Yeh et al. 101 (5): e7 -- Pediatrics.		√
VERMONT OXFORD NETWORK STEROID STUDY, G. 2001. Early postnatal dexamethasone therapy for the prevention of chronic lung disease. *Pediatrics,* 108**,** 741-748.	√	
VIJAYAKUMAR, E., SOLL, R. F., BRACKEN, M. B. & SINCLAIR, J. C. 1994. Early postnatal corticosteroids for the prevention of chronic lung disease (CLD): Meta-analysis of clinical trials. *Pediatric Research,* 37**,** 276A.	√	
WALSH, M. C., YAO, Q., HORBAR, J. D., CARPENTER, J. H., LEE, S. K. & OHLSSON, A. 2006. Changes in the use of postnatal steroids for bronchopulmonary dysplasia in 3 large neonatal networks. *Pediatrics,* 118.		√
WALTHER, F. J., FINDLAY, R. D. & DURAND, M. 2003. Adrenal suppression and extubation rate after moderately early low-dose dexamethasone therapy in very preterm infants. *Early Human Development,* 74**,** 37-45.		√

WILSON-COSTELLO, D., WALSH, M. C., LANGER, J. C., GUILLET, R., LAPTOOK, A. R., STOLL, B. J., SHANKARAN, S., FINER, N. N., VAN MEURS, K. P., ENGLE, W. A. & DAS, A. 2009. Impact of postnatal corticosteroid use on neurodevelopment at 18 to 22 months' adjusted age: Effects of dose, timing, and risk of bronchopulmonary dysplasia in extremely low birth weight infants. *Pediatrics,* 123.		√
YEH, T. F. 1997. Early postnatal dexamethasone therapy for the prevention of chronic lung disease in preterm infants with respiratory distress syndrome: A multicenter clinical trial. *PEDIATRICS,* 100**,** art.-e3.		√
YEH, T. F. 1998. Early dexamethasone therapy in preterm infants: A follow-up study. *PEDIATRICS,* 101**,** art.-e7.		√
YEH, T. F., LIN, Y. J., LIN, H. C., HUANG, C. C., HSIEH, W. S., LIN, C. H. & TSAI, C. H. 2004. Outcomes at school age after postnatal dexamethasone therapy for lung disease of prematurity. *New England Journal of Medicine,* 350, 1304-1313.		√

IVF Siblings List

Sibling Title	***Direct***	**Indirect**
AFNAN, M. 2009. Identifying real differences in live birth rates between HMG and rFSH in IVF. *REPRODUCTIVE BIOMEDICINE ONLINE,* 18**,** S25-S30.		√
AGRAWAL, R., CONWAY, G. S., ENGMANN, L., BEKIR, J. S. & JACOBS, H. S. 1998. Implications of using follicle-stimulating hormone preparations depleted of luteinizing hormone to achieve follicular growth in in vitro fertilization. *Gynecol Endocrinol,* 12**,** 9-15.	√	
AGRAWAL, R., HOLMES, J. & JACOBS, H. S. 2000. Follicle-stimulating hormone or human menopausal gonadotropin for ovarian stimulation in in vitro fertilization cycles: a meta-analysis. *Fertil Steril,* 73**,** 338-43.	√	
ANDERSEN, A. N., DEVROEY, P. & ARCE, J. C. 2006. Clinical outcome following stimulation with highly purified hMG or recombinant FSH in patients undergoing IVF: a randomized assessor-blind controlled trial. *Hum Reprod,* 21**,** 3217-27.	√	
BOSCH, E. 2008. Highly purified hMG versus recombinant FSH in ovarian hyperstimulation with GnRH antagonists - a randomized study. *HUMAN REPRODUCTION,* 23**,** 2346-2351.		√
BRUYNESTEYN, K., BONSE, G. J., BRAAT, D. D. M., FAUSER, B., DEVROEY, P. & VAN GENUGTEN, M. L. L. 2005. Economic evaluation of the administration of follitropin- with a pen device. *Reproductive BioMedicine Online,* 11**,** 26.	√	
COOMARASAMY, A., AFNAN, M., CHEEMA, D., VAN DER VEEN, F., BOSSUYT, P. M. & VAN WELY, M. 2008. Urinary hMG versus recombinant FSH for controlled ovarian hyperstimulation following an agonist long down-regulation protocol in IVF or ICSI treatment: a systematic review and meta-analysis. *Hum Reprod,* 23**,** 310-5.	√	
DEVROEY, P., MANNAERTS, B., SMITZ, J., COELINGH BENNINK, H. & VAN STEIRTEGHEM, A. 1993. First established pregnancy and birth after ovarian stimulation with recombinant human follicle stimulating hormone (Org 32489). *Hum Reprod,* 8**,** 863-5.	√	
FILICORI, M. M. D., COGNIGNI, G. E. M. D., POCOGNOLI, P. M. D., TABARELLI, C. M. D., FERLINI, F. M. D., PERRI, T. M. D. & PARMEGIANI, L. B. 2003. Comparison of controlled ovarian stimulation with human menopausal gonadotropin or recombinant follicle-stimulating hormone *1. *Fertility and Sterility,* 80**,** 390.	√	
FRITZSCHE, H., SCHMEDEMANN, R. K. A., MICHELMANN, H. W. & SIEBZEHNRUBL, E. 2005. The cumulus-oocyte complex and its significance for IVF treatment. *Geburtshilfe Und Frauenheilkunde,* 65**,** 850-861.		√

GERMOND, M., DESSOLE, S., SENN, A., LOUMAYE, E., HOWLES, C., BELTRAMI, V., DEVROEY, P. & COELINGH BENNINK, H. 1992. Successful in-vitro fertilisation and embryo transfer after treatment with recombinant human FSH. *The Lancet,* 339**,** 1170.	√	
HOMPES, P. G. A., BROEKMANS, F. J., HOOZEMANS, D. A. & SCHATS, R. 2008. Effectiveness of highly purified human menopausal gonadotropin vs. recombinant follicle-stimulating hormone in first-cycle in vitro fertilization-intracytoplasmic sperm injection patients. *Fertility and Sterility,* 89**,** 1685-1693.		√
KEYE, W. R., WEBSTER, B., DICKEY, R., SOMKUTI, S., CRAIN, J. & SCOBEY, M. J. 2005. Subcutaneously administered Menopur(R), a new highly purified human menopausal gonadotropin, causes significantly fewer injection site reactions than Repronex(R) in subjects undergoing in vitro fertilization. *Reproductive Biology and Endocrinology,* 3.		√
KILANI, Z. 2003. A prospective, randomized, controlled trial comparing highly purified hMG with recombinant FSH in women undergoing ICSI: ovarian response and clinical outcomes. *HUMAN REPRODUCTION,* 18**,** 1194-1199.		√
LLOYD, A., KENNEDY, R., HUTCHINSON, J. & SAWYER, W. 2003. Economic evaluation of highly purified menotropin compared with recombinant follicle-stimulating hormone in assisted reproduction. *Fertility and Sterility,* 80**,** 1108-1113.		√
PLATTEAU, P., ANDERSEN, A. N., LOFT, A., SMITZ, J., DANGLAS, P. & DEVROEY, P. 2008. Highly purified HMG versus recombinant FSH for ovarian stimulation in IVF cycles. *Reproductive Biomedicine Online,* 17**,** 190-198.	√	
PLATTEAU, P., SMITZ, J., ALBANO, C., SORENSEN, P., ARCE, J. C. & DEVROEY, P. 2004a. Exogenous luteinizing hormone activity may influence the treatment outcome in in vitro fertilization but not in intracytoplasmic sperm injection cycles. *Fertil Steril,* 81**,** 1401-4.	√	
PLATTEAU, P., SMITZ, J., ALBANO, C., SORENSEN, P., ENGELS, S., ARCE, J. C. & DEVROEY, P. 2004b. *Differential effects of LH activity administration in IVF and ICSI cycles*.	√	
SMITZ, J., ANDERSEN, A. N., DEVROEY, P. & ARCE, J. C. 2007. Endocrine profile in serum and follicular fluid differs after ovarian stimulation with HP-hMG or recombinant FSH in IVF patients. *Hum Reprod,* 22**,** 676-87.	√	
THE EUROPEAN AND ISRAELI STUDY GROUP ON HIGHLY PURIFIED MENOTROPIN VERSUS RECOMBINANT FOLLICLE-STIMULATING, H. 2002. Efficacy and safety of highly purified menotropin versus recombinant follicle-stimulating hormone in in vitro fertilization/intracytoplasmic sperm injection cycles: a randomized, comparative trial a *1. *Fertility and Sterility,* 78**,** 520.	√	
TREW, G. H. 2007. Comparing highly purified hMG and rFSH in patients undergoing IVF. *Human Reproduction,* 22**,** 1797-1798.		√
WECHOWSKI, J., CONNOLLY, M., SCHNEIDER, D., MCEWAN, P. & KENNEDY, R. 2009. Cost-saving treatment strategies in in vitro fertilization: a combined economic evaluation of two large randomized clinical trials comparing highly purified human menopausal gonadotropin and recombinant follicle-stimulating hormone alpha. *Fertil Steril,* 91, 1067-76.		√

Appendix Seven

IDEATel Publications

Peer Reviewed Publications:

1. Bakken S, Grullon-Figueroa L, Izquierdo R, Lee NJ, Morin P, Palmas W, Teresi J, Weinstock RS, Shea S, Starren J. Development, validation, and use of English and Spanish versions of the telemedicine satisfaction and usefulness questionnaire. JAMIA 2006;13:660-667.
2. Hilliman CA, Cimino JJ, Lai AM, Kaufman DR, Starren JB, Shea S, for the IDEATel Consortium. The effects of redesigning the IDEATel architecture on glucose uploads. Telemedicine and e-Health 2009;15:248-254.
3. Homenko DR, Morin PC, Eimicke JP, Teresi JA, Weinstock RS. Food insecurity and food choices in rural older adults with diabetes receiving nutrition education via telemedicine. Journal of Nutrition Education and Behavior. In press.
4. Izquierdo R, Lagua CT, Meyer S, Ploutz-Snyder RJ, Palmas W, Eimicke JP, Kong J, Teresi JA, Shea S, Weinstock RS. Telemedicine intervention effects on waist circumference and body mass index in the IDEATel Project. Diabetes Technology & Therapeutics 2009. In press.
5. Izquierdo R, Meyer S, Starren J, Goland R, Teresi J, Shea S, Weinstock RS. Detection and remediation of medically urgent situations using telemedicine case management for older patients with diabetes mellitus. Therapeutics and Clinical Risk Management 2007;3(3):1-5.
6. Kaufman D, Pevzner J, Hilliman C, Weinstock RS, Shea S, Starren J. Redesigning a telehealth diabetes management program for a digital divide seniors population. Home Health Care Management & Practice 2006;18(3):223-234.
7. Kaufman DR, Patel VL, Hilliman C, Morin PC, Pevzner J, Weinstock RS, Goland R, Shea S, Starren J. Usability in the real world: assessing medical information technologies in patients' homes. J Biomed Inform 2003;36:45-60.
8. Kaufman DR, Pevzner J, Rodriguez M, Cimino JJ, Ebner S, Field L, Moreno V, McGuinness C, Weinstock RS, Shea S, Starren J. Understanding workflow in teleheath video visits: observations from the IDEATel Project. J Biomed Informatics. In press.
9. Lai AM, Kaufman DR, Starren JB, Shea S, for the IDEATel Consortium. Evaluation of a remote training approach for teaching seniors to use a telehealth system. International Journal of Medical Informatics 2009;78:732-744.
10. Lai AM, Starren JB, Kaufman DR, Mendoca EA, Palmas W, Nieh J, Shea S, for the IDEATel Consortium. The Remote Patient Education in a Telemedicine Environment Architecture (REPETE). Telemedicine and e-Health 2008;14:355-361.
11. Moran A, Palmas W, Field L, Bhattarai J, Schwartz JE, Weinstock RS, Shea S. Cardiovascular autonomic neuropathy is associated with microalbuminuria in older patients with type 2 diabetes. Diabetes Care 2004;27:972-977.
12. Moran A, Palmas W, Pickering TG, Schwartz JE, Field LS, Weinstock RS, Shea S. Ambulatory blood pressure and office blood pressure taken with an automated device are independently associated with albuminuria in older subjects with type 2 diabetes mellitus. Hypertension 2006;47(5):955-961.
13. Morin PC, Wolff LT, Eimicke JP, Teresi JA, Shea S, Weinstock RS. Record media used by primary care providers in medically underserved areas of upstate New York was not pivotal to clinical result in the Informatics for Diabetes Education and Telemedicine [IDEATel] project. Inform Prim Care 2009;17:103-112.

14. Palmas W, Moran A, Pickering T, Eimicke JP, Teresi J, Schwartz JE, Field L, Weinstock RS, Shea S. Association of ambulatory pulse pressure with progression of urinary albumin excretion in a multiethnic cohort of older patients with type 2 diabetes mellitus. Hypertension 2006;48(2):301-308. [Editorial: Knudson ST, Anderson NH, Mogensen CE. Ambulatory pulse pressure and progression of albuminuria in type 2 diabetes: evidence provided, new questions emerge. Hypertension 2006;48:207-208.]

15. Palmas W, Pickering T, Eimicke JP, Moran A, Teresi J, Schwartz JE, Field L, Weinstock RS, Shea S. Value of ambulatory arterial stiffness index and 24-hour pulse pressure to predict progression of albuminuria in elderly people with diabetes mellitus. Am J Hypertension 2007;20:493-500.

16. Palmas W, Pickering TG, Moran A, Teresi J, Schwartz JE, Eguchi K, Field LS, Weinstock RS, Shea S. Nocturnal blood pressure elevation predicts progression of albuminuria in type-2 diabetes. J Clin Hypertension 2008;10:12-30.

17. Palmas W, Pickering TG, Teresi J, Schwartz JE, Field LS, Weinstock RS, Shea S. Telemedicine home blood pressure measurements and progression of albuminuria in elderly people with diabetes. Hypertension 2008;51:1-7.

18. Palmas W, Pickering TG, Teresi J, Schwartz JE, Moran A, Weinstock RS, Shea S. Ambulatory blood pressure monitoring and all-cause mortality in elderly people with diabetes. Hypertension 2009;53:120-127. [Editorial. Bursztyn M, Ben-Div UZ. Diabetes mellitus and 24-hour ambulatory blood pressure monitoring: broadening horizons of risk assessment. Hypertension 2009;53:110-11.]

19. Palmas W, Shea S, Starren J, Teresi JA, Ganz ML, Burton TM, Pashos CL, Blustein J, Field L, Morin PC, Izquierdo RE, Silver S, Eimicke JP, Lantigua RA, Weinstock RS. Medicare Payments, health care service use, and telemedicine implementation costs in a randomized trial comparing telemedicine case management with usual care in medically underserved participants with diabetes mellitus (IDEATel). J Am Med Inform Assoc. In press.

20. Palmas W, Teresi J, Morin P, Wolff LT, Field L, Eimicke JP, Capps L, Prigollini A, Orbe I, Weinstock RS, Shea S. Recruitment and enrollment of rural and urban medically underserved elderly into a randomized trial of telemedicine case management for diabetes care. Telemedicine and e-Health 2006;12(5):601-607.

21. Palmas, W, Teresi J, Weinstock RS, Shea S. Acceptability to primary care providers of a telemedicine diabetes case management intervention. J Telemedicine and Telecare 2008;14:306-308.

22. Robinson KS, Morin PC, Shupe J, Izquierdo R, Ploutz-Snyder R, Meyer S, Teresi J, Starren J, Shea S, Weinstock RS. Use of three computer training methods in elderly underserved rural patients enrolled in a diabetes telemedicine program. Computers, Informatics, Nursing, In press.

23. Sandberg J, Izquierdo R, Goland R, Morin PC, Palmas W, Larson CD, Shea S, Weinstock RS. A qualitative study of the experiences and satisfaction of direct telemedicine providers in diabetes care. Telemedicine and e-Health. 2009;15:742-750.

24. Sarkar IN, Starren J. Desiderata for personal electronic communication in clinical systems. JAMIA 2002. 9(3):2009-216.

25. Shea S, Starren J, Weinstock RS, Knudson PE, Teresi J, Holmes D, Palmas W, Field L, Goland R, Tuck C, Hripcsak G, Capps L, Liss D. Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel) Project: rationale and design. JAMIA 2002;9:49-60.

26. Shea S, Weinstock RS, Starren J, Teresi J, Palmas W, Field L, Morin P, Goland R, Izquierdo RE, Wolff LT, Ashraf M, Hilliman C, Silver S, Meyer S, Holmes D, Petkova E Capps L, Lantigua RA, for the IDEATel Consortium. A randomized trial comparing telemedicine case management with usual care in older, medically underserved patients with diabetes mellitus. JAMIA 2006;13:40-51.

27. Shea S, Weinstock RS, Teresi J, Palmas W, Starren J, Cimino JJ, Lai AM, Field L, Morin PC, Goland R, Izquierdo RE, Ebner S, Silver S, Petkova E, Kong J, Eimicke JP, for the IDEATel Consortium. A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus: 5 year results of the IDEATel Study. JAMIA 2009;16:446-456.

28. Shea S. Health delivery system changes required when integrating telemedicine into existing treatment flows of information and patients. J Telemedicine and Telecare 2006; 122 (Suppl. 2):S2:85-90.

29. Starren J, Hripcsak G, Sengupta S, Abbruscato CR, Knudson PE, Weinstock RS, Shea S. Columbia University's Informatics for Diabetes and Telemedicine (IDEATel) Project: technical implementation. JAMIA 2002;9:25-36.

30. Starren J, Tsai C, Bakken S, Aidala A, Morin PC, Hilliman C, Weinstock RS, Goland R, Teresi J, Shea S, for the IDEATel Consortium. The role of nurses in installing telehealth technology in the home. Comput Inform Nurs 2005;23:181-189.

31. Starren, J. Federal Demonstration Project Tackles Diabetes with Telemedicine. Caring 2005;24(1):6-11.

32. Trief PM, Morin PC, Izquierdo R, Teresi J, Eimicke J, Goland R, Starren J, Shea S, Weinstock RS. Depression and glycemic control in elderly ethnically diverse patients with diabetes: The IDEATel Project. Diabetes Care 2006;29(4):830-835.

33. Trief PM, Morin PC, Izquierdo R, Teresi J, Starren J, Shea S, Weinstock R. Marital quality and diabetes outcomes: The IDEATel Project. Families, Systems & Health. 2006;24(3):318-331. [Commentary: Anderson BJ. Families and chronic illness research: targeting transitions and tools – commentary on Trief et al. (2006). Families, Systems, & Health 2006;24(3):332-335.]

34. Trief PM, Sandberg J, Izquierdo R, Morin PC, Shea S, Brittain R, Feldhousen EB, Weinstock RS. Diabetes managements assisted by telemedicine: patient perspectives. Telemedicine and e-Health 2008;14:647-655.

35. Trief PM, Teresi J, Izquierdo R, Morin P, Goland R, Field L, Eimicke J, Brittain R, Starren J, Shea S, Weinstock R. Psychosocial outcomes of telemedicine case management for elderly patients with diabetes: IDEATel, a randomized trial. Diabetes Care 2007;30:1266-68.

36. Trief PM, Teresi JA, Eimicke SP, Shea S, Weinstock RS. Improvement in diabetes self-efficacy and gyclaemic control using telemedicine in a sample of older, ethnically diverse individuals who have diabetes: the IDEATel Project. Age and Aging 2009;38:219-225.

37. Tsai CC, Starren J. Patient participation in electronic medical records. JAMA 2001;285 (13):1765. JAMA Online.

38. Tudiver F, Wolff LT, Morin PC, Teresi J, Palmas W, Starren J, Shea S, Weinstock RS. Primary care providers' perceptions of home diabetes telemedicine care in the IDEATel Project. Journal of Rural Health 2007;23:55-61.

39. Weinstock RS, Izquierdo R, Goland R, Palmas W, Teresi JA, Eimicke JP, Shea S and the IDEATel Consortium. Lipid treatment in ethnically diverse underserved older adults with diabetes: statin use, goal attainment and health disparities in the IDEATel telemedicine project (Letter to the Editor). J Amer Geriatric Soc. In press.

40. West SP, Lagua C, Trief PM, Izquierdo R, Weinstock RS. Goal setting using telemedicine in rural underserved older adults with diabetes: Experiences from the IDEATel Project. Telemedicine and e-Health. In press.

Book Chapters:

1. Brennan PF, Starren J. Telehealth and consumer health informatics. Computer Applications in Health Care and Biomedicine, Third Edition. 2006.

2. Shea S, Starren J, Weinstock RS. Case Study: Informatics for Diabetes Education and Telemedicine (IDEATel). In: Detmer D, Steen E, eds. The Academic Health Center: Leadership and Performance. Cambridge University Press, 2005: 270-271.

3. Starren J, Weinstock RS, Palmas W, Izquierdo RE, Morin P, Kaufman D. Diabetes (Chapter 16). In: Wootton R, Dimmick SL, Kedvar JC, eds. Home Telehealth: Connecting Care Within the Community, Royal Society of Medicine Press, 2006:170-83.

4. Weinstock RS, Knudson PE, Shea S. Using the internet and telemedicine to expand access to health care. In: Evans CH, Osterweis M, eds. Insurance and Beyond: Expanding Access to Quality Health Care. Association of Academic Health Centers; Washington, DC, 2001: 67-79.

Conference Papers

1. Kaufman DR, Starren J, Patel VL, Morin PC, Hilliman C, Pevzner J, Weinstock R.S, Goland R, Shea S. A cognitive framework for understanding barriers to the productive use of a diabetes home telemedicine system. Proceedings of the American Medical Informatics Association Annual Fall Symposium 2003;356-360.

2. Lai A, Nieh J, Laine A, Starren J. Thin client performance for remote 3-D image display, Proceedings of the American Medical Informatics Association Annual Fall Symposium. Hanley & Belfus. 2003;904-908.

3. Lai AM, Kaufman DR, Starren J. Training digital divide seniors to use a telehealth system: A remote training approach. Proceedings of the American Medical Informatics Association Annual Fall Symposium 2006;459-63.

4. Starren J, Abbruscato C R, Porter C, Ring G. Making it work: Success factors in multi-vendor projects. HIMSS Proceedings 2002.

5. Starren J, Sengupta S, Hripcsak G, Ring G, Klerer R, Shea S. Making grandma's data secure: A security architecture for home telemedicine. Proceedings of the American Medical Informatics Association Annual Fall Symposium 2001: 657-661.

Abstracts:

1. Chused AE, Payne P, Starren J. Novel techniques for survey and classification studies to improve patient centered websites. Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2006;891.

2. Deitzer J, Payne PR, Starren JB. Coverage of clinical trials tasks in existing ontologies. Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2006;903.

3. Izquierdo R, Weinstock RS, Meyer S, Starren J, Goland R, Teresi J, Shea S. Critically important clinical interventions in an interactive diabetes case management model of primary care providers and a diabetes team using telemedicine in the Upstate New York cohort of the Informatics for Diabetes Education and Telemedicine Project (IDEATel). The Endocrine Society 87th Annual Meeting, San Diego, CA. P3-422, 2005.

4. Kaufman DR, Cimino J, Starren JB, Rodriguez M, McGuinness C, Moreno V, Weinstock RS, Shea S. Understanding dimensions of complexity in nurse case managers workflow in a telemedicine program. Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2007;1006.

5. Kaufman DR, Starren J. A methodological framework for evaluating mobile health devices. Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2006;978.

6. Lai A, Starren J, Shea S. A novel solution for remote training of home telemedicine patients. American Telemedicine Association Annual Meeting, 2006.

7. Lai A, Starren J, Shea S. Architecture for remote training of home telemedicine patients. Proceedings of the American Medical Informatics Association Annual Fall Symposium 2005;1015.

8. Lai AM, Nieh J, Starren JB. REPETE2: A next generation home telemedicine architecture. Proceedings of the American Medical Informatics Association Annual Fall Symposium. 2007:1020.

9. Pevzner J, Kaufman DR, Hilliman C, Shea S, Weinstock RS, Starren J. Developing computer skills and competencies in seniors. Proceedings of the American Medical Informatics Association Annual Fall Symposium 2005;1078.

10. Starren J, Hilliman C, Weinstock RS, Shea S. Theater style demonstration: The Informatics for Diabetes Education and Telemedicine (IDEATel) Project. Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2006;1195.

11. Starren J, Lee N, Bakken S. Standardizing remote monitoring data using the LOINC data standard. American Telemedicine Association, 2006.

12. Starren J, Shea S, Weinstock RS, Goland R, Morin PC, Wolff LT, Izquierdo R, Meyer S, Teresi J, Palmas W, Field L, Silver S, Hilliman C, Ashraf M, for the IDEATel Consortium. Effect of telemedicine management of diabetes in the Medicare population: The IDEATel Project. Tromso Telemedicine and e-Health Conference – TteC, 89, 2004.

13. Starren J, Shea S, Weinstock RS, Knudson P, Morin PC, Teresi J, Holmes D, Palmas P, et al. The IDEATel Project: Up and Running. J Telemed. e-Health 2002;8(2):212.

14. Starren JB, Kaufman DR, Payne PR. Human computer interaction issues in clinical trials management systems. Proceedings of the American Medical Informatics Association Annual Fall Symposium, 2006;1109.

15. Trief PM, Izquierdo R, Morin PC, Teresi J, Shea S, Starren J, Weinstock RS. Does depression predict intervention outcomes – The IDEATel Project. Diabetes 2005;54(Suppl 1):A465, 2005.

16. Trief PM, Morin PC, Izquierdo R, Teresi J,Shea S, Weinstock RS. Marital quality and diabetes outcomes of IDEATel: A telemedicine intervention for the elderly. Psychosomatics 2005;67(1), A73.

17. Trief PM, Sandberg JG, Izquierdo R, Morin P, Shea S, Brittain RA, Feldhousen EB, Weinstock RS. The telemedicine experience of elderly diabetes patients: a qualitative study. Annals of Behavioral Medicine, 2007, 33 (Suppl): S096.

18. Trief PM, Weinstock RS, Shea S, Starren J, Teresi J, for the IDEATel Group. The effects of telemedicine case management for elderly diabetes patients on psychosocial outcomes: The IDEATel Project. Diabetes 2006; 55 Supplement (1) A432.

19. Wolff LT, Tudiver T, Palmas W, Starren J. Goland R, Morin P, Teresi J, Weinstock R, Shea S. Provider's perceptions of *Informatics for Diabetes Education and Telemedicine (IDEATel)*, a diabetes home telemedicine system: Preliminary report. The North American Primary Care Group annual meeting, Banff, Canada, October 26, 2003.

BMS Publications

<u>Peer Reviewed Publications</u>:

1. Chan, C. L. W., Ho, R. T. H., Fu, W. & Chow, A. Y. M. (2006a). Turning curses into blessings: An eastern approach to psychosocial oncology. Journal of Psychosocial Oncology, 24, 15-32.

2. Chan, C. L. W., Ho, R. T. H., Lee, P. W. H., Cheng, J. Y. Y., Leung, P. P. Y., Foo, W., Chow, L. W. H., Sham, J. S. T. & Spiegel, D. (2006b). A randomized controlled trial of psychosocial interventions using the psychophysiological framework for Chinese breast cancer patients. Journal of Psychosocial Oncology, 24, 3-26.

3. Chan, T. H. Y., Ho, R. T. H., & Chan, C. L. W. (2007). Developing an outcome measurement for meaning-making intervention with Chinese cancer patients. Psycho-Oncology, 16(9): 843-850.

4. Ho, R. T. H., Potash, J. S., Fu, W., Wong, K. P. L. & Chan, C. L. W. (2010). Changes in breast cancer patients after psychosocial intervention as indicated in drawings. Psycho-Oncology, 19, 353-360.

5. Ho, R. T., Potash, J. S., Fu, W., Wong, K. P. & Chan, C. L. (2009). Changes in breast cancer patients after psychosocial intervention as indicated in drawings. Psychooncology.

6. Ho, R. T. H., Chan, C. L. W., Ho, S. M. Y., Lee, P. W. H., Sham, J. S. T. & Chow, L. W. C. (2005). #57 Psychoneuroendocrine outcomes of psychosocial interventions for breast cancer patients in Hong Kong: Phase I results. Brain Behavior and Immunity, 19.

7. Ho, R. T. H. (2005). Effect of Dance Movement therapy on Chinese Cancer Patients - A Pilot Study in Hong Kong. The Arts in Psychotherapy, 32(5): 337-345.

8. Ho, R. T. H. (2005). Regaining Balance Within: Dance Movement Therapy With Chinese Cancer Patients In Hong Kong. American Journal of Dance Therapy, 27(2): 87-99.

9. Ho, R. T. H., Chan, C. L. W., & Ho, S. M. Y. (2004). Emotional control in Chinese female cancer survivors. Psycho-Oncology, 13(11), 808-817.

10. Hsiao, F. H., Jow, G. M., Kuo, W.H., Chang, K.J., Liu, Y.F., Ho, R. T. H., Ng, S. M., Chan, C. L. W., Lai, Y. M., Chen. Y.T. (2012). The Effects of Psychotherapy on Psychological Well-Being and Diurnal Cortisol patterns in Breast Cancer Survivors. Psychotherapy and Psychosomatics, 81:173–182. DOI: 10.1159/000329178.

11. Liu, C. J., Hsiung, P. C., Chang, K. J., Liu, Y. F., Wang, K. C., Hsiao, F. H., Ng, S. M. & Chan, C. L. W. (2008). A study on the efficacy of body-mind-spirit group therapy for patients with breast cancer. Journal of Clinical Nursing, 17, 2539-2549.

12. Liu, Y. (2005). Personality, Self-evaluation and Emotional Adaptation of 63 Patients with Breast Cancer. Zhongguo xin li wei sheng za zhi = Chinese mental health journal., 19, 820.

<u>Book Chapters</u>:

1. Ho, A.H.Y., Lo, P.H.Y., Leung, P.P.Y., Ho R.T.H. and Chan, C.L.W. (2009). Efficacy Studies of the Body-Mind-Spirit Model on Serving Cancer Patients. In: Chan, L.W., Fan, F.M. & Leung, P.Y., The Holistic Body-Mind-Spirit Well-being Model – Chinese Culture and Group Counselling. (pp. 187-199). Beijing, China Light Industrial Press (in Chinese).

2. Ho, R. T. H., Chan C.L.W. Lo P.H.Y., Leung, P. Y. Leung (2009).The Efficacy of the Body-Mind-Spirit Intervention and Social Support groups on Chinese breast cancer patients, In M.Y. Lee,

C.L.W. Chan, S.M. Ng, P. P.Y. Leung (Eds.), Integrative Body-Mind-Spirit Social Work: An Empirically Based Approach to Assessment and Treatment. (pp.217-234). New York: The Oxford University Press.

Conference Papers:

1. Ho, R.T.H. & Sham, T.O.T. (2012). What words can tell besides number in a randomized controlled trial: preliminary qualitative findings of a dance/movement therapy program on the quality of life and symptom cluster among Chinese cancer patients undergoing radiotherapy.

2. Ho, R. T. H., Potash, J. S., Fu, W. Chan, C. L. W., Cheung I.K.M. (2009). Images Speak More Than Words: Integration of Drawings in the Assessment Process for Understanding the Changes of Cancer Patients after a Psychosocial Intervention. Psychooncology, 18 (Suppl 2): S51-S52.

3. Ho, R.T.H., Chan, C. L. W., Ho, S. M. Y., Lee, P. W. H., Sham, J. S. T., Chow, L. W. C., et al. (2007). Active Treatment with Professionals Yields Promising and Sustained Advantages for Chinese Breast Cancer Patients: A Randomized Controlled Trial of Psychosocial Interventions in Hong Kong. Psycho-Oncology, 16 (9):S13.

4. Lo, P. H. Y., Chan, C. L. W., & Ho, R.T.H. (2008). Linking health practices and spirituality in Chinese breast cancer survivors. Psycho-Oncology, 17 (Suppl. 2): S100.

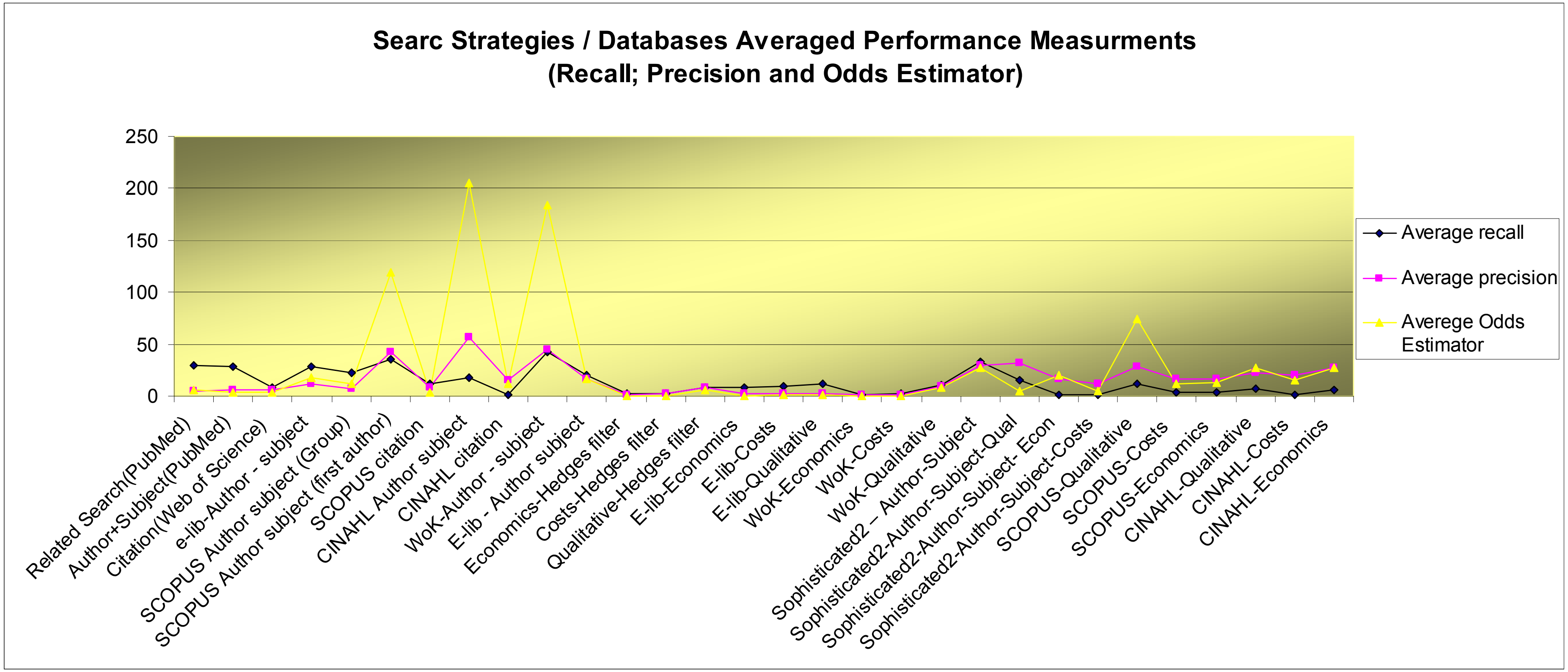
Meeting Abstracts:

1. Ho, R. T. H., Cheung, I. K. M., Chan, C. L. W. & Yip, P. S. F. (2012). Chinese breast cancer patients undergoing radiotherapy: Role of fatigue and its influences on their quality of life. Asian-Pacific Journal of Clinical Oncology, 8(Suppl. 3): S224.

2. Lo, P. H. Y., Ho, R. T. H., Chan, J. S. M., Cheung, I. K. M. & Chan, C. L. W. (2012). When social support rises above all: Determinants of end-of-life-treatment and long-term anxiety and depression for Chinese female cancer survivors. Asian-Pacific Journal of Clinical Oncology, 8(Suppl. 3): S131.

Appendix Eight

The Association Between Odds Estimator and Precision



Appendix Nine

Information Sheet

Information Retrieval for Systematic Reviews

Hello,

My name is Faten Hamad, a PhD student at Aberystwyth University. I'm currently working on a research project that aims to evaluate different search methods for identifying RCTs and other kinds of studies of different research designs that are associated or/and related to a seed study (usually an RCT).

With most systematic reviews (SR) the emphasis has been placed on quantitative research designs, particularly randomised control trials as these provide the least possible bias. These help to answer the question – is one intervention better, on average, than another intervention? Several different types of studies may need to be clearly identified and retrieved to increase not just the robustness and reliability of SR, but transfer of knowledge into practice. These studies handle different issues on the same topic such as economic issues, qualitative research on patient or professional attitudes, and process implementation. These associated studies can be grouped under one term that reflect their relationship which is "Sibling studies", and I have classified these as either direct or indirect siblings. Importantly, sibling studies provide information on how to implement an intervention in an efficient and effective manner, and thus they complement each other and the RCT.

The main focus of this research is the direct siblings, which are the studies that are based on or emerge from the seed study and aim to investigate other aspect that may interfere, affect or explain the intervention output using either the same or a different research design. Moreover, sibling studies must share at least one author of the seed study (if not, I have designated these as indirect siblings).

I created a gold standard using relative recall (a pool of direct siblings that are retrieved from all databases which is used in this research) in order to compare my actual dataset against it. I chose five RCTs, and tested different search strategies, search filters and databases.

I would be grateful for your help in checking my set of direct siblings for comprehensiveness, in order to be able to assess individual database retrieval performance. If I missed some important studies that should be included in the set, I would be grateful for the reference details. I have attached the set of direct siblings. It would also be helpful to know if there are any sibling studies that do not share an author with the chosen seed study and which are therefore indirect siblings.

If you have any questions or would like further details regarding the research then please contact me:

Faten F. Hamad
E-mail: fatoon_82@yahoo.com

Supervisor Details:
Christine Urquhart
E-mail: ahcjurquhart@btinternet.com
University account: cju@aber.ac.uk

References

Acid, S., De Campos, L. M., Fernández-Luna, J. M., & Huete, J. F. (2003). An information retrieval model based on simple Bayesian networks. *International Journal of Intelligent Systems*, *18*(2), pp. 251-265.

Alhenshiri, A., Shepherd, M., & Watters, C. (2010). Improving web search for information gathering: visualization in effect. Proceedings of the forth workshop of Human-Computer Interaction and Information Retrieval, HCIR 2010: 22 August 2010, New Brunswick, NJ USA.

Agoritsas, T., Merglen, A., Delphine S Courvoisier, D. S., Combescure, C., Garin, N., Perrier, A., & Perneger, T. V. (2012). Sensitivity and predictive value of 15 pubmed search strategies to answer clinical questions rated against full systematic reviews. *Journal of Medical Internet Research*, *14*(3), e85. doi:10.2196/jmir.2021.

Ananiadou, S., Kell, D. B., & Tsujii, J.-ichi. (2006). Text mining and its potential applications in systems biology. *Trends in biotechnology*, *24*(12), pp. 571-579

Ananiadou, S., Procter, R., Rea, B., & Sasaki, Y. (2007). Supporting systematic reviews using text mining. *Third International Conference on e-Social Science.* Ann Arbor, Michigan, USA. Retrieved 16 September 2009 from http://www.nactem.ac.uk/assert/projectReports2.php.

Ananiadou, S., Rea, B., Okazaki, O., Procter, R., & Thomas, J. (2009). Supporting systematic reviews using text mining. *Social Science Computer Review; 27*(4), pp. 509-523. doi:10.1177/0894439309332293.

Andronis, L., Barton, P., & Bryan, S. (2009). Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. *Health Technology Assessment (Winchester, England)*, *13*(29), iii, ix-xi, 1-61. doi:10.3310/hta13290.

Atsaros, G., Spinellis, D., & Louridas, P. (2008). Site-Specific versus general purpose web search engines: A comparative evaluation. In *Proceedings of the 2008 Panhellenic Conference on Informatics, PC: 28 - 30 August, 2008, IEEE Computer Society, Washington, DC* (pp. 44-48). USA, Washington, DC: IEEE Computer Society.

Avenell, a, Handoll, H. H., & Grant, a M. (2001). Lessons for search strategies from a systematic review, in The Cochrane Library, of nutritional supplementation trials in patients after hip fracture. *The American Journal of Clinical Nutrition*, *73*(3), pp. 505-510.

Baeza-Yatez, R., & Robeiro-Neto, B. (1999). *Modern Information Retrieval*. Essex: Addison Wesely Longman.

Bates, T., Anić, A., Marusić, M., & Marusić, A. (2004). Authorship criteria and disclosure of contributions: comparison of 3 general medical journals with different author contribution forms. *The journal of the American Medical Association: JAMA*, *292*(1), pp. 86-88.

Baykoucheva, S. (2010). Selecting a database for drug literature retrieval: A comparison of MEDLINE, Scopus, and Web of Science. *Science & Technology Libraries*, *29*(4), pp. 276-288.

Berkman, N., & Viswanathan, M. (2008). Development of a tool to evaluate the quality of observational studies. The 16th Cochrane Colloquium. Freiburg, Germany. (Special issue of Elsevier 2008, 'Abstract of the 16th Cochrane Collooquium "Evidence in the era of globalization"'. *German Journal for Evidence and Quality in Health Care*, *102*, pp. 7 – 99.

Bernstam, E. (2001). MedlineQBE (Query-by-Example). *Proceedings / AMIA ... Annual Symposium. AMIA Symposium*, pp. 47-51.

Borgman, C. L. (1999). What are digital libraries? Competing visions. *Information Processing and Management, 35*(3), pp. 227 – 243.

Borlund, P. (2003). The concept of relevance in IR. *Journal of the American Society for Information Science and Technology*, *54*(10), pp. 913-925.

Boulenger, S., Nixon, J., Drummond, M., Ulmann, P., Rice, S., & de Pouvourville, G. (2005). Can economic evaluations be made more transferable? *The European Journal Of Health Economics : HEPAC : Health Economics in Prevention and Care*, *6*(4), pp. 334-46.

Breslow, N. (1981). Odds Ratio estimators when the data are sparse. *Biometrika*, *68*(1), 73.

Buckland, M., & Gey, F. (1994). The relationship between Recall and Precision. *Journal of the American Society for Information Science*, *45*(1), pp. 12-19.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ (Clinical Research ed.), 321*(7262), pp. 694-696.

Cattan, M., White, M., Bond, J., & Learmouth, A. (2005). Preventing social isolation and loneliness among older people: a systematic review of health promotion interventions. *Ageing and Society, 25*(1), pp. 41-67.

Chaudhuri, S., Das, G., Hristidis, V., & Weikum, G. (2004). Probabilistic ranking of database query results. *Proceedings of the Thirtieth international conference on Very large data bases - Volume 30* (VLDB '04): 29 August - 3 September 2004, Toronto, Canada (pp. 888-899). VLDB Endowment, Toronto, Canada.

Cheng, K., Preston, C., Ashby, D., O'Hea, U., & Smyth, R. L. (1998). Time to publication as full reports of abstracts of randomized controlled trials in cystic fibrosis. *Pediatric pulmonology, 26*(2), 101-5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9727760.

Chowdhury, G. G. (2004). Introduction to modern information retrieval (2nd ed.). London: Facet Publishing.

Chi, E. H., & Mytkowicz, T. (2008). Understanding the efficiency of social tagging systems using information theory. In: Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia: 19 - 21 June 2008, Pittsburgh, PA, USA (pp. 81-88). New York, NY: ACM.

Clarke, C. L. A., Cormack, G. V., & Tudhope, E. A. (2000). Relevance ranking for one to three term queries. *Information Processing and Management, 36*, pp. 291-311.

Coiera, E., Westbrook, J. I., & Rogers, K. (2008). Clinical decision velocity is increased when meta-search filters enhance an evidence retrieval system. *Journal of the American Medical Informatics Association : JAMIA, 15*(5), pp. 638-646. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2528038/.

Convey J. (1992). Online information retrieval: An introductory manual to principles and practice (4th ed.). London: Library Association Publishing Ltd.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new medical research council guidance. *BMJ*, 337, a1655. doi: 10.1136/bmj.a1655.

Croft, W. B. (1987). Approaches to intelligent information retrieval. *Information Processing and Management, 23*(4), pp. 249-254. Cited in Ingwersen, P. (1992). Information Retrieval Interaction. London: Taylor Graham.

Day, D., Furlan, A., Irvin, E., & Bombardier, C. (2005). Simplified search strategies were effective in identifying clinical trials of pharmaceuticals and physical modalities. *Journal of Clinical Epidemiology, 58*(9), pp. 874-881.

Dee, C. R. (2007). The development of the Medical Literature Analysis and Retrieval System (MEDLARS). *Journal of the Medical Library Association, 95*(4), pp.417-425.

Devadason, F. J., & Lingam, P. P. (1996). A methodology for the identification of information needs of users. The Challenge of Change: Libraries and Economic Development: Proceedings of 62nd IFLA General Conference and Council: 25 – 31 August 1996, Beijing, China.

Dickersin, K., Scherer, R., Suci, E. S. T., & Gil-Montero, M. (2002). Problems with indexing and citation of articles with group authorship. *The journal of the American Medical Association: JAMA, 287*(21), pp. 2772-2774. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12038908

Dixon-Woods, M., Fitzpatrick, R., & Roberts, k. (2001). Including qualitative research in systematic reviews: opportunities and problems. *Journal of Evaluation in Clinical Practice, 7*(2), pp. 125-133.

Donald, O. C. (2007). Looking for information: a survey of research on information seeking, needs, and behaviour (2nd ed.) [online]. London: Elsevier Ltd. Retrieved 20 June 2010 from http://books.google.co.uk.

Drummond, M., Barbieri, M., Cook, J., Glick, H. A., Lis, J., Malik, F., Reed, H. D., Rutten, F., Sculpher, M., & Severens, J. (2009). Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. Value in Health,12(4), pp. 409-418.

EBSCO Publishing. Visual search. [Online]. Retrieved 9 October 2011 form http://www.ebscohost.com/academic/visual-search.

Essers, B. A. B., Seferina, S. C., Tjan-Heijnen, V. C. G., Severens, J. L., Novák, A., Pompen, M., Oron, U. H., & Joore, M. A. 2010. Transferability of model-based economic evaluations: The case of Trastuzumab for the adjuvant treatment of HER2-Positive early breast cancer in the Netherlandsvhe_683 375..380. *Value in Health, 13*(4), pp. 375-380.

European and Israeli Study Group on Highly Purified Menotropin versus Recombinant Follicle-Stimulating Hormone.(2002). Efficacy and safety of highly purified menotropin versus recombinant follicle-stimulating hormone in vitro fertilization/intracytoplasmic sperm injection cycles: a randomized, comparative trial. *Fertility and Sterility, 78*(3): pp. 520-528.

Evans, D. (2002). Database searches for qualitative research. *Journal of the Medical Library Association : JMLA, 90*(3), pp. 290-293.

Falagas, M. E., Pitsouni, E. I., Malietzis, G. a, & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 22*(2), pp. 338-342.

Fallowfield, L., Fleissig, A., Edwards, R., West, A., Powles, T. J., Howell, A., & Cuzick, J. (2001). Tamoxifen for the prevention of breast cancer: psychosocial impact on women participating in two randomized controlled trials. *American Journal of Clinical Oncology, 9*(7), pp. 1885-1892.

Fernández, E., & García, A. M. (2003). Accuracy of referencing of Spanish names in Medline. *Lancet, 361*(9354), pp. 351-352. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12559897.

Field, A. (2009). Discovering Statistics Using SPSS (3rd ed.). London: SAGE publications.

Flemming, K., & Briggs, M. (2007). Electronic searching to locate qualitative research: evaluation of three strategies. *Journal of advanced nursing, 57*(1), pp. 95-100. doi:10.1111/j.1365-2648.2006.04083.x

Flottorp, S., Havelsrud, K., & Oxman, A. D. (2003). Process evaluation of a cluster randomized trial of tailored interventions to implement guidelines in primary care--why is it so hard to change practice? *Family Practice, 20*(3), pp. 333-339.

Furnas, G. W., Fake, C., von Ahn, L., Schachter, J., Golder, S., Fox, K., Davis, M., Marlow, C., & Naaman, M. (2006). Why do tagging systems work? CHI '06

Extended Abstracts on Human Factors in Computing Systems: 22 - 27 April 2006, Montréal, Québec, Canada (pp. 36-39). New York, NY: ACM.

Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all. *Journal of Information Science, 30*(4), pp. 378-391.

Glanville, J. M., Lefebvre, C., Miles, J. N. V., & Camosso-Stefinovic, J. (2006). How to identify randomized controlled trials in MEDLINE: ten years on. *Journal of the Medical Library Association : JMLA, 94*(2), pp. 130-136.

Glenton, C., Carlsen, B., Eccles, M., Francis, J., Grimshaw, J., & Pope, C. (2008). Why don't GPs follow guidelines? Increasing our understanding of complex interventions by combining qualitative and quantitative reviews. The 16th Cochrane Colloquium. Freiburg, Germany. (Special issue of Elsevier 2008, Abstract of the 16th Cochrane Collooquium: Evidence in the era of globalization. *German Journal for Evidence and Quality in Health Care, 102*, pp. 7 – 99.

Glenton, C., Lewin, S., & Scheel, I. B. (2011).Still too little qualitative research to shed light on results from reviews of effectiveness trials: A case study of a Cochrane review on the use of lay health workers. *Iplementation Science, 6*:53. doi: 10.1186/1748-5908-6-53.

Golder, S., McIntosh, H. M., Duffy, S., & Glanville, J. (2006). Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. *Health Information and Libraries Journal, 23*(1), pp. 3-12.

Golder, S., & Loke, Y. (2009). Search strategies to identify information on adverse effects: a systematic review. *Journal of the Medical Library Association : JMLA, 97*(2), pp. 84-92.

Golder, S., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32*(2), pp. 198-208.

Goss, C., Lowenstein, S., Roberts, I., & DiGuiseppi, C. (2007). Identifying controlled studies of alcohol-impaired driving prevention: designing an effective search strategy. *Journal of Information Science, 33*(2), pp. 151-162.

Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), pp. 213-228.

Gray, A., Raikou, M., McGuire, A., Fenn , P., Stevens, R., Cull, C., Stratton, I., Adler, A., Holman , R., & Turner, R. (2000). Cost effectiveness of an intensive blood glucose control policy in patients with type 2 diabetes: economic analysis alongside randomised controlled trial (UKPDS 41). *BMJ, 320*(7246), pp. 1373-1378.

Green, S. B., & Salkind, N. J. (2005). Using SPSS for Windows and Macintosh: Analysing and understanding data (4th ed.). New Jersey: Pearson education, Inc.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. Milbank Quart, 82(4), pp. 581-629. Quoted in May, C. (2006). A rational model for assessing and evaluating complex interventions in health care. *BMC Health Services Research, 6*: 86. doi:10.1186/1472-6963-6-86

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (eds)', Syntax and Semantic, 3, pp. 41- 58. Quoted in Xu, Y.,C., & Chen, Z. (2006). Relevance Judgment – What Do Information Users Consider beyond Topicality?. *Journal of the American Society for Information Science and Technology, 57*(7), pp. 961-973.

Gwizdka, J. (2009). Assessing cognitive load on web search tasks. T*he Ergonomics Open Journal, 2*(1), pp. 114 – 123.

Hammond, T., Hannay, T., Lund, B., & Scott, J. (2005). Social bookmarking tools: A general review. D-Lib Magazine, 11(4). Retrieved from http://www.dlib.org/dlib/april05/hammond/04hammond.html.

Hansen, H. P. (Speaker). (2008). Qualitative Research Methods "Patients and Research". Podcast. The 16th Cochrane Colloquium. Freiburg, Germany. Retrieved 21 November 200 from http://www.cochrane.org/multimedia/colloquium_2008/slidecasts/plenary1_2_helle_hansen.html.

Harden, A., Garcia, J., Oliver, S., Rees, R., Shepherd, J., Brunton, G., & Oakley, A. (2004). Applying systematic review methods to studies of people's views: an example from public health research. *Journal of Epidemiology and Community Health, 58*(9), pp. 794-800.

Haynes, R. B., McKibbon, K. A., Wilczynski, N. L., Walter, S. D., & Werre, S. R. (2005). Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ (Clinical Research ed.), 330*(7501), pp. 1179- 1186.

Haynes, R. B., & Wilczynski, N. L. (2004). Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ (Clinical Research ed.), 328*(7447), pp. 1040-1042.

Health Information Research Unit (HiRU). Hedges. HiRU 2007a [Updated on 2007 June 29]. Retreived 15th Septemeber 2009 form http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx.

Hersh, W. R., Hickam, D. H., Hayens, R.B., & McKibbon, K. A. (1994). A performance and failure analysis of SAPHIRE with MEDLINE test collection. *Journal of the American Medical Informatics Association, 1*(1), pp. 51-60.

Heimstra, D. (2009). Information retrieval models. In: Goker A, Davies J. Information Retrieval: Searching in the 21st Century. John Wiley and Sons.

Herskovic, J., Tanaka, L., Hersh, W., & Bernstram, E. V. (2007). A day in the life of PubMed: Analysis of a typical day's query log. *Journal of the American Medical Informatics Association: JAMIA, 14*(2), pp. 212-220.

Higgins, JPT. & Green, S (eds). (2011). Cochrane handbook for systematic reviews of interventions Version 5.1.0. The Cochrane Collaboration. Retrieved 29 April 2012 from www.cochrane-handbook.org.

Hoogendam, A., de Vries Robbé, P. F., Stalenhoef, A. F. H., & Overbeke, a J. P. M. (2009). Evaluation of PubMed filters used for evidence-based searching: validation using relative recall. *Journal of the Medical Library Association : JMLA, 97*(3), pp. 186-193.

Hopewell S, ClarkeMJ, Stewart L, Tierney J. Time to publication for results of clinical trials. Cochrane Database of Systematic Reviews 2007; Issue 2. Art. No.: MR000011. doi: 10.1002/14651858.MR000011.pub2.

Howes, F., Doyle, J., Jackson, N., & Waters, E. (2004). Evidence-based public health: the importance of finding "difficult to locate" public health and health promotion intervention studies for systematic reviews. *Journal of Public Health, 26*(1), pp. 101-104.

Hripcsak, G. Wilcox, A. Reference Standards, Judges, and Comparison Subjects: roles for experts in evaluating system performance. Journal of American Medical Informatics Association; 9(1), pp. 1-15

Huang, A. W. C., & Chuang, T. R. (2009). Social tagging, online communication, and Peircean semiotics: a conceptual framework. *Journal of Information Science, 35*(3), pp. 340-357.

Huić, M., Marušić, M., & Marušić, A. (2011). Completeness and changes in registered data and reporting bias of randomized controlled trials in ICMJE journals after trial registration policy. *PloS One, 6*(9), e25258. doi:10.1371/journal.pone.0025258

Hulscher, M. E. J. L., Laurant, M. G. H., & Grol, R. P. T. M. (2003). Process evaluation on quality improvement interventions. *Quality & safety in health care, 12*(1), pp. 40-46.

Ingwersen, P. (1992). Information Retrieval Interaction. London: Taylor Graham.

International Committee of Medical Journal Editors (ICMJE). (2012). Retrieved from http://www.icmje.org/.

Janes, J. W. (1994). Other people's judgments: A comparison of users' and others' judgments of document relevance, topicality, and utility. *Journal of the American Society for Information Science, 45*(3), pp. 160-171.

Joanna Briggs Institute. (2008). Joanna Briggs Institute Reviewers' Manual: 2008 edition. [e-book] Adelaide: Joanna Briggs Institute. Retrieved 18 March 2010 from http://www.joannabriggs.edu.au/pdf/JBIReviewManual_CiP11449.pdf .

Kagolovsky, Y., & Moehr, J. R. (2003). Terminological problems in information retrieval. *Journal of Medical Systems, 27*(5), pp. 399-408.

Kim, J. (2009). Describing and predicting information-seeking behavior on the Web. *Journal of the American Society and Information Science Technology, 60*(4), pp. 679-93.

Kowalski, G. J., & Maybury, M. T. (2000). Information storage and retrieval (2nd ed.). Boston: Kluwer Academic Publishers. Retrieved 1 May 2010 from http://books.google.co.uk.

Land, J. M., & Altman, D. G. (2000). Statistics Notes: The odds ratio. *BMJ, 320*, p. 1468.

Lemeshow, A. R., Blum, R. E., Berlin, J. a, Stoto, M. a, & Colditz, G. a. (2005). Searching one or two databases was insufficient for meta-analysis of observational studies. *Journal of Clinical Epidemiology, 58*(9), 867-73. doi:10.1016/j.jclinepi.2005.03.004

Lesk M. (1996). The Seven ages of information retrieval. IFLA UDT Core Programme Occasional Pape. Retrieved from: http://archive.ifla.org/VI/5/op/index.htm [Accessed 22th May 2010].

Lewin, S., Glenton, C., & Oxman, a. D. (2009). Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. *BMJ, 339*, b3496. doi:10.1136/bmj.b3496

Liddy, E. D. (2005). Automatic document retrieval (2nd ed.). In Encyclopedia of Language and Linguistics. Elsevier Press. Retrieved from http://surface.syr.edu/istpub/51/

Lin, J., DiCuccio, M., Grigoryan, V., & Wilbur, W. J. (2008). Navigating information spaces: A case study of related article search in PubMed. *Information Processing & Management, 44*(5), pp. 1771-1783.

Lin, J., & Wilbur, W. J. (2007). PubMed related articles: a probabilistic topic-based model for content similarity. *BMC Bioinformatics, 8*, 423. doi:10.1186/1471-2105-8-423

Liu, X., & Altman, R. B. (1998). Updating a bibliography using the related articles function within PubMed. *Proceedings / AMIA ... Annual Symposium. AMIA Symposium,* (5), pp. 750-754.

Liu, C.-J., Hsiung, P.-C., Chang, K.-J., Liu, Y.-F., Wang, K.-C., Hsiao, F.-H., Ng, S.-M., et al. (2008). A study on the efficacy of body-mind-spirit group therapy for patients with breast cancer. *Journal of Clinical Nursing, 17*(19), pp. 2539-49.

Lu, Z., Kim, W. & Wilbur, W. J. (2009). Evaluation of query expansion using MeSH in PubMed. *Information Retrieval, 12*(1), pp. 69–80.

MacDonald, J., Bath, P., & Booth, A. (2011). Information overload and information poverty: challenges for healthcare services managers? *Journal of Documentation, 67*(2), pp. 238-263.

McGowan, J., Sampson, M., & Lefebvre, C. (2010). Evidence based library and information practice. *Evidence Based Library and Information Practice, 5*(1), PP. 149-154.

Manning, C. D., Raghavan, P., & Schütze, H. (2009). An introduction to information retrieval. Cambridge: Cambridge University Press. Retrieved from http://nlp.stanford.edu/IR-book/information-retrieval-book.html.

Marušić, A., & Marušić, M. (2010). A contribution to the authorship debate: Can we trust definitions and declarations?. *The Journal of the European Medical Writers Association, 19*(1), pp. 14-17.

Mathes, A. (2004). Folksonomies -- cooperative classification and communication through shared metadata. Computer Mediated Communication.

May, C. (2006). A rational model for assessing and evaluating complex interventions in health care. *BMC Health Services Research, 6*, 86. doi:10.1186/1472-6963-6-86

May, C. R., Mair, F., Finch, T., MacFarlane, A., Dowrick, C., Treweek, S., Rapley, T., et al. (2009). Development of a theory of implementation and integration: Normalization Process Theory. *Implementation Science : IS, 4*, 29. doi:10.1186/1748-5908-4-29

May, C. R., Mair, F. S., Dowrick, C. F., & Finch, T. L. (2007). Process evaluation for complex interventions in primary care: understanding trials using the normalization process model. *BMC Family Practice, 8*, 42. doi:10.1186/1471-2296-8-42

Mays, N., & Pope, C. (1995). Qualitative Research: Rigour and qualitative research. *BMJ, 311*: 109. doi: http://dx.doi.org/10.1136/bmj.311.6997.109

McGowan, J., & Sampson, M. (2005). Systematic reviews need systematic searchers. *Journal of the Medical Library Association : JMLA, 93*(1), pp. 74-80.

McGowan, J., Sampson, M., & Lefebvre, C. (2010).An evidence based checklist for the Peer Review of Electronic Search Strategies (PRESS EBC). *Evidence Based Library and Information Practice, 5*(1), pp. 149-154.

McKibbon, K. A., Wilczynski, N. L., & Haynes, R. B. (2006). Developing optimal search strategies for retrieving qualitative studies in PsycINFO. *Evaluation & the Health Professions, 29*(4), pp. 440-454.

Milette, K., Roseman, M., & Thombs, B. D. (2011). Transparency of outcome reporting and trial registration of randomized controlled trials in top psychosomatic and behavioral health journals: A systematic review. J*ournal of Psychosomatic Research, 70*(3), pp. 205-17.

Moher, D., Tsertsvadze, A., Tricco, A. C., Eccles, M., Grimshaw, J., Sampson, M., & Barrowman, N. (2007). A systematic review identified few methods and strategies describing when and how to update systematic reviews. *Journal of Clinical Epidemiology, 60*(11), pp. 1095-1104.

Morrison, D., & Mair, F. S. (2011). Telehealth in practice: using Normalisation Process Theory to bridge the translational gap. *Primary Care Respiratory Journal : Journal of the General Practice Airways Group, 20*(4), pp. 351-352.

Muddamalle, R. M. (1998). Natural language versus controlled vocabulary in information retrieval: A Case Study in Soil Mechanics. *Journal of the American Society for Information Science, 49*(10), pp. 881-887.

National Library of Medicine (USA). PubMed Help. Appendix: Computation of Related Articles. [Online]. Retrieved 15 November 2009 from http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helppubmed&part=pubmedhelp#pubmedhelp.Computation_of_Relat.

National Library of Medicine. (2011). MEDLINE/PubMed Data Element (Field) Descriptions. Retrieved 04 April 2012 from http://www.nlm.nih.gov/bsd/mms/medlineelements.html#au.

National Library of Medicine. (2010). Authorship in MEDLINE. Retrieved 04 April 2012 from http://www.nlm.nih.gov/pubs/factsheets/authorship.html.

National Library of Medicine. (2012a). PubMed Tutorial: Automatic Term Mapping. Retrieved 25 February 2012 from http://www.nlm.nih.gov/bsd/disted/pubmedtutorial/020_040.html.

National Library of Medicine. (2012b). PubMed help. Retreived 17 August 2012 from http://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.I3_Full_Author_trans.

National Library of Medicine. (2012c). MEDLINE Data Changes — 2013. Retrieved 15 December 2012 from http://www.nlm.nih.gov/pubs/techbull/nd12/nd12_medline_data_changes_2013.html.

Neufeld, M. L., & Cornog, M. (1986). Database history: From dinosaurs to compact discs. *Journal of the American Society for Information Science, 37*(4), pp. 183-190. doi:10.1002/asi.4630370404

NHS Centre for Reviews and Dissemination (CRD). (2009). Systematic Reviews: CRD's guidance for undertaking reviews in health car. CRD, University of York.

Noyes, J, & Popay, J. (2007). Directly observed therapy and tuberculosis: how can a systematic review of qualitative research contribute to improving services? A qualitative meta-synthesis. *Journal of Advanced Nursing, 57*(3), pp. 227-243.

Nunn, R. (2008). Evidence-based medicine and limits to the literature search. *Journal of Evaluation in Clinical Practice, 14*(5), pp. 672-678.

Oakley A., Strange V., Bonell C., Allen E., & Stephenson J. (RIPPLE Study Team). (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ, 332*(7538), pp. 413-416.

Ogilvie, D., Hamilton, V., Egan, M., & Petticrew, M. (2005). Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go? *Journal of Epidemiology and Community Health*, *59*(9), pp. 804-808.

Ounger, P., & Boddy, K. (2008). When is a search not a search? A comparison of searching the AMED complementary health database via EBSCOhost, OVID and DIALOG. *Health Information and Libraries Journal, 26*(2), pp. 126-135.

Palmer, A. J., Annemans, L., Roze, S., Lamotte, M., Lapuerta, P., Chen, R., Gabriel, S., et al. (2004). Cost-effectiveness of early irbesartan treatment versus control (standard antihypertensive medications excluding ACE inhibitors, other angiotensin-2 receptor antagonists, and dihydropyridine calcium channel blockers) or late irbesartan treatment in patien. *Diabetes Care*, *27*(8), pp. 1897-903.

Palmer, A. J., Valentine, W. J., Chen, R., Mehin, N., Gabriel, S., Bregman, B., & Rodby, R. a. (2008). A health economic analysis of screening and optimal treatment of nephropathy in patients with type 2 diabetes and hypertension in the USA. *Nephrology, Dialysis, Transplantation : Official Publication of the European Dialysis and Transplant Association - European Renal Association*, *23*(4), pp. 1216-1223.

Pao, M. L., & Worthen, D. B. (1989). Retrieval effectiveness by semantic and citation searching. *Journal of the American Society for Information Science*, *40*(4), pp. 226-235.

Papaioannou, D., Sutton, A., Carroll, C., Booth, A., & Wong, R. (2009). Literature searching for social science systematic reviews: consideration of a range of search techniques. *Health Information and Libraries Journal*, *27*(2), 114-22.

Parkhill, A. F., Clavisi, O., Pattuwage, L., Chau, M., Turner, T., Bragge, P., & Gruen, R. (2011). Searches for evidence mapping: effective, shorter, cheaper. *Journal of the Medical Library Association : JMLA*, *99*(2), pp. 157-60.

Pawson, R., Greenhalgh, T., Harvey, G. & Walshe K. (2004). *Realist synthesis: an introduction*. Manchester: ESRC Research Methods Programme. Retrieved 15 March 2010 from http://www.ccsr.ac.uk/methods/publications/documents/RMPmethods2.pdf.

Pope, C., Royen, P., & Baker, R. (2002). Qualitative methods in research on healthcare quality. *Quality and Safety in Health Care*, *11*(2), pp. 148-152.

Raghavan, V., Bollmann, P., & Jung, G. S. (1989). A critical investigation of recall and precision as measures of retrieval system performance. *ACM Transactions on Information Systems*, *7*(3), pp. 205-229.

Relevo, R. (2012). Chapter 4: Effective search strategies for systematic reviews of medical tests. *Journal of General Internal Medicine, 27*(supplement 1), pp. 28-32. doi: 10.1007/s11606-011-1873-8.

Robinson, K. A., & Dickersin, K. (2002). Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *International Journal of Epidemiology*, *31*(1), pp. 150-153.

Royle, P., & Waugh, N. (2003). Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system. *Health Technology Assessment (Winchester, England)*, *7*(34), iii, ix-x, pp. 1-51.

Royle, P., & Waugh, N. (2005). A simplified search strategy for identifying randomised controlled trials for systematic reviews of health care interventions: a comparison with more exhaustive strategies. *BMC Medical Research Methodology*, *5*, 23. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1183214&tool=pmcentrez&rendertype=abstract

Ruiz-Pérez, R., Delgado López-Cózar, E., & Jiménez-Contreras, E. (2002). Spanish personal name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *Journal of the Medical Library Association : JMLA*, *90*(4), pp. 411-430.

Sackett, D. L., & Rosenberg, W. M. C. (1995). On the need for evidence-based medicine. *Journal of Public Health*, *17*(3), pp. 330-334.

Sainani, K. (2008). Mining Biomedical Literature: Using computers to extract knowledge nuggets. *Biomedical Computation Review, 4*(3), pp. 16–27. Available from: http://www.citeulike.org/group/3266/article/4642422 [Accessed 13th September 2009].

Sampson, M. J. (2009). *Updating Searches for Systematic Reviews*. PhD thesis. Aberystwyth University, Aberystwyth.

Sampson, M., Barrowman, N. J., Moher, D., Clifford, T. J., Platt, R. W., Morrison, A., Klassen, T. P., et al. (2006a). Can electronic search engines optimize screening of search results in systematic reviews: an empirical study. *BMC Medical Research Methodology*, *6*, 7. doi: 10.1186/1471-2288-6-7

Sampson, M., & McGowan, J. (2006). Errors in search strategies were identified by type and frequency. *Journal of Clinical Epidemiology*, *59*(10), pp. 1057-1063.

Sampson, M., & McGowan, J. (2011). Inquisitio validus index medicus: A simple method of validating MEDLINE systematic review searches. *Research Synthesis Methods*, *2*(2), pp. 103-109.

Sampson, M., McGowan, J., Cogo, E., Grimshaw, J., Moher, D., & Lefebvre, C. (2009). An evidence-based practice guideline for the peer review of electronic search strategies. *Journal of Clinical Epidemiology*, *62*(9), pp. 944-52.

Sampson, M., McGowan, J., Cogo, E., & Horsley, T. (2006b). Managing database overlap in systematic reviews using Batch Citation Matcher: case studies using Scopus. *Journal of the Medical Library Association : JMLA*, *94*(4), pp. 461-463.

Sampson, M., McGowan, J., Tetzlaff, J., Cogo, E., & Moher D. (2008a). No consensus exists on search reporting methods for systematic reviews. *Journal of Clinical Epidemiology*, *61*(8), pp. 748-754.

Sampson, M., Shojania, K. G., McGowan, J., Daniel, R., Rader, T., Iansavichene, A. E., Ji, J., et al. (2008b). Surveillance search techniques identified the need to update systematic reviews. *Journal of Clinical Epidemiology*, *61*(8), pp. 755-762.

Sampson, M., Zhang, L., Morrison, A., Barrowman, N. J., Clifford, T. J., Platt, R. W., Klassen, T. P., et al. (2006c). An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Medical Research Methodology*, *6*, 33. doi: 10.1186/1471-2288-6-33

Saracevic, T. (2007a). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, *58*(13), pp. 1915-1933.

Saracevic, T. (2007b). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science*, *58*(13), pp. 2126-2144.

Sassi, F., Archard, L., & McDaid, D. (2002). Searching literature databases for health care economic evaluations: how systematic can we afford to be? *Medical Care, 4*(5), pp. 387-394.

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: Toward a dynamic, situational definition. *Information Processing & Management, 26*, 755-775. Quoted in Borlund, P. (2003). The Concept of Relevance in IR. *Journal of the American Society for Information Science and Technology, 54*(10), pp. 913-925.

Shaw, R. L., Booth, A., Sutton, A. J., Miller, T., Smith, J. a, Young, B., Jones, D. R., et al. (2004). Finding qualitative research: an evaluation of search strategies. *BMC Medical Research Methodology*, *4*, 5. doi: 10.1186/1471-2288-4-5

Shea, S., Weinstock, R. S., Starren, J., Teresi, J., Palmas, W., Field, L., Morin, P., et al. (2006). A randomized trial comparing telemedicine case management with usual care in older, ethnically diverse, medically underserved patients with diabetes mellitus. *Journal of the American Medical Informatics Association : JAMIA*, *13*(1), pp. 40-51.

Singhal A. (2001). Modern Information Retrieval: A Brief Overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 24*(4), pp. 35-43.

Smalheiser, N. R., & Torvik, V. I. (2009). Author name disambiguation. *Annual Review of Information Science and Technology*, *43*(1), pp. 1-43.

Sparck, J. K., Walker, S., & Robertson, S. E. (2000). A probabilistic model of information retrieval: development and comparative experiments. *Information Processing & Management*, *36*(6), pp. 779-808.

SPSS [Computer program]. Chicago, Illinois: SPSS Inc., 2010.

Stokes, P., Foster, A., & Urquhart, C. (2009). Beyond relevance and recall: testing new user-centred measures of database performance. *Health Information and Libraries Journal*, *26*(3), pp. 220-321.

Strandell, A., Lindhard, A., & Eckerlund, I. (2005). Cost-effectiveness analysis of salpingectomy prior to IVF, based on a randomized controlled trial. *Human Reproduction (Oxford, England)*, *20*(12), pp. 3284-3292.

Subirana, M., Sola´, I., Garcia, JM., Gich, I., & Urru´tia, G. (2005). A nursing qualitative systematic review required MEDLINE and CINAHL for study identification. *Journal of Clinical Epidemiology, 58*, pp. 20-25.

Sykes, J. (2001). *The Value of Indexing.* Factiva, Dow Jones and Reuters. Retrieved 24 May 2010 from http://www.factiva.com/infopro/indexingwhitepaper.pdf.

Tedd, L. A. (2006). Library management systems. In Bowman, J. H. (Ed.), *British Librarianship and Information Work 1991-2000* (pp.452-471). Aldershot: Ashgate.

Takeda, a, Loveman, E., Harris, P., Hartwell, D., & Welch, K. (2008). Time to full publication of studies of anti-cancer medicines for breast cancer and the potential for publication bias: a short systematic review. *Health Technology Assessment (Winchester, England)*, *12*(32), iii, ix-x, pp. 1-46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18831948

Taljaard, M., McGowan, J., Grimshaw, J. M., Brehaut, J. C., McRae, A., Eccles, M. P., & Donner, A. (2010). Electronic search strategies to identify reports of cluster randomized trials in MEDLINE: low precision will improve with adherence to reporting standards. *BMC Medical Research Methodology*, *10*, 15. doi:10.1186/1471-2288-10-15

Taylor, M. (2009). What is sensitivity analysis?. *What is…? Series*. Retrieved 25 June 2011 form http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_sens_analy.pdf

The CONSORT Statement. (2012). Retrieved from (http://www.consort-statement.org/consort-statement/).

The Vermont Oxford Network Steroid Study Group. (2001). Early postnatal dexamethasone therapy for the prevention of chronic lung disease. *Pediatrics 108*, pp. 741-748.

Thomson Reuters. (2012). Web of knowledge (WoK). Retrieved from http://apps.webofknowledge.com/.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*, 45. doi: 10.1186/1471-2288-8-45

Thomas, J., McNaught, J., & Ananiadou, S. (2011). Applications of text mining within systematic reviews. *Research Synthesis Methods, 2*(1), pp. 1-14. doi: 10.1002/jrsm.27.

Vakkari, P., & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, *56*(5), pp. 540-562.

Vermeire, E., Hearnshaw, H., Rätsep, A., Levasseur, G., Petek, D., van Dam, H., van der Horst, F., et al. (2007). Obstacles to adherence in living with type-2 diabetes: an international qualitative study using meta-ethnography (EUROBSTACLE). *Primary Care Diabetes*, *1*(1), pp. 25-33.

Wilczynski, N. L., Haynes, R. B., Lavis, J. N., Ramkissoonsingh, R., & Arnold-Oatley, A. E. (2004). Optimal search strategies for detecting health services research studies in MEDLINE. *CMAJ : Canadian Medical Association journal*, *171*(10), pp. 1179-1185.

Wilczynski, N. L., Marks, S., & Haynes, R. B. (2007). Search strategies for identifying qualitative studies in CINAHL. *Qualitative Health Research*, *17*(5), pp. 705-710.

Wilczynski, N.L., McKibbon, K.A., & Haynes, R. B. (2007). Response to Glanville et al.: how to identify randomized controlled trials in MEDLINE: ten years on'. *Journal of Medical Library Association, 95*(2), pp. 117-118.

Wilczynski, N. L., McKibbon, K. A., & Haynes, R. B. (2011). Search filter precision can be improved by NOTing out irrelevant content. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, pp. 1506-1513.

Wong, S. S. L., Wilczynski, N. L., & Haynes, R. B. (2006). Optimal CINAHL search strategies for identifying therapy studies and review articles. *Journal of Nursing Scholarship : An Official Publication of Sigma Theta Tau International Honor Society of Nursing / Sigma Theta Tau*, *38*(2), pp. 194-199.

Wong, S. S.-L., Wilczynski, N. L., Haynes, R. B., & Ramkissoonsingh, R. (2003). Developing optimal search strategies for detecting sound clinical prediction studies in MEDLINE. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, pp. 728-732.

World Health Organisation (WHO). (2012). International Clinical Trials Registry Platform (ICTRP). Retrieved from http://www.who.int/ictrp/search/en/.

Xu, Y. C., & Chen, Z. (2006). Relevance judgment: What do information users consider beyond topicality? *Journal of the American Society for Information Science and Technology*, *57*(7), pp. 961-973.

Yoshii, A., Plaut, D. a, McGraw, K. a, Anderson, M. J., & Wellik, K. E. (2009). Analysis of the reporting of search strategies in Cochrane systematic reviews. *Journal of the Medical Library Association : JMLA*, *97*(1), pp. 21-29.

Zhang, L., Ajiferuke, I., & Sampson, M. (2006). Optimizing search strategies to identify randomized controlled trials in MEDLINE. *BMC Medical Research Methodology*, *6*, 23. doi:10.1186/1471-2288-6-23

Zimmer, L. (2006). Qualitative meta-synthesis: a question of dialoguing with texts. *Journal of Advanced Nursing*, *53*(3), pp. 311-318.