Canadian Coordinating Office for Health Technology Assessment

The Pediatric Economic Database Evaluation (PEDE) Project

Wendy J. Ungar, MSc PhD[1]

Maria T. Santos, MHSc[2]

May 2002

1. Population Health Sciences, The Hospital for Sick Children Research Institute and the Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
2. Population Health Sciences Department, The Hospital for Sick Children Research Institute

REVIEWERS

These individuals kindly provided comments on this report.

CCOHTA Scientific Advisory Panel Reviewers

Mr. Doug Coyle
Principal Investigator
Clinical Epidemiology Unit
Loeb Research Institute
Ottawa Civic Hospital
Ottawa, Ontario

Dr. Jeff Scott
Provincial Medical Officer of Health
Province of Nova Scotia
Halifax, Nova Scotia

This report is a review of existing public literature, studies, materials and other information and documentation (collectively the "source documentation") which are available to the CCOHTA. The accuracy of the contents of the source documentation on which this report is based is not warranted, assured or represented in any way by CCOHTA and CCOHTA does not assume responsibility for the quality, propriety, inaccuracies, or the reasonableness of any statements, information or conclusions contained in the source documentation.

CCOHTA takes sole responsibility for the final form and content of this report. The statements and conclusions in this report are those of CCOHTA and not of its Panel members or reviewers.

Authorship

As Principal Investigator, Dr. W. Ungar conceived the study design, supervised the data collection and was responsible for citation adjudication and all decisions regarding data management. Dr. Ungar developed the analytic plan, supervised the statistical analysis, drafted the technical report and responded to reviewers' comments. Dr. Ungar was responsible for the final version of the report. Ms. M. Santos, Study Coordinator, conducted the data collection, contributed to the analytic plan, performed the statistical analysis and contributed to the draft and final reports. Ms. Santos commented on interpretation of all findings and was responsible for creating and maintaining the PEDE database. Ms. Santos approved the final version of the report.

Acknowledgements

This research was supported by funding from The Canadian Coordinating Office for Health Technology Assessment, The Hospital for Sick Children Research Institute, and from in-kind support from The Institute of Health Economics. The researchers are grateful to Ms. Amy Lee, The Hospital for Sick Children, and Ms. Tania Stafinski, The Institute of Health Economics, who conducted the quality appraisals. We thank Dr. Devidas Menon for his contribution to this research project and the CCOHTA reviewers for valuable feedback. Expert advice from the following panel members is gratefully acknowledged: Dr. Upton Allen, The Hospital for Sick Children, Dr. Ruth Collins-Nakai, University of Alberta Hospital, Dr. Cam Donaldson, University of Calgary, Dr. Tom Einarson, University of Toronto, Dr. David Feeny, Institute of

Health Economics and Dr. George Torrance, Innovus Research Inc. We thank Ms. Mahdie Seyed and Ms. Renee Henderson for their technical assistance and Ms. Elizabeth Uleryk for her expertise in Information Sciences. Thanks to Mr. Bruce Brady (CCOHTA) for managing this project as the CCOHTA liason.

EXECUTIVE SUMMARY

Introduction

Economic evaluations of health care interventions and services are becoming increasingly frequent to inform clinical and health policy decision-making. In recent years, standard methods have been recommended for conducting health economic studies, especially pharmaceutical evaluations. However, the applicability of standard methods to special populations such as children must be examined. Thus the Pediatric Economic Database Evaluation (PEDE) Project was conceived. The two research objectives were:

1. To analyze trends of published economic evaluations in pediatric health care over a 20-year period
2. To critically appraise the quality of published pediatric economic evaluations in order to identify gaps and areas for future methodological investigation.

Methods

With funding from The Hospital for Sick Children (HSC) Research Institute, a comprehensive pediatric health economic database of 787 publications for January 1, 1980 to December 31, 1999 was established. The database includes key characteristics for each citation, including year, target population, ICD-9-CM disease class, age group, experimental intervention, intervention category, health outcomes and analytical technique and is linked to a bibliographic database containing the full citation information and abstract.

The examination of trends in the literature was accomplished by performing various one-way frequency distributions and two-way cross-tabulations on variables within the database. To conduct the quality appraisal, a draft quality appraisal instrument was developed and subjected to review by a panel of experts. The instrument was pilot tested and the final quality appraisal questionnaire was completed for a random sample of 150 publications by two independent appraisers. Test-retest and inter-rater reliability of the quality appraisal instrument were assessed. Descriptive statistics, correlations and analysis of variance were used to describe the quality of the sample of publications.

Results

A total of 787 papers were published over the study period with an average annual growth rate of 22%. Over half of the papers (440) were published from 1995 to 1999. Whereas infectious diseases and congenital anomalies were most frequently studied in 1980-84, complications of pregnancy and perinatal conditions displayed increasing prominence over time. Infectious diseases commonly studied included hepatitis B, general vaccination strategies, Hemophilus influenzae type B, measles, and varicella. Common pregnancy and perinatal conditions were cardiac abnormalities, low birth weight, prematurity, respiratory conditions, Downs syndrome, and congenital hip dislocation. By 1995-99, infectious diseases, congenital anomalies and complications of pregnancy accounted for 47% of all publications and health prevention, health treatments and detection interventions accounted for 70% of all interventions. Cost-effectiveness analysis (CEA) was the most common technique used, accounting for a majority of evaluations in all time intervals. The proportion of studies using CEA increased by 23% while the proportion

of studies using cost benefit analysis (CBA) decreased from 31% in 1980-84 to just 12% in 1995-99. Cost utility analysis (CUA) remained the least common analytic technique at all times. A 57-item pediatric quality appraisal questionnaire was developed and demonstrated excellent test-retest reliability (ICC = 0.92, 95% CI 0.71-0.98) and very good inter-rater reliability (ICC = 0.75, 95% CI 0.66-0.81). The sampled papers demonstrated good quality for the following domains: Target Population, Economic Evaluation, Discounting, Conclusions and Comparators. The papers were of poor quality with respect to Conflict of Interest, Incremental Analysis, Sensitivity Analysis and Perspective. Over one third (38%) were rated as very good to excellent according to the global rating while 44% were good or fair. Only the Costs & Resource Use, Outcomes and Conclusions domains demonstrated improvement over time but the effect was slight. Analytic technique was a significant predictor of domain score for most of the study design-related domains. Except for Time Horizon and Perspective, studies designated as CUAs demonstrated the highest domain scores and were significantly higher than those of CEAs and most cost minimization analyses (CMA).

Discussion

The largest proportion of studies represented health prevention interventions rather than treatments of specific diseases. This is consistent with the understanding that the incidence of pediatric disease is rare and the prevention of future pediatric and adult illness is emphasized. The growth in evaluations of diagnostic and treatment strategies for congenital and perinatal conditions is expected to continue with the advent of expensive new technology to support early life and as research advances in genomics are translated into genetic testing and treatment.

Most outcome measures used in the pediatrics economic literature represented intermediate outcomes. Final outcomes such as mortality, life years gained and QALYs are particularly problematic for the pediatric population. Currently no published health economic models exist for measuring life years gained over long time horizons that include periods of maturation, development and rapid physiological change.

The quality appraisal demonstrated a disappointing performance. Incremental analysis and sensitivity analysis were often missing and the overall economic analysis was usually poorly done. Statements regarding funding sources and investigator independence were usually absent. These findings reiterate the importance of adherence to health economic guidelines that promulgate the conduct and publication of high quality economic evaluations. Furthermore, peer reviewers who critique manuscripts for publication in medical journals should become familiar with these guidelines and should insist on high quality manuscripts.

Conclusions

The PEDE Project represents the first attempt to catalogue, analyse and appraise economic evaluations performed in a pediatric population. The findings point to the importance of approaching pediatric health economic research with a child-centered framework. Although this research revealed dramatic growth in the volume of publications over time, there remains much room for improvement in the quality of pediatric economic evaluations. Opportunities for future research relate to the development of statistical models to capture health benefits over time horizons that include development and maturation. The PEDE database, containing 787 citations,

will be of value to decision-makers as well as researchers undertaking methods research, a systematic review or meta-analysis of a particular intervention or condition. The conceptual knowledge derived from this project will foster research initiatives that will enhance the quality of economic assessment in children.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ABBREVIATIONS

ANOVA	Analysis Of Variance
CBA	Cost Benefit Analysis
CEA	Cost Effectiveness Analysis
CCOHTA	Canadian Coordinating Office of Health Technology Assessment
CMA	Cost Minimization Analysis
CUA	Cost Utility Analysis
EBM	Evidence-Based Medicine
DARE	Database of Abstracts of Reviews of Effectiveness
HSC	Hospital for Sick Children
HEED	Health Economic Evaluations Database
HTA	Health Technology Assessment
IPA	International Pharmaceutical Abstracts
NHS EED	National Health Services Economic Evaluation Database
QALY	Quality Adjusted Life Year
PEDE	Pediatric Economic Database Evaluation
RCT	Randomized Controlled Trials

1 INTRODUCTION

Economic evaluations of health care interventions and services are becoming increasingly frequent to inform clinical and health policy decision-making.[1-4] It is vital to ensure that health economic evaluations used for allocating health care have a strong methodological foundation and are of the highest quality. In recent years, standard methods have been recommended for conducting health economic studies, especially pharmaceutical evaluations.[5,6] However, the applicability of standard methods to special populations must be examined. In particular, child health differs from adult health in important ways. These differences include the vulnerability of children as they grow and develop, their reliance on parents, teachers and others to provide access to health care, the different ways they manifest disease and the ways in which they interact with the health care system and their immediate environment.[7-11] Understanding these aspects of child health is an important prerequisite for developing appropriate tools to assess the economic benefit of health interventions, programs and services directed at children.[12,13]

The gaps in knowledge and methodology that are unique to pediatric economic evaluation can only be revealed by a close examination of the literature. Thus the Pediatric Economic Database Evaluation (PEDE) Project was conceived.

1.1 Study Objectives

The purpose of the Pediatric Economic Database Evaluation (PEDE) Project was to identify pediatric economic evaluations in the literature and to summarize the trends and development of this field of research in the pediatric population. The two research objectives were:

1. To analyze trends of published economic evaluations in pediatric health care over a 20-year period
2. To critically appraise the quality of published pediatric economic evaluations in order to identify gaps and areas for future methodological investigation.

This investigation will provide some insight into the disease categories, target populations, age groups, types of interventions, and types of economic analysis that have characterized pediatric economic evaluations for the past twenty years. The ultimate purpose of the project is to identify areas requiring methodological improvement to establish the direction for a program of research in the conduct of economic evaluations in a pediatric population.

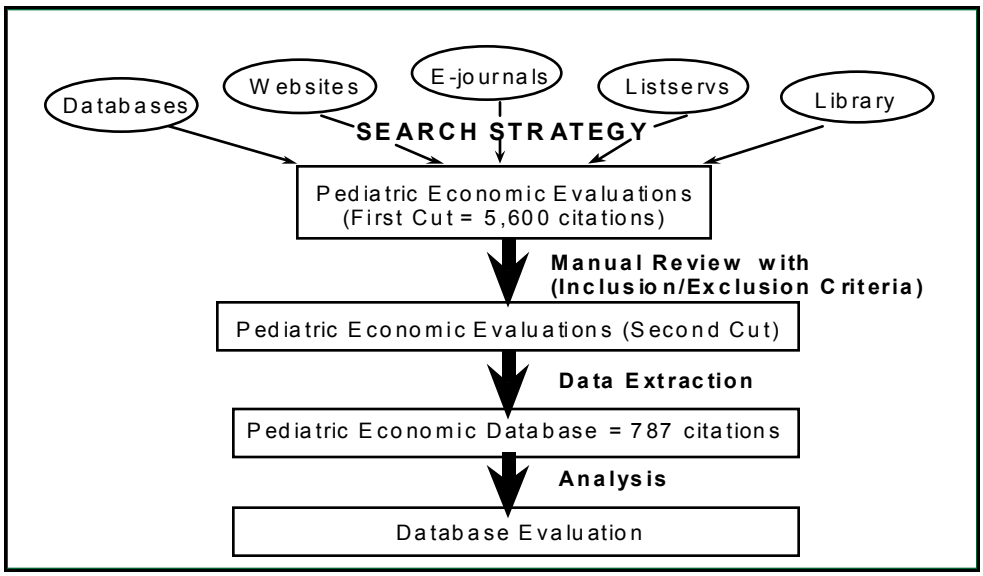
1.2 Background

With start-up funding provided by the HSC Research Institute, an initial phase of the PEDE project was completed in July 2000. This work consisted of establishing a comprehensive pediatric economic evaluation database of 787 publications for the 20-year period January 1, 1980 to December 31, 1999.[14] The completed steps are illustrated in Figure 1.1 and included:

1. Development, testing and validation of a highly sensitive search strategy to retrieve citations of pediatric economic evaluations from various sources

2. Manual review of 5,600 citations generated by the multi-database search strategy and application of the inclusion/exclusion criteria
3. Data extraction from eligible citations
4. Creation of the pediatric economic database
5. Inter-rater and test-retest reliability assessment of the citation selection process

Figure 1.1: Overview of Database Development



Our goal was to establish a database that was fully inclusive. To achieve this, a comprehensive search of diverse sources was undertaken. The sources for the database citations included electronic journal databases such as MEDLINE®, EMBASE® and NHS EED, web sites and other electronic and hard copy sources.

The details of the methods employed in steps 1-5 are described in a Technical Report[15] available from the authors and are summarized in Appendix 1. Additional information on Internet websites for grey literature, journals included and the categorized target populations are included in Appendices 2 through 4. With these preliminary steps completed, the database was ready for addressing the specific research questions outlined above.

2 LITERATURE REVIEW OF QUALITY APPRAISALS OF ECONOMIC EVALUATIONS

2.1 Literature Review

Several investigators have examined the scope or quality of economic evaluations in the medical literature.[16-31] The first systematic review of economic evaluations was published in 1980.[16] The investigators listed over 500 papers published from 1966 to 1978, demonstrating a rapid increase in publication rate over time. The growth in publications was greater in medical than non-medical journals with a preference for CEA over CBA. Concern was expressed regarding the quality implications of the rapid growth rate. In the first quality appraisal of economic evaluations, 47 articles published between January 1982 and November 1987 were compared on seven criteria.[17] Most CBAs and CEAs did not include comprehensive costs and sensitivity analysis was usually omitted from CBAs. Adams et al. appraised the quality of economic evaluations performed alongside randomized controlled trials (RCTs) published between 1966 and 1988.[18] A random sample of 51 papers demonstrated an average quality score of 0.32 and an average completeness score of 0.52 on a scale of zero to one. Studies published in later years exhibited higher scores. The most common deficiencies were improper allocation of overhead costs, absence of sensitivity analysis and lack of aggregation of treatment costs and consequences. Udvarhelyi et al. examined the quality of 27 economic evaluations published in 1978-80 and 1985-87 in general medicine, general surgery and medical sub-specialty journals.[19] An assessment of the presence of six fundamental principles revealed a performance rated as 'fair' by the investigators. The most common problem was failure to state and test assumptions with sensitivity analysis. Articles in general medicine journals were of higher quality than those appearing in general surgery or medical sub-specialty journals. No significant improvement was noted over time. The quality of 65 pharmacoeconomic studies published in six pharmacy journals between January 1985 and December 1990 was assessed according to 10 methodological criteria.[20] In 55% of the articles, cost saving was misinterpreted as cost-effectiveness. Only three of the 10 criteria were fulfilled by the majority of studies. The major limitations were lack of identification of relevant costs and consequences, lack of discounting, failure to perform an incremental analysis and failure to conduct sensitivity analysis. Many studies used the term "cost-effective" inappropriately. In a quality appraisal of 51 CUAs reported between 1980 to 1991, 46 studies were considered appropriate, with 16 rated as technically average, 13 above average and 17 below average.[21] Sacristan et al. reviewed the quality of selected published pharmacoeconomic studies for the purpose of constructing a quality appraisal checklist.[22] The authors concluded that while the number of pharmacoeconomic studies has increased progressively, there has not been a parallel increase in study quality.

The first bibliography of economic evaluations was reported in 1993, comprised of studies published from 1979 to 1990.[23] The list included 3,206 papers divided into studies and other publications, including reviews, descriptions of methodology, letters, and editorials. All publications were classified into 250 clinical subject categories. Quality of the publications was not reviewed. This health economic bibliography was later updated to include publications from 1991 to 1996.[24] The search resulted in 3,539 health economic publications which were divided into the same categories described above. Studies were further classified by study type, publication type and medical function.

In a study comparing the quality of 90 economic studies published in selected pharmacy, medical, and health economics journals from 1989-1993, health economics journals scored highest on a 13-item checklist, followed by medical journals, with pharmacy journals significantly lower.[25] Items relating to ethical issues and study perspective were the most common deficiencies. A study evaluating the quality of 44 CEAs and CBAs reported in the radiology literature from 1989 to 1995 found that adherence to methodologic standards was not optimal and quality did not improve over the study period.[26] A quality assessment of 54 economic evaluations of pharmaceuticals published in *PharmacoEconomics* between 1992-95 revealed a mean quality score of 3.0 (maximum 4.0).[27] CUA studies were of highest quality, followed by CBA, CEA and CMA. Some of the same investigators examined the quality of 51 pharmacoeconomic abstracts published in select journals between 1990-94.[28] Quality scores improved slightly between 1990 and 1994, with medical journals scoring highest, followed by pharmacy journals and health economic journals. Trends in economic evaluation were assessed in a study of the UK Health Economic Evaluations Database (HEED), comparing studies published between 1992 and 1996.[29] While the study did not examine quality according to established criteria, it did find that the proportion of RCTs, considered a high quality study design, did not increase over the study period, nor did discounting and sensitivity analysis.

Some recent quality appraisals have concentrated on particular aspects of health economic evaluation. Focusing on the completeness of uncertainty analysis, Briggs and Sculpher examined 93 papers listed in MEDLINE® in 1992 and observed that only 25% assessed uncertainty.[30] A recent review focused exclusively on health outcome measures used in economic evaluations published from 1986 to 1996.[31] In a sample of 455 studies included in HEED, major differences in the choice of health outcome measures were observed across disease categories. There was no evidence suggesting that the outcome measures most often recommended in guidelines, the quality-adjusted life year (QALY) and life-years gained, have become more common. As a reflection of the increasing trend to conduct systematic reviews and meta-analyses, the Consensus on Health Economic Criterialist (CHEC) Project was launched in 2000 to develop a checklist for use with systematic reviews of observational and experimental health economic clinical studies. To date, a checklist has not been released.[32]

The quality appraisal studies discussed above used a variety of appraisal methods and techniques. While there was some consistency among them with respect to the elements included, most of the appraisal instruments remained subjective with little information provided regarding validity and reliability. Nevertheless, these previously conducted appraisals reflect an ongoing concern regarding the quality of economic evaluations. The implication of poorly conducted health economic studies is that incomplete or unreliable information is used by decision-makers resulting in less than optimal resource allocation. The field of health economics continues to grow, and in an era of increasing financial constraint in the health care system, the need for high quality investigations becomes more acute. Most of the quality appraisals conducted to date reveal significant deficiencies in the use of health economic methodology. Frequent problems were insufficient cost itemization, lack of testing of assumptions through the use of sensitivity analysis, failure to specify incremental costs and consequences and abuse of the term 'cost-effective' as an imprecise adjective rather than as an analytic technique. While many of the studies described above suggest that little improvement has occurred over time, these

studies were usually limited to target journals or short time periods. In the one study conducted over a time period exceeding 12 years, the authors did observe an increase in quality over time.[18] There remains a need for a comprehensive appraisal of the literature over a sufficiently long time interval to determine if health economic research is improving and thus becoming of better value to decision makers. Most importantly, none of the previous studies focused on the pediatric population and the particular methodologic challenges of conducting health economic evaluations in children. The need persists for an investigation of the 'state of the art' of health economic evaluation in children.

2.2 Comparison of PEDE Project to other Health Economic Database Initiatives

While there are a number of health economic databases that provide researchers with the ability to search for specific types of economic evaluations, they are hampered by limitations that are overcome in the PEDE Database. An on-line electronic database of the bibliography of economic evaluations published by Elixhauser and Luce in 1993 was created by the US Centre for Disease Control and Prevention.[33] The WONDER database included 3,206 economic evaluations published from 1979 through 1990.[23] The database was recently retired due to Y2K incompatibility. The University of York maintains three databases: the National Health Services Economic Evaluation Database (NHS EED),[34] the Database of Abstracts of Reviews of Effectiveness (DARE),[35] and the Health Technology Assessment (HTA) database.[36] Unlike DARE and HTA, the NHS EED consists exclusively of economic evaluations and starting in 2000, became part of the Cochrane Library.[37] The HEED is a joint initiative between the UK Office of Health Economics and the International Federation of Pharmaceutical Manufacturers' Associations. HEED contains more than 19,700 citations, with 50% reviewed by academic researchers using a structured format.[38]

The database generated for the PEDE Project is distinct from the NHS EED and the HEED databases in several significant respects. First, a number of on-line databases (HealthSTAR, CANCERLIT®, IPA, Mental Health Abstracts and PubMed [PreMEDLINE]) were incorporated into the PEDE Project literature search that are not included in the HEED or NHS EED search procedures. Second, with regard to the period of coverage, the NHS EED database references literature published after 1994. Although, the HEED database maintains some publications dating back to 1967, the main emphasis of their database dates from 1992 onward. Currently, the PEDE project includes pediatric economic evaluations from 1980 to 1999. This twenty-year time frame permits the analysis of trends in the literature over time. The third distinct feature of the PEDE Project database involves the search strategy used for retrieving literature. The PEDE Project is limited to the pediatric population. Because the search was narrowed in this way, an expansive search strategy could be used to identify economic evaluations and 'explode' more relevant terms than would otherwise be possible if the whole population was included. Finally, it has been determined that NHS EED alone is insufficient for identifying all relevant pediatric economic evaluations. The PEDE Project database contains 787 full pediatric economic evaluations published between 1980 to 1999. While, the NHS EED contains over 600 publications related to pediatrics, these citations include costing papers, methodological studies, and reviews in addition to full economic evaluations. Due to its greater comprehensiveness over a 20-year interval, the PEDE database demonstrates notable superiority over existing databases for the conduct of pediatric citation research.

3 METHODS

3.1 PEDE Database

The creation of the PEDE project database is summarized in Appendix 1. Data entry and data management of eligible publications were performed using Microsoft ACCESS 97. Data were extracted and entered into the PEDE database from the 787 eligible citations. The following data fields were used for data entry:

- Notes
- Citation identification number
- Journal
- Target population
- Population category
- ICD-9-CM disease class[39]
- Age group
- Experimental intervention
- Intervention category
- Primary health outcome
- Summarized outcome category
- Analytical technique

Detailed definitions of each variable can be found in Appendix 1. Data from the above fields were entered into an MS ACCESS 97 database for each citation. In addition to the data contained in the ACCESS database, the citation identification number was used to link the record to its full citation, including MeSH terms and abstract, in EndNote. To maintain consistency throughout the database, decision rules were followed for each data field. The PEDE database is thus a linked MS ACCESS – EndNote electronic database of 787 citations.

3.2 Pediatric Economic Evaluation Trend Analysis

The first research objective of the PEDE Project was to examine trends in the pediatric economics literature from 1980 to 1999. This was accomplished by performing various one-way frequency distributions and two-way cross-tabulations on the variables in the database. The results were summarized in descriptive tables.

3.3 Quality Appraisal of the Pediatric Health Economic Literature

The second research objective of the PEDE project was to conduct a quality appraisal of pediatric health economic evaluations published from 1980 to 1999. To conduct the quality appraisal, a quality appraisal instrument was developed that could be used to extract data from publications and score the quality of each publication. A draft instrument was subjected to an external review by a panel of experts. The panel of experts also provided feedback on proposed scoring methods. Following expert review, the instrument was pilot tested in 10 publications by three independent experienced appraisers. The ratings were compared. Discrepancies were addressed and a comprehensive set of decision rules was developed to accompany the final instrument. The final quality appraisal questionnaire was completed for a random sample of

publications by two of the independent experienced appraisers who participated in the pilot test. Test-retest and inter-rater reliability of the quality appraisal instrument were assessed. The following sections describe each of the above steps in detail.

3.3.1 Development of the quality appraisal instrument

The initial quality appraisal instrument was developed through a process of item selection, domain construction, assignment of response options and development of a scoring scheme.

Based on the various checklists in the literature that evaluated the quality of health economic evaluations,[17-22,25-28,40-45] a comprehensive questionnaire was constructed. The types of literature that were appraised using these checklists encompassed research published in specific pharmacy, medical, or health economics journals[20,22,25-28] or economic studies employing a particular analytic technique (e.g. CBAs, CEAs, or CUAs).[17,19,21,43,45] In addition to checklists found in peer-reviewed journals, those available from CCOHTA and relevant textbooks[40,41] were incorporated. Lastly, questions applicable specifically to the pediatric population were newly formulated since available checklists lacked items that address pediatric and parental outcomes, parental indirect costs, a maturation time horizon, the unit of analysis, and other concerns related to pediatric economic evaluations. Thus the draft quality appraisal instrument was based on currently recommended checklists that have been modified to incorporate outcomes, costing, measurement and analysis issues pertaining to the pediatric population. With the understanding that many issues in economic evaluation methodology remain controversial, a deliberately comprehensive approach was taken to include all items of possible relevance.

The questionnaire domains were conceived based on a careful review of established checklists.[17-20,22,25-28,40,41,43-45] The questions were grouped into the following domains: Economic Evaluation, Comparators, Target Population, Time Horizon, Perspective, Costs and Resource Use, Outcomes, Quality of Life, Analysis, Discounting, Incremental Analysis, Sensitivity Analysis, Conflict of Interest, and Conclusions.

Within the questionnaire, all the responses were categorical. Most of the categorical items had response options that followed the following format: yes (explicitly stated); yes (inferred from text, tables, or figures); no; unknown/not stated/can't tell; or not applicable. The distinction between explicitly and implicitly stated responses was based on checklists from the literature that used similar approaches.[22,25,27,28,43] Also included were several questions that were purely descriptive and were not used in calculating the quality score. These items added information about methodology that could further characterize the literature and explain the observed results.

3.3.2 External review

An expert panel was assembled to assess the face and content validity of the draft questionnaire and the proposed scoring approach, i.e. to ensure that all relevant items were included and that the choice of items and scoring technique were sensible and credible (Appendix 5). The reviewers were chosen to represent expertise in health economic methods, quality appraisal of health economic publications and pediatric research. The external review was comprised of three stages: 1) the experts were asked to comment on the wording of the questions and recommend the deletion or addition of items, 2) the experts were asked to rate each question as having high,

medium, or low importance for inclusion in the questionnaire and 3) the experts were asked to comment on the various options for scoring the questionnaire and/or provide suggestions for alternative scoring schemes.

The experts were asked to rate 48 quantitative items in the questionnaire as having high, medium or low importance for inclusion. The distribution of ratings for each item was determined. To evaluate how well the reviewers' responses correlated with one another, the frequency for each possible pattern of response distribution was calculated. "High" agreement among experts was specified if five or more experts agreed on the rating. If four experts agreed on the rating, a "medium" level of agreement was assigned. If agreement was found between three experts or fewer, the level of agreement was designated as "low". Table 3.1 illustrates that there was high concordance of responses for only 42% of the questionnaire items, medium agreement for 31% of the items and low agreement for 27% of the 48 items for which seven reviewers provided a rating.

Table 3.1: Assessment of agreement among expert reviewers

Potential response distribution patterns			Number of items sharing the same response distribution	Concordance
0	0	7	2	High *(5 or more experts agree on the rating)*
0	1	6	10	
0	2	5	6	
1	1	5	2	
			n = 20 items, frequency = 42%	
0	3	4	5	Medium *(4 experts agree On the rating)*
1	2	4	10	
			n =15 items, frequency = 31%	
1	3	3	4	Low *(3 experts agree On the rating)*
2	2	3	9	
			n = 13 items, frequency = 27%	

All seven experts provided a rating for n=48 items

The agreement between the experts' rating of highly important items and published economic evaluation checklists was evaluated by calculating the per cent observed agreement and a kappa coefficient. A rating was classified as high if five or more experts rated the item as having **high** importance and if the item appeared in four or more published checklists.

Table 3.2: Agreement between the experts and the literature

Item has high importance:

		Experts		
		Yes	No	Total
Literature	Yes	11	4	15
	No	5	28	34
	Total	16	32	48

Observed agreement = 0.81
Chance-expected agreement = 0.56
Kappa = 0.57
95% Confidence Interval = (0.32, 0.82)

The level of agreement between the experts and the literature in rating the items as having high importance was moderate (kappa value of 0.57). There was much variation in opinion among the experts regarding which items were important to include. There were 28 items that four or fewer experts rated as important for inclusion and that also appeared in fewer than four published checklists. As these items represent contentious issues rather than items for which there was consensus for exclusion, almost all of these items were retained. We took a comprehensive approach and retained items considered important by the experts or by the literature. Some of the items rated as important by the experts but not by the literature pertained to the pediatric population and were therefore absent from published checklists. There were no items that five or more experts rated as having low importance.

3.3.3 Pilot testing of the questionnaire

Based on the comments provided by the external reviewers, modifications were made to construct the pilot questionnaire. During the pilot test, ten articles representative of the various analytic techniques (i.e. CEA, CBA, CUA, and CMA) and the five-year periods (i.e. 1980-84, 1985-89, 1990-94, and 1995-99) were appraised. The pilot version of the questionnaire had 64 items. Pilot-testing a questionnaire with maximum comprehensiveness enabled us to subsequently delete items that were found to be redundant, invalid or difficult to interpret. Three independent appraisers with experience in health economic evaluation and critical appraisal reviewed each article (AL, TS and WU). The scores for each question were compared among the three appraisers. Table 3.5 illustrates the number of questions for each article in which all three, two out of three, or no appraisers agreed on the response. The appraisers had a high tendency to agree. Following the pilot test, discussions were held to address discrepancies and review the evaluation process of each appraiser. Based on these discussions, detailed decision rules for completing the questionnaire were created. These decision rules were incorporated into the final questionnaire. These decision rules were developed to ensure consistent application of the economic principles by the two independent appraisers conducting the quality appraisal. The final questionnaire and decision rules are presented in Appendix 6.

Table 3.3: Pilot test results

Paper ID#	Author	Year	No. questions with 3/3 agreed		No. questions with 2/3 agreed		No. questions with 0/3 agreed	
370	Knight et al.	1982	30	47%	28	44%	6	9%
351	Karam et al.	1986	32	50%	28	44%	4	6%
440	Marchese et al.	1986	28	44%	32	50%	4	6%
529	Phaosavasdi et al.	1987	29	45%	27	42%	8	13%
7	Aggarwal et al.	1994	28	44%	31	48%	5	8%
243	Glotzer et al.	1994	30	47%	28	44%	6	9%
519	Paul et al.	1995	36	56%	22	34%	6	9%
314	Horton et al.	1996	35	55%	25	39%	4	6%
499	Oelberg et al.	1998	30	47%	31	48%	3	5%
713	Tsevat et al.	1999	32	50%	28	44%	4	6%

3.3.4 Questionnaire scoring

There was a consensus among the members of the expert panel that a linear additive total score would not be useful. As all of the study domains represented important components of an economic evaluation, it would be inappropriate to allow a high sub-score for a particular domain to compensate for a low sub-score on another domain. It was also agreed that weighting the domains equally to calculate a total score would be difficult to justify. The experts agreed that domain sub-scores would be of value. The 57 items in the final questionnaire (Appendix 6) were mapped onto 14 domains as illustrated in Table 3.4.

Table 3.4: Mapping of questionnaire items onto 14 domains

Domain	Question #
Economic Evaluation	1 to 3
Comparators	4 to 8
Target Population	9 and 10
Time Horizon	11 and 12
Perspective	13 to 15
Costs and Resource Use	16 to 21
Outcomes	22 to 29
Quality of Life	30 to 32
Analysis	33 to 42
Discounting	43 and 44
Incremental Analysis	45 to 47
Sensitivity Analysis	48 to 51
Conflict of Interest	52 and 53
Conclusions	54 to 56
Global Assessment	57

The number of items in a domain varied from 2 to 10 with an average of 3.8. The questionnaire was designed to maximize the number of scorable items. Among the 57 items in the final questionnaire, 46 have response options that are scorable as follows:

Table 3.5: Response options for scorable items

Response Option	Score
Yes (explicitly stated)	1.0
Yes (inferred from text, tables, or figures)	0.5
No	0
Unknown/Not stated/Can't tell	0
Not applicable	Not included

A score for each domain was calculated by taking the mean score of all the scorable items in that domain. Thus each domain score had a range from zero to one. If an item was 'not applicable' for that publication, then that item was not included in the domain score calculation. The questionnaire does not have a total score, as a single summary score would mask high and low scores for individual domains.

In the final version of the questionnaire, in addition to the 46 scorable items that mapped onto 14 domains, there were 10 items that added descriptive information about methodology to further characterize the literature and explain the observed results. In addition, a global quality assessment was included as the final questionnaire item. This yielded a total of 57 items.

3.3.5 Quality appraisal

For the quality appraisal, a 20% random sample of articles, stratified by five-year period, was selected from the PEDE database for appraisal. A total of 150 articles, in which the journal name, author and year were concealed, were distributed to two independent appraisers, one at the Hospital for Sick Children (AL) and the other at the Institute for Health Economics (TS). A database was developed using ACCESS 97 to input the data from the completed questionnaires. The final quality appraisal instrument underwent both inter-rater and test-retest reliability assessment.

a) Inter-rater reliability

Inter-rater reliability is an indication of the consistency in responses between multiple users of the quality appraisal questionnaire. For the inter-rater reliability assessment, AL completed appraisals of 150 papers, while TS completed appraisals for 149 papers. The results are presented for the 149 papers appraised by both reviewers. Inter-rater reliability was assessed by calculating an intra-class correlation coefficient for each domain. These analyses permitted the detection of systematic differences between the raters. Ideally it would be optimal to calculate intra-class correlations for each item because of the possibility that item scores within a domain may oppose each other, thus biasing the domain coefficient. However, the current approach was chosen because calculating 57 intra-class correlation coefficients would present unnecessary complexity and would lead to difficulty in interpretation. Although a total questionnaire score was not recommended for the formal quality appraisal analysis, a total questionnaire score was used as a

convenient quantitative summary measure for assessing inter-rater reliability. To calculate the total questionnaire score, the scores from each of the 46 scorable items were summed together and then divided by the total number of applicable items among the 46 scorable items.

Each intra-class correlation coefficient was calculated from a repeated measures Analysis of Variance (ANOVA) in SPSS[46] using a two-way mixed model and a consistency type of index. Ninety-five percent confidence intervals were computed for the intra-class coefficient.

b) Test-retest reliability

The purpose of the test-retest is to measure the consistency of responses within a rater. In this case, AL re-appraised ten papers from the initial batch approximately two months later. Similar to testing the inter-rater reliability, an intra-class coefficient was calculated using the repeated measures ANOVA. The test-retest reliability was evaluated for the total questionnaire score and the domain scores.

c) Preparing the final dataset for analysis

Since two datasets existed comprised of the appraisals for each of the two raters, it was necessary to create a final dataset to conduct the quality appraisal analysis. To achieve this, a third reviewer (WU) examined papers with major discrepancies between the two raters. Publications were selected in which the absolute difference in total score between AL and TS was greater than or equal to eight and the difference of at least one key domain score was greater than or equal to 0.6. This represented 17% variation in scores and was observed in 26/150 (17%) of publications. Within each of the 26 selected papers, individual questions within key domains were reviewed to establish where the discrepancies existed. WU selected the response to be used in the final dataset. The final dataset integrates TS' values and any final responses adjudicated by WU.

d) Quality appraisal analysis

A series of statistical tests were applied to the final dataset using SAS version 8.02.[47] Descriptive statistics for each domain score, including mean, standard deviation, minimum, maximum, median and inter-quartile range were calculated. Further analyses were undertaken on seven key domains considered to be highly relevant and most often discussed in published quality appraisals: Comparators, Perspective, Costs and Resource Use, Outcomes, Analysis, Incremental Analysis and Sensitivity Analysis. The non-parametric Spearman correlation coefficient, representing the degree of association between two variables, was calculated between all the individual domain scores and between key domain scores and five-year period.

One-way frequency tests were performed on the global impression score (questionnaire item # 57). In addition, the Spearman correlation coefficient was determined between the global score and five-year period and between the global score and key domain scores. ANOVA was performed to determine whether variances in the key domain scores could be explained by the analytic technique. Duncan's multiple-range test for post hoc comparisons was used to assess the statistical significance of differences between pairs of mean domain scores. Finally, one-way frequency distributions were performed on the remaining descriptive items that were not scored within the domains. A p value < 0.05 was considered statistically significant for all inferential tests.

4 RESULTS

In accordance with the research objectives, the results of the PEDE project analysis are divided into sections related to the analysis of trends in pediatric economic evaluation between 1980 to 1999 and the analysis of the quality of a random sample of the publications over this period.

4.1 Trend Analysis of PEDE Database

This analysis examined trends over time with respect to the volume of publications, the disease category, the type of intervention, the age groups studied, the health outcomes measured, the analytic technique used and the journal type. Results of each analysis are presented below.

4.1.1 Volume of publications

There was significant growth in the number of published pediatric economic evaluations over the study period. The annual number grew ten-fold, from eight publications in 1980 to 82 in 1999 (Table 4.1, Figure 4.1). A total of 787 papers were published over the study period, indicative of an average annual growth rate of 22%. As indicated in Table 4.2, over half of the papers (440) were published from 1995 to 1999. The volume of publications roughly doubled every five years.

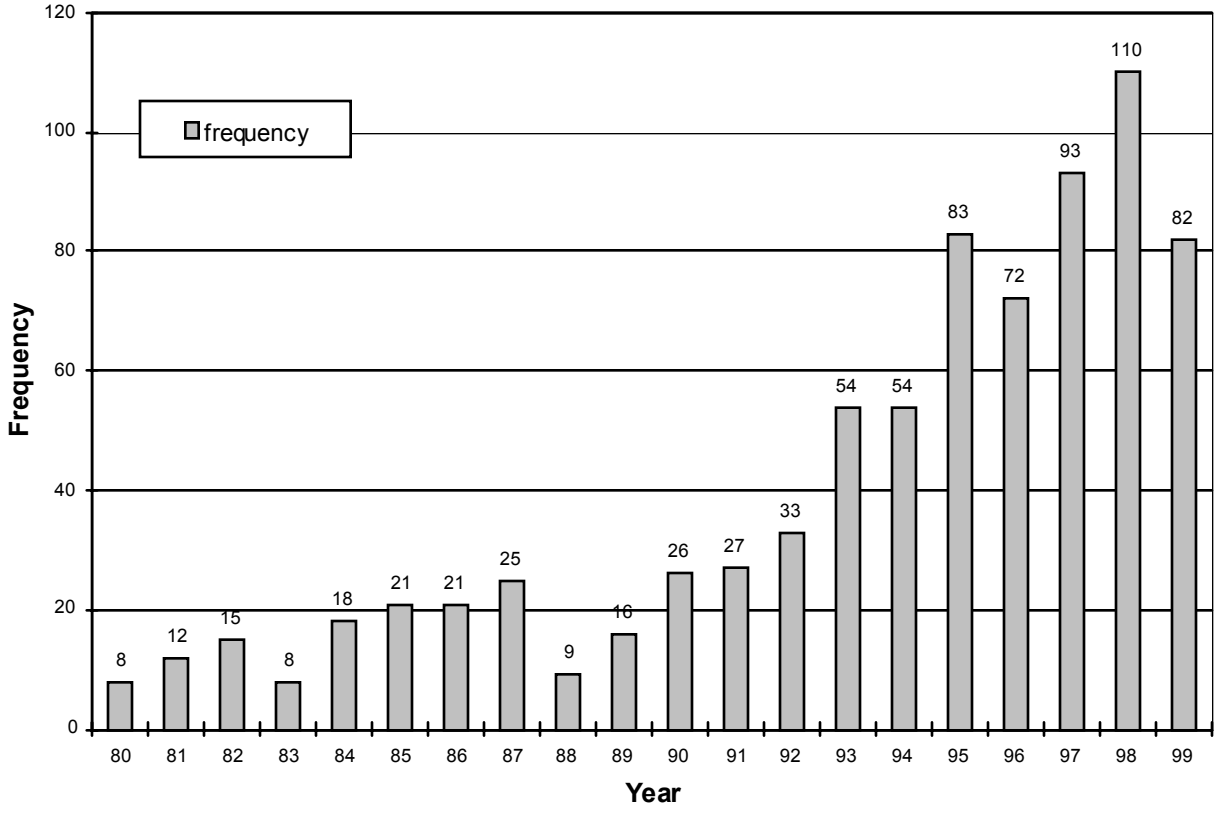
Table 4.1: Publications per year

Year	n	%	% change
80	8	1%	
81	12	2%	50%
82	15	2%	25%
83	8	1%	-47%
84	18	2%	125%
85	21	3%	17%
86	21	3%	0%
87	25	3%	19%
88	9	1%	-64%
89	16	2%	78%
90	26	3%	63%
91	27	3%	4%
92	33	4%	22%
93	54	7%	64%
94	54	7%	0%
95	83	11%	54%
96	72	9%	-13%
97	93	12%	29%
98	110	14%	18%
99	82	10%	-25%
Total	787	100%	--
Average annual growth			22%

Table 4.2: Publications per five-year period

Five-year period	n	%	% change
1980-84	61	8%	0%
1985-89	92	12%	51%
1990-94	194	25%	111%
1995-99	440	56%	127%
Total	787	100%	--
Average five-year growth			96%

Figure 4.1: Volume of publications over time



4.1.2 Disease category

The various conditions studied in the database were mapped onto ICD-9-CM disease categories. Whereas infective and parasitic diseases and congenital anomalies were most frequently studied in 1980-84, complications of pregnancy and perinatal conditions displayed increasing prominence over time. Specific conditions commonly studied in the infective and parasitic disease category include hepatitis B, general vaccination strategies, Hemophilus influenzae type B, measles, and varicella. Common pregnancy and perinatal conditions that are treated or screened for include cardiac abnormalities, low birth weight, prematurity, respiratory conditions, Down syndrome, and congenital hip dislocation. By 1995-99, infective and parasitic diseases, congenital anomalies, and complications of pregnancy, childbirth, and the puerperium accounted for 47% of all publications. More than half the economic evaluations in almost all disease categories were published between 1994-99.

4.1.3 Intervention type

As seen in Table 4.3, health prevention interventions accounted for one third of publications in 1980-84. By 1995-99, health prevention, health treatments and detection interventions accounted for 70% of all publications.

Table 4.3: Publications per intervention type by five-year period

Intervention Type	1980-84		1985-89		1990-94		1995-99		Total	
	n	%	n	%	n	%	n	%	n	%
Health prevention	20	32%	26	25%	53	25%	126	27%	225	26%
Health treatment	9	14%	16	16%	63	29%	129	27%	217	25%
Detection	9	14%	20	20%	42	20%	74	16%	145	17%
Health program	6	10%	21	21%	19	9%	45	9%	91	11%
Health care delivery	3	5%	6	6%	12	6%	36	8%	57	7%
Surgical	4	6%	1	1%	8	4%	33	7%	46	5%
Dental	6	10%	8	8%	12	6%	8	2%	34	4%
Diagnosis	2	3%	3	3%	4	2%	17	4%	26	3%
Educational	4	6%	1	1%	2	1%	7	1%	14	2%
Total	63	100%	102	100%	215	100%	475	100%	855	100%

Note: The total does not reflect the total number of records (787) since some of the publications included multiple intervention types.

A cross-tabulation of intervention type across disease category indicates that a majority (63%) of health prevention interventions was for infectious disease. Economic evaluations of interventions to prevent complications of pregnancy were also common. The largest proportions of health treatments were for infectious disease and perinatal conditions. Detection/diagnostic interventions were focused mostly on congenital anomalies, complications of pregnancy, and infectious diseases. Studies of health programs were devoted to general health as well as perinatal conditions and complications of pregnancy. Over a third (35%) of assessments of surgical interventions were directed at congenital anomalies and were also common for digestive and genitourinary conditions.

4.1.4 Age group

Shifting trends in the age groups studied were visible over the study period (Table 4.4). Whereas studies of perinates and children under 12 years declined between 1980 to 1999, the proportion of economic assessments studying infants and adolescents increased. Nevertheless, by 1995-99, studies of children accounted for more than one third (36%) of all pediatric economic evaluations.

Table 4.4: Publications per age group by five-year period

Age Category	1980-84		1985-89		1990-94		1995-99		Total	
	n	%	n	%	n	%	n	%	n	%
Perinate (antenatal period, up to 7 days of life)	14	17%	24	19%	43	15%	80	11%	161	13%
Neonate (new born to ≤ 1 month)	8	10%	21	16%	46	16%	82	11%	157	13%
Infant (1 month to ≤ 1 year)	13	16%	20	16%	55	20%	147	21%	235	19%
Child (> 1 year to 12 years)	34	41%	42	33%	95	34%	261	36%	432	36%
Adolescent (13 to 18 years)	12	14%	18	14%	37	13%	123	17%	190	16%
Adult (19 and older)*	2	2%	3	2%	6	2%	24	3%	35	3%
Total	83	100%	128	100%	282	100%	717	100%	1,210	100%

Note: The total does not reflect the total number of records (787) since several publications included multiple age groups.

* Publications that involved adults and included a separate analysis for the pediatric population.

A cross-tabulation of intervention type by age group demonstrated that 57% of health prevention interventions were directed at children and infants. The largest proportion of health treatments (40%) was aimed at children. The majority of detection and diagnosis interventions were directed towards perinates (34%). Health programs were fairly evenly distributed across age groups.

4.1.5 Outcome measures

The specific outcome measures reported in the literature were categorized as indicated in Table 4.5. The most common category of outcome measure remained cases of disease/condition/ abnormality over the study period. The proportions of papers studying cases of death and cases of cures/improvements/healing declined slightly over time while changes in physiologic measures and cases of complications/adverse events became more common as outcome measures.

Table 4.5: Publications per outcome category by five-year period

Summary Outcome	1980-84		1985-89		1990-94		1995-99		Total	
	n	%	n	%	n	%	n	%	n	%
Cases of disease/condition/ abnormality	32	42%	63	53%	109	42%	236	40%	440	42%
Cases of death (all causes)	11	14%	23	19%	39	15%	61	10%	134	13%
Cases of cures/ improvements/ healing	13	17%	6	5%	29	11%	55	9%	103	10%
Changes in physiologic measure	3	4%	10	8%	26	10%	53	9%	92	9%
Cases of complications/ adverse events	4	5%	7	6%	19	7%	60	10%	90	9%
Life years gained	4	5%	1	1%	7	3%	34	6%	46	4%
QALYs, or similar unit	2	3%	2	2%	8	3%	21	4%	33	3%
Time outcome*	0	0%	2	2%	4	2%	25	4%	31	3%
Changes in behavioural/ social	3	4%	1	1%	8	3%	17	3%	29	3%
Cases of vaccination	1	1%	2	2%	3	1%	8	1%	14	1%
Health service/ process outcome	2	3%	2	2%	3	1%	6	1%	13	1%
Cases of injury	0	0%	0	0%	4	2%	5	1%	9	1%
Changes in quality of life	1	1%	0	0%	0	0%	4	1%	5	0%
Total	76	100%	119	100%	259	100%	585	100%	1,039	100%

*This category refers to days in a state or days absent from a state, time to achieve an outcome or to recover.

A cross-tabulation of intervention type by outcome category demonstrated that 58% of assessments of health prevention interventions used an outcome categorized as cases of disease/condition/abnormality. Studies of health treatments commonly utilized cases of disease/condition/abnormality and cases of cures/improvements as outcome measures. Seventy-four percent of assessments of health programs were fairly evenly distributed across the outcome categories, cases of disease/condition/abnormality, cases of death and changes in physiologic measures.

4.1.6 Analytic technique

Marked changes in the type of analytic technique used over time was evident, as seen in Table 4.6. While cost-effectiveness analysis was the most common technique used, accounting for a majority of evaluations in all time intervals, the proportion of studies using CEA increased by 23 percentage points while the proportion of studies using CBA decreased from 31% in 1980-84 to just 12% in 1995-99. CUA remained the least common analytic technique used over all time intervals.

Table 4.6: Publications per analytic technique by five-year period

Analytic Technique	1980-84		1985-89		1990-94		1995-99		Total	
	n	%	n	%	n	%	n	%	n	%
CEA	32	52%	67	73%	154	79%	331	75%	584	74%
CBA	19	31%	20	22%	24	12%	53	12%	116	15%
CMA	8	13%	3	3%	8	4%	33	8%	52	7%
CUA	2	3%	2	2%	8	4%	23	5%	35	4%
Total	61	100%	92	100%	194	100%	440	100%	787	100%

A cross-tabulation of intervention type by analytic technique indicated that CEA was the most common technique for all intervention types. Cost-benefit analyses tended to be used to evaluate health prevention and detection/diagnostic interventions. Cost-utility analyses were usually applied to evaluations of health prevention interventions and health treatments.

4.1.7 Journal type

Publication of economic evaluations by journal type is presented in Table 4.7. Publication in journals of pediatrics/perinatal medicine was the most common venue for all time intervals and increased as a proportion of the total over time. Pediatric economic evaluations appearing in other sub-specialty journals also increased over time while the proportion of publications in public health and general medicine journals decreased over time. There was also an increase in the proportion of publications of pediatric economic evaluations in journals devoted to health economics/health policy/methods.

Table 4.7: Publications per journal type by five-year period

Journal Type	1980-84		1985-89		1990-94		1995-99		Total	
	n	%	n	%	n	%	n	%	n	%
Pediatrics/Perinatal Medicine	17	28%	23	25%	58	30%	156	36%	254	32%
Sub-specialty Medicine	13	21%	17	18%	37	19%	136	31%	203	26%
Public Health	9	15%	21	23%	42	22%	53	12%	125	16%
General Medicine	15	25%	17	18%	27	14%	55	13%	114	15%
Health Economics/Health Policy/ Methods	0	0%	3	3%	15	8%	22	5%	40	5%
Dentistry	5	8%	6	7%	7	4%	5	1%	23	3%
Pharmacology	2	3%	3	3%	3	2%	11	3%	19	2%
Other	0	0%	2	2%	3	2%	1	0%	6	1%
Total	61	100%	92	100%	192	100%	439	100%	784	100%

Missing records=3

A cross-tabulation of journal type by analytic technique demonstrated that while CEAs tended to be concentrated in pediatrics/perinatal medicine journals, CBAs and CUAs were more evenly distributed across journals of pediatrics/perinatal medicine, sub-specialty medicine, public health and general medicine.

4.2 Quality Appraisal Analysis

The quality appraisal analysis was conducted on a random sample of papers, stratified by five-year period. In comparing the random sample to the parent dataset, there were no significant differences with respect to distributions for age group, analytic technique, intervention category, ICD-9 CM disease class and journal type. The quality appraisal analysis included assessments of inter-rater and test-retest reliability, descriptive statistics of domain scores, correlations among domains and between domains and five-year period and global score. Also included were descriptive statistics of analytic technique, costing methods, uncertainty analysis and the six-level global rating (questionnaire item #57). Each analysis is described below.

4.2.1 Inter-rater reliability

Two independent reviewers conducted quality appraisals of a random sample of publications spanning the 20-year study period. The results are presented for the 149 papers appraised by both reviewers. The results of the inter-rater reliability assessment are displayed in Table 4.8.

Table 4.8: Inter-rater reliability

Domain	Intraclass Correlation Coefficient	95% Confidence Interval
Economic Evaluation	0.53	(0.41, 0.64)
Comparators	0.53	(0.41, 0.64)
Target Population	0.15	(-0.01, 0.31)
Time Horizon	0.25	(0.09, 0.39)
Perspective	0.81	(0.74, 0.86)
Costs & Resource Use	0.50	(0.37, 0.61)
Outcomes	0.28	(0.13, 0.43)
Analysis	0.52	(0.39, 0.63)
Discounting	0.85	(0.80, 0.89)
Incremental Analysis	0.47	(0.34, 0.59)
Sensitivity Analysis	0.83	(0.77, 0.87)
Conflict of Interest	0.82	(0.76, 0.87)
Conclusions	0.37	(0.22, 0.50)
Total Questionnaire Score	0.75	(0.66, 0.81)

The inter-rater reliability, as assessed by the intra-class correlation coefficient, was 0.75 for the overall questionnaire. High inter-rater reliability was observed for the following domains: Discounting, Sensitivity Analysis, Conflict of Interest and Perspective. Low inter-rater reliability was observed for Target Population, Time Horizon, Outcomes and Conclusions. The remaining domains demonstrated intermediate inter-rater reliability.

4.2.2 Test-retest reliability

One reviewer (AL) re-appraised 10 papers two months after the initial appraisal. The results of the test-retest reliability assessment are displayed in Table 4.9.

Table 4.9: Test-retest reliability

Domain	Intraclass Correlation Coefficient	95% Confidence Interval
Economic Evaluation	0.86	(0.55, 0.96)
Comparators	0.83	(0.46, 0.96)
Target Population	-0.29	(-0.76, 0.37)
Time Horizon	0.71	(0.20, 0.92)
Perspective	0.81	(0.40, 0.95)
Costs & Resource Use	0.77	(0.31, 0.94)
Outcomes	0.88	(0.59, 0.97)
Analysis	0.83	(0.46, 0.96)
Discounting	1.00	(1.00, 1.00)
Incremental Analysis	0.76	(0.29, 0.93)
Sensitivity Analysis	0.89	(0.62, 0.97)
Conflict of Interest	0.79	(0.35, 0.94)
Conclusions	0.58	(-0.03, 0.88)
Total Questionnaire Score	0.92	(0.71, 0.98)

The test-retest reliability was 0.92 for the overall questionnaire. High test-retest reliability was observed for almost all domains. Low test-retest reliability was observed for Target Population and Conclusions.

4.2.3 Quality of publications

The quality of the sampled publications of pediatric economic evaluations can be determined by examining the domain scores, which varied from zero to one, and the global rating. The mean scores for all domains in the quality appraisal questionnaire are presented in Table 4.10.

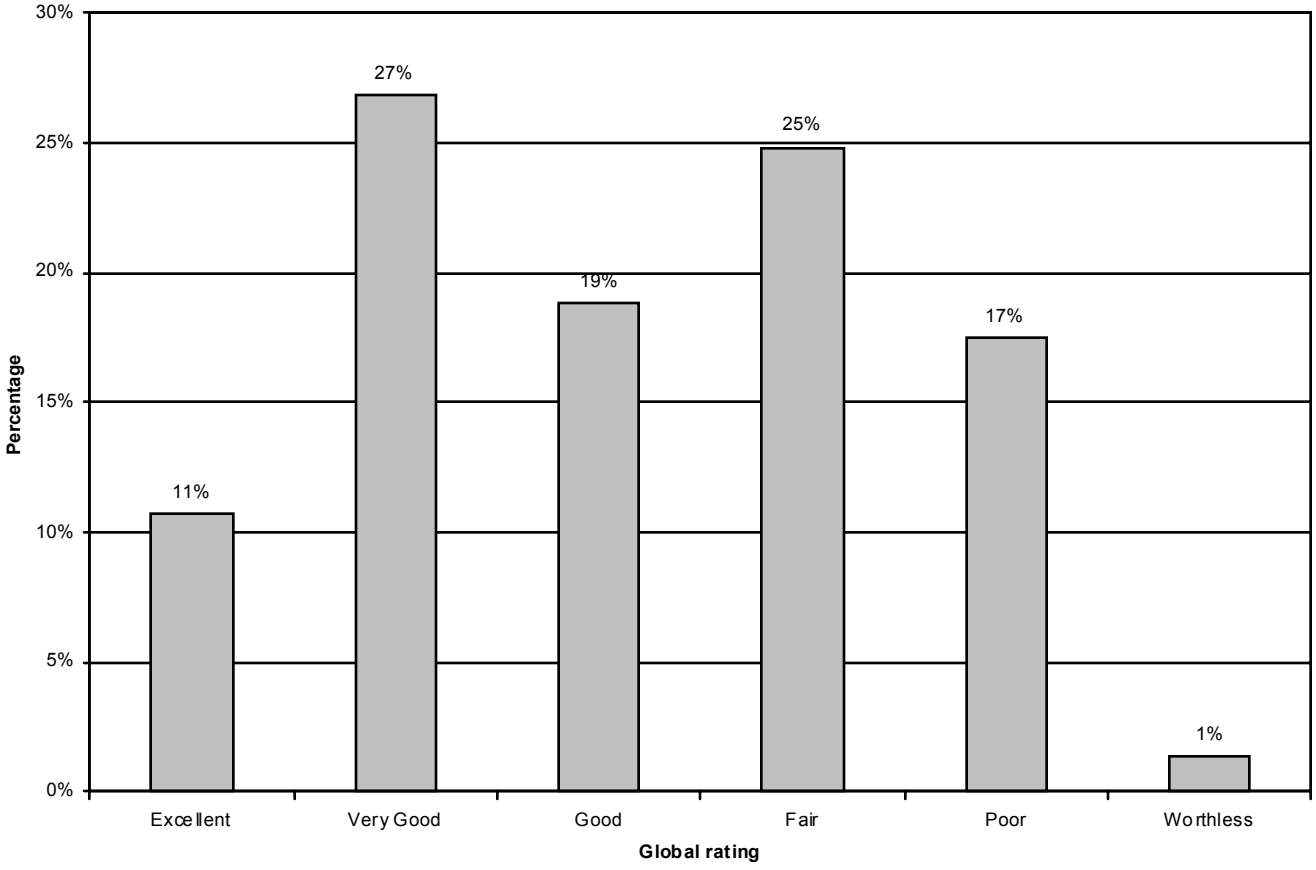
Table 4.10: Domain scores

Domain	Mean	Standard Deviation
Economic Evaluation	0.76	0.31
Comparators	0.74	0.17
Target Population	0.77	0.21
Time Horizon	0.50	0.29
Perspective	0.39	0.27
Costs & Resource Use	0.62	0.30
Outcomes	0.57	0.18
Analysis	0.50	0.28
Discounting (n=73)	0.78	0.41
Incremental Analysis	0.35	0.28
Sensitivity Analysis	0.46	0.41
Conflict of Interest	0.32	0.43
Conclusions	0.75	0.24

All domains except Discounting were represented by at least one item in all of the publications. The Discounting domain was relevant only for 73 publications. The sampled papers demonstrated good quality for the following domains: Discounting, Target Population, Economic Evaluation, Conclusions and Comparators. The papers were poor quality with respect to Conflict of Interest, Incremental Analysis and Perspective. The remaining domains displayed quality scores around the mid-point of 0.5.

Figure 4.2 displays the global rating frequency distribution. Just over one third (38%) were rated as very good to excellent.

Figure 4.2: Global rating



To determine whether quality improved over time, we computed the correlations between domain scores and five-year period (Table 4.11).

Table 4.11: Correlation of domain scores with five-year period

Domain	Spearman correlation coefficient	p
Economic Evaluation	0.03	0.6865
Comparators	0.13	0.1253
Target Population	-0.03	0.7398
Time Horizon	0.03	0.7309
Perspective	0.13	0.1116
Costs & Resource Use	0.16	0.0461
Outcomes	0.18	0.0250
Analysis	0.14	0.0852
Discounting	-0.03	0.7770
Incremental Analysis	0.02	0.7726
Sensitivity Analysis	0.07	0.3905
Conflict of Interest	0.13	0.1286
Conclusions	0.23	0.0049
Global Score	0.15	0.0594

While most domain scores showed no change over time, there were weak but significant positive correlations observed between five-year period and Costs & Resource Use, Outcomes and Conclusions. Table 4.12 presents the correlation between the global rating and each domain score.

Table 4.12: Correlation of domain scores with global rating

Domain	Spearman correlation coefficient	P
Economic Evaluation	0.63	< 0.0001
Comparators	0.53	< 0.0001
Target Population	0.48	< 0.0001
Time Horizon	0.48	< 0.0001
Perspective	0.64	< 0.0001
Costs & Resource Use	0.66	< 0.0001
Outcomes	0.40	< 0.0001
Analysis	0.83	< 0.0001
Discounting	0.60	< 0.0001
Incremental Analysis	0.72	< 0.0001
Sensitivity Analysis	0.83	< 0.0001
Conflict of Interest	0.25	0.0025
Conclusions	0.57	< 0.0001
Total Questionnaire Score	0.92	< 0.0001

Statistically significant correlations between the global rating and domain scores were observed for all domains. Strong correlations were observed for Analysis, Sensitivity Analysis, Incremental Analysis and Costs and Resource Use.

Analyses were undertaken to examine the correlations between the scores of key domains specified in section 3.3.5 d) and the other domains. The strongest correlations were observed between Analysis and other domains (Table 4.13). All domains demonstrated statistically significant correlations with Analysis. The strongest correlations were found with Sensitivity Analysis, with Costs and Resource Use, with Incremental Analysis and with Economic Evaluation.

Table 4.13: Correlation of domain scores with analysis score

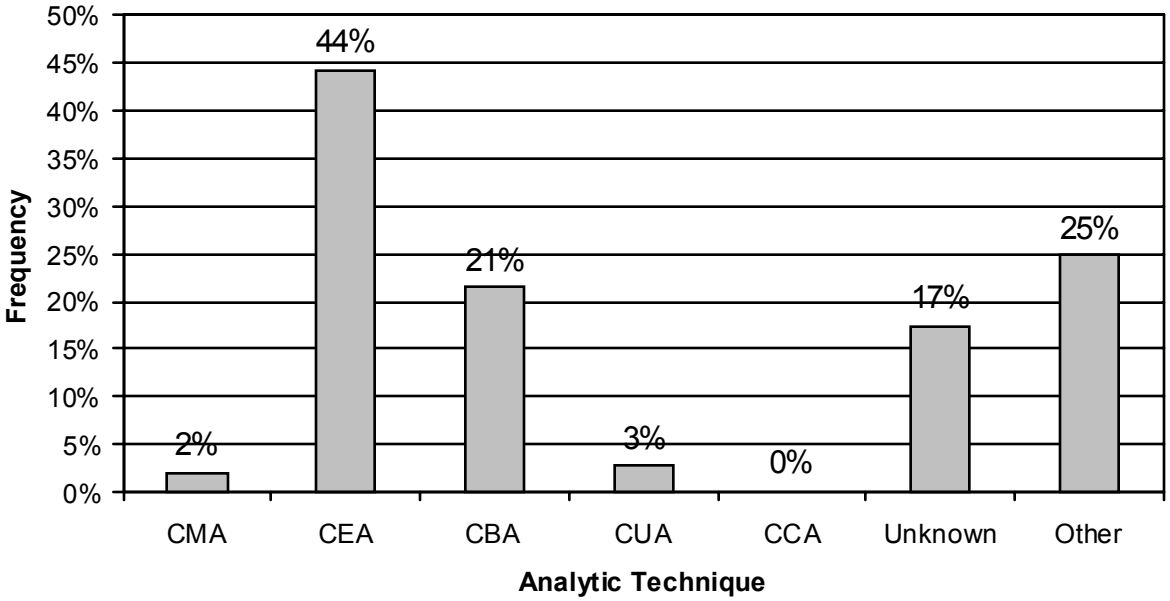
Domain	Spearman correlation coefficient	p
Economic Evaluation	0.61	< 0.0001
Comparators	0.50	< 0.0001
Target Population	0.41	< 0.0001
Time Horizon	0.44	< 0.0001
Perspective	0.58	< 0.0001
Costs & Resource Use	0.70	< 0.0001
Outcomes	0.36	< 0.0001
Discounting	0.49	< 0.0001
Incremental Analysis	0.66	< 0.0001
Sensitivity Analysis	0.76	< 0.0001
Conflict of Interest	0.19	0.0177
Conclusions	0.47	< 0.0001

4.2.4 Publication characteristics

In addition to the questionnaire items that were included within the domains, a number of questions were included for the purpose of description and further characterization of the publication (see section 3.3.3).

Figure 4.3 presents a frequency distribution of the analytic technique. Within the 149 appraised papers, there were 168 economic analyses, as some papers reported more than one analytic technique.

Figure 4.3: Analytic technique



Cost-effectiveness analysis and cost-benefit analysis comprised two thirds of analyses (66%). A high proportion of papers (42%) were designated as 'unknown' or 'other' as they did not specify a conventional analytic technique. This speaks to the knowledge and skill level of the author. Upon further examination, a CEA was implicit in the majority of unclassified papers.

An analysis of variance compared the mean domain scores by analytic technique. To ensure that a primary technique was assigned to every paper, the analytic technique listed in the PEDE database was used, rather than the appraiser's designation. When more than one analytic technique was used in a study, the PEDE database listed the most significant and/or highest quality analysis. Table 4.14 presents the mean domain scores by analytic technique for those domains that had a statistically significant p value for the ANOVA model F test ($p < 0.05$).

Table 4.14: Domain score by analytic technique

Domain	CUA (n=9)	CBA (n=23)	CEA (n=111)	CMA (n=6)
Economic Evaluation	0.97 (A)	0.91 (A,B)	0.71 (B)	0.79 (A,B)
Time Horizon	0.64 (A,B)	0.70 (A)	0.44 (B)	0.63 (A,B)
Perspective	0.47 (A,B)	0.62 (A)	0.35 (B)	0.28 (B)
Analysis	0.74 (A)	0.59 (A,B)	0.47 (B)	0.42 (B)
Discounting	0.78 (A)	0.70 (A)	0.31 (B)	0.00 (B)
Incremental Analysis	0.59 (A)	0.48 (A,B)	0.32 (B)	0.10 (C)
Sensitivity Analysis	0.85 (A)	0.73 (A)	0.39 (B)	0.22 (B)

Note: Means with the same letter are not significantly different. The p-value for a significant difference was $p < 0.05$.

Analytic technique was a significant predictor of domain score for most of the study design-related domains. Except for Time Horizon and Perspective, studies designated as CUAs demonstrated the highest domain scores and were significantly higher than those of CEAs and most CMAs. Cost-benefit analyses demonstrated the highest scores for Time Horizon and Perspective. CBAs were also superior to CEAs for Discounting and Sensitivity Analysis.

The questionnaire includes several descriptive questions regarding costing methods. A wide variety of methods were used to assign values to direct cost items. The frequencies of these methods are presented in Table 4.15.

Table 4.15: Direct cost valuation

Direct costs valuation	Frequency	%
Charges or fees	66	44%
Market or wholesale prices, replacement costs	66	44%
Average cost	58	39%
Fixed, overhead, capital or administrative costs	47	32%
Assumption, opinion, expert panel	40	27%
Unknown/Not stated	37	25%
Opportunity cost	21	14%
Deflated charges	7	5%
Other	7	5%

Most studies undertook multiple costing methods, including the use of charges, fees, market prices and average costs. Expert opinion or assumptions were used in 27% of direct cost valuations. Indirect costs, also referred to as productivity costs, were also valued using multiple methods. The variety of methods used to value productivity costs is exhibited in Table 4.16.

Table 4.16: Productivity cost valuation

Productivity costs valuation	Frequency	%
Not applicable	117	79%
Opportunity cost approach	16	11%
Average statistical wage from population database	12	8%
Market value approach	2	1%
Other	2	1%
Friction cost method	0	0%

The majority of publications (79%) did not include productivity costs. In the studies that did include this cost category, an opportunity cost approach was most often used.

The results of the published studies were sometimes reported with more than one unit of analysis. As displayed in Table 4.17, the results were most frequently expressed as costs and outcomes per *child*. Alternatively, costs and outcomes were expressed per patient sample or per population.

Table 4.17: Unit of analysis

Unit of Analysis	Frequency	%
Per child or patient or case	96	64%
Per patient sample or study sample	41	28%
Per population	23	15%
Per parent	1	1%
Per family	1	1%
Per joint parent-child	1	1%

Although most papers (52%) failed to perform any uncertainty analysis, of those that did, one-way sensitivity analysis was performed almost exclusively (Table 4.18).

Table 4.18: Uncertainty analysis

Sensitivity analysis	Frequency	%
None	77	52%
One-way sensitivity analysis	65	44%
Multi-way sensitivity analysis	7	5%
Bootstrapping or Monte Carlo simulations	3	2%
Two-way sensitivity analysis	0	0%
Other	0	0%

5 DISCUSSION

5.1 Trends in Pediatric Economic Evaluations

While the volume of published economic evaluations in the pediatric population is far less than that of the adult population, this research reflects the increasing importance of economic evaluations to assess the economic benefit of interventions, programs and services for children. Significant growth in the number of published pediatric economic evaluations was observed over the twenty-year study period, with the number of publications doubling every five years.

A number of findings from this research underscore the differences in health care between the pediatric and adult populations. The largest proportion of studies represented health prevention interventions rather than treatments of specific diseases or disorders. This is consistent with an epidemiological model of child health where the incidence of pediatric disease is rare and the prevention of future pediatric illness and illness occurring in adult years is emphasized. The high frequency of health prevention interventions is supported by the finding that infective and parasitic disease accounted for the largest proportion of ICD-9-CM categories. Among these studies, malaria control and vaccination strategies for hepatitis B, Hemophilus influenzae type B, measles, and varicella among children aged 1 to 12 years were the most prevalent.

This research project included adult pregnant or breast-feeding women as long as outcomes were measured in the offspring. Interventions in this population, although administered to adults, were aimed at influencing child health. Taken together, the three ICD-9-CM categories of congenital anomalies, complications of pregnancy and perinatal conditions accounted for 30% of all publications, a larger proportion than infectious disease. These studies evaluated preventative interventions, treatments, detection and diagnostic strategies, health programs and surgical procedures. Over half of all economic evaluations of detection/diagnostic strategies (53%), over one third of all health programs (36%) and over one third of all surgical interventions (37%) were for congenital anomalies, complications of pregnancy and perinatal conditions. Among the disease categories, congenital anomalies, complications of pregnancy and perinatal conditions demonstrated the largest publication growth rates between 1990-94 and 1995-99, with the number of studies of congenital anomalies increasing by 231%. The growth in economic evaluations of diagnostic and treatment strategies for congenital and perinatal conditions is expected to continue with the advent of expensive new technology to support early life and as research advances in genomics are translated into genetic testing and treatment.

Outcome measures that examined the number of cases of a disease, a condition or an abnormality were the most frequent and are consistent with the preponderance of health prevention interventions. These measures represent intermediate outcomes. More final outcomes would include mortality, life years gained and QALYs. While the use of final outcomes represented only 20% of the total, they did demonstrate a sustained growth in frequency over the study period. Final outcome measures are particularly problematic for the pediatric population, unless the study is focused on a terminal disease of childhood such as a severe musculoskeletal disorder or cancer. Currently no published health economic models exist for measuring life years gained over long time horizons that include periods of maturation, development and rapid physiological change.

A particular challenge in building the PEDE database was assigning an analytic technique. Given the high frequency of evaluations of preventive measures for infectious disease, many of the studies were self-labeled as cost-benefit analyses. In addition, the early publications pre-dated the widespread acceptance of the conventional analytic techniques defined as CEA, CBA, CUA and CMA. These early studies tended to be self-labeled as cost-benefit analysis regardless of the actual analytic technique used. To circumvent the problem of mis-labeled studies, the research team ignored the authors' label and assigned one based on a careful reading of the methods used. Thus many so-called CBAs that examined the incremental cost of cases prevented were re-labeled as CEAs. However, if a so-called CBA monetarized the outcome measure (even if it did so poorly), the study remained a CBA. Our results indicated that true CEAs accounted for a large majority of publications. CUAs were rare, even in 1995-99. This is probably due to the difficulty in ascertaining life years gained, as discussed above, as well as the challenge of estimating utilities for health states in children. The persistent misuse of the terms cost-benefit analysis and cost-effectiveness analysis continues to hamper the conduct of literature searches for the purpose of methodological research.

5.2 Quality Appraisal of the Pediatric Economic Evaluation Literature

A key research objective was to determine if the observed growth in publications of pediatric economic evaluations in the 1980s and 1990s was accompanied by an increase in study quality. Existing quality appraisal checklists and instruments for assessing adult economic evaluations were deemed inadequate for appraising the pediatric economic evaluation literature. Thus a vital aspect of the research was the development of a valid and reliable instrument for critically appraising the quality of pediatric economic evaluations. This proceeded through a systematic process of item selection and item reduction, review by a panel of experts, pilot-testing and reliability assessment. Whereas there is good consistency among published instruments used for appraising the adult health economic literature, there was little agreement among experts regarding which items were essential for inclusion in a pediatric instrument. Our comprehensive development process resulted in a 57-item instrument with a high degree of test-retest and inter-rater reliability.

There was consensus among the members of our expert panel that it would be inappropriate to assume that all items in the questionnaire had equal weight to calculate a simple total score. It was also problematic to assign weights to individual items to determine a single summary score. A reasonable compromise between presenting individual item scores and a single summary score was to score the instrument according to fourteen domains, each pertaining to a critical aspect of economic evaluation and scored from zero to one. Domains closely related to the elements of economic evaluation, i.e. Time Horizon, Perspective, Costs and Resource Use and Outcomes, demonstrated quality scores around the midpoint of 0.5. Domains related to analysis, including Analysis, Incremental Analysis and Sensitivity Analysis fared poorly. The lowest average domain score of 0.32 was observed for Conflict of Interest. The absence of sponsorship and conflict of interest declarations in the majority of papers may be due to publication prior to the present movement toward full disclosure. Even today, full disclosure of financial relationships is not required by many journals. There is much room for improvement in the quality of pediatric economic evaluations. The observed quality of the publications may have been related to the unique challenges of conducting economic evaluations in children.

The performance of the domain scores was reflected in the global rating, the last item in the questionnaire. Forty-four percent of publications were rated as good or fair. Most of the domains showed no improvement over time. Only Outcomes, Costs & Resource Use and Conclusions were weakly but significantly correlated with five-year period.

Classifying the papers by analytic technique proved to be a challenge, with 43% of analyses unclassified by the appraiser. Upon inclusion in the PEDE database, these papers were assigned a primary analytic technique based on a close reading of the methods used. As discussed above, many unlabelled studies and mis-labelled CBAs were classified as CEAs. The distribution of analytic technique in the sample chosen for quality appraisal mirrored the distribution in the whole PEDE database, with the majority of analyses being CEAs. Despite their small numbers, CUAs demonstrated significantly higher quality than the other types of analyses, with high domain scores observed for Economic Evaluation, Discounting and Sensitivity Analysis. The true CBAs also showed superior quality to the CEAs.

The results of the quality appraisal of pediatric economic evaluations published from 1980 to 1999 were similar to earlier appraisals of the adult literature that demonstrated a disappointing performance. Incremental analysis and sensitivity analysis were often missing and the overall analysis was poorly done. Costing and outcome assessment exhibited mediocre results. These findings reiterate the importance of adherence to publicly available health economic guidelines that promulgate the conduct and publication of high quality economic evaluations. Furthermore, peer reviewers who critique manuscripts for publication in medical journals should become familiar with these guidelines and should insist on high quality manuscripts. Resource allocation decisions for the pediatric population are ill-served by poor quality economic evaluations.

5.3 Study Limitations

There were a number of limitations encountered in this research project. In building the full PEDE database, a complex search strategy was developed in order to include all relevant, publicly available studies. It is possible that the database missed economic evaluations conducted in the adolescent age group, when adolescents were studied together with adults or when adolescents were not identified with a "pediatric" keyword in the title or abstract. While every attempt was made to obtain relevant grey literature through perusal of web-sites, newsletters and textbooks, it is possible that some pertinent papers were missed. Foreign language articles were included in the PEDE database as long as an English abstract was available. However, foreign language articles were excluded from the quality appraisal sample. Despite this relatively small potential for missing citations, the multi-stage exhaustive search strategy employed in the building of the PEDE database, has resulted in a database that is vastly more comprehensive and inclusive than existing ones.

A particular challenge in identifying pediatric economic evaluations of health care interventions relates to the way 'health' is defined. In the pediatric population, health is intricately connected to behaviour and development. Many health problems in children are manifested as behavioural changes for which social services and educational interventions are advocated. Multiple care settings are involved in the delivery of care to children. These include physicians' offices, clinics, schools, homes and community agencies. Our focus on scanning only the medical literature resulted in the exclusion of economic evaluations of psychological or social service interventions that may be relevant for child health. A greater understanding of child health that

integrates psychological, educational and social service interventions will necessitate a broader approach to assessing health outcomes in children.

Finally, the PEDE database currently includes all publications from January 1, 1980 through December 31, 1999. It is expected that the number and quality of publications may change rapidly over the next few years. The PEDE database must be updated annually and quality appraisals of samples of papers repeated periodically to track changes over time and to help pinpoint the methodological deficits.

5.4 Future Research

The PEDE project has been beneficial in identifying avenues for future research. Future projects will focus on the distinct challenges of conducting health economic research in the pediatric population and the importance of approaching pediatric health economic research with a child-centered framework. For example, unlike the adult population, most economic studies in children focus on evaluating preventive interventions such as vaccination programs. Studies of preventive interventions that involve a short time horizon utilize an outcome measure such as cases averted and are thus classified as CEAs. However, most preventive interventions involve high up-front costs with health benefits deferred over a long time horizon, usually extending into adult life.

An important challenge remains determining the optimal analytic technique for this population. With CUA there are inherent difficulties in determining life years gained and estimating utilities in children. Cost-benefit analysis can monetarize benefits accruing over a lifetime for assessments of incremental net gain and can facilitate resource allocation decisions for screening/prevention activities across different diseases. However, there has been little research into willingness-to-pay and other methodologies for determining the benefits of preventive interventions over long time horizons in children.

The PEDE database revealed a paucity of studies using final outcome measures and CUAs. If CUAs are to be used more frequently in pediatrics, then models for predicting health outcomes and life years gained over long time horizons spanning periods of development and maturation must be developed. Additional research can examine the validity of novel outcome measures that reflect an integration of physiologic, behavioural and sociologic aspects of health status. Other issues relating to the use of parent proxy respondents for the reporting of health resource use, quality of life and health status need to be studied. This research points to the need to re-visit standard economic evaluation guidelines to ensure that issues pertaining specifically to the pediatric population are addressed.

In addition to the methodological research described above, the PEDE database can serve as a valuable research tool to academic researchers, health policy analysts and decision-makers alike. The results from the multi-national studies reported in the database may be used to inform health policy decisions in Canada and elsewhere. The studies may be of particular value with respect to planning of vaccination or other preventive and screening programs in developing countries. In the expanding field of health economics, many researchers focus on methods research. The PEDE database is a valuable tool for exploration of methodological challenges in the conduct of pediatric health economic evaluation. In addition, researchers interested in preparing a systematic review and/or meta-analysis of a particular pediatric disorder or aspect of economic evaluation will derive great benefit from such a resource.

Plans are underway to update the PEDE database on an annual basis and develop a web-based, user-friendly searchable database that may be accessed by any interested individuals. Appropriate links to pediatric health services research groups, health economic associations and information synthesis organizations, such as the Cochrane Collaboration, will be created.

6 CONCLUSIONS

The PEDE Project represents the first ever attempt to catalogue, analyze and appraise economic evaluations performed in a pediatric population. This original research project revealed dramatic growth in the volume of published pediatric economic evaluations over a 20-year period. Comprehensive in scope, the PEDE database is characterized by a large majority of CEAs. While evaluations of preventative interventions for infectious diseases constituted the single largest category of publication, the combined categories of congenital anomalies, complications of pregnancy and perinatal conditions demonstrated the greatest growth. The growth in economic evaluations of diagnostic and treatment strategies for congenital and perinatal conditions is expected to continue with the advent of expensive new technology to support early life and as research advances in genomics are translated into genetic testing and treatment.

This project involved the development of a valid and reliable 57-item instrument for the quality appraisal of pediatric economic evaluations. The use of this instrument in appraising the quality of 149 publications revealed mediocre quality with poor performance in areas related to analysis, including incremental analysis and sensitivity analysis. Few publications included disclosures of relationships to sponsors. There was little improvement in the quality of publication over the twenty-year period. Although few in number, CUAs demonstrated higher quality compared to CBAs and CEAs. These findings reiterate the importance of adherence to publicly available health economic guidelines that promulgate the conduct and publication of high quality economic evaluations. Furthermore, medical journals should insist on high quality manuscripts. Resource allocation decisions for the pediatric population are ill-served by poor quality economic evaluations.

Opportunities for future research relate to the development of statistical models to determine health benefits such as life years gained over long time horizons characterized by periods of development and maturation as well as research into utility assessment and willingness-to-pay approaches for children's interventions. The PEDE database, containing 787 citations, will be of great value for informing decision-makers as well as to researchers undertaking methods research, a systematic review or meta-analysis of a particular intervention or condition. The conceptual knowledge derived from this project will enhance the quality of economic assessment and will foster research initiatives that will ultimately improve health policy and allocation decision-making for children.

7 REFERENCES

1. Drummond M. Evaluation of health technology: economic issues for health policy and policy issues for economic appraisal. **Soc Sci Med** 1994;38(12):1593-600.

2. Power EJ, Eisenberg JM. Are we ready to use cost-effectiveness analysis in health care decision-making? A health services research challenge for clinicians, patients, health care systems, and public policy. **Med Care** 1998;36(5 Suppl):MS10-147.

3. Detsky AS. Using cost-effectiveness analysis for formulary decision making: from theory into practice. **Pharmacoeconomics** 1994;6(4):281-8.

4. Hurst J. The impact of health economics on health policy in England, and the impact of health policy on health economics, 1972-1997. **Health Econ** 1998;7 Suppl 1:S47-61.

5. Torrance GW, Blaker D, Detsky A, Kennedy W, Schubert F, Menon D, et al. Canadian guidelines for economic evaluation of pharmaceuticals. Canadian Collaborative Workshop for Pharmacoeconomics. **Pharmacoeconomics** 1996;9(6):535-59.

6. Jacobs P, Bachynsky J, Baladi JF. A comparative review of pharmacoeconomic guidelines. **Pharmacoeconomics** 1995;8(3):182-9.

7. Jameson EJ, Wehr E. Drafting national health reform legislation to protect the health interests of children. **Stanford Law Pol Rev** 1993;5(1):152-76.

8. Starfield B, Newacheck P. Children's health status, health risks and use of health services. In: Schlesinger MJ, Eisenberg L, editors. **Children in a changing health system: assessments and proposals for reform**. Baltimore: Johns Hopkins University Press; 1990.

9. Perrin JM. Introduction. In: Hobbs N, Perrin JM, editors. **Issues in the care of children with chronic illness**. San Francisco: Jossey-Bass; 1985. p.1-10.

10. Bronfenbrenner Y, Moen P, Garbarino J. Families and communities. In: Parke HR, editor. **Review of child development research**. Chicago: University of Chicago Press; 1984.

11. Halfon N, Newacheck PW. Characterizing the social impact of asthma in children. In: Weiss KB, Buist AS, Sullivan SD, editors. vol 38 in **Lung biology and health in disease** series. New York: Marcel Dekker; 2000.

12. Forrest CB, Simpson L, Clancy C. Child health services research. Challenges and opportunities. **JAMA** 1997;277(22):1787-93.

13. Agency of Health Care Policy and Research. **Child health services: building a research agenda. Report to the Committee on Appropriations, U.S. House of Representatives**. Rockville (MD): The Agency; 1997. AHCPR publ no 97-R055.

14. Santos MT, Ungar WJ. **Searching the cyber-library: the use of advanced on-line literature search methods in the conduct of a paediatric economic evaluation literature review**. Toronto: Hospital for Sick Children Research Institute; 1999.

15. Santos MT, Ungar WJ. **The Pediatric Economic Database Evaluation (PEDE) Project.** Toronto: Division of Population Health Sciences, Hospital for Sick Children Research Institute; 2000. Technical report.

16. Warner KE, Hutton RC. Cost-benefit and cost-effectiveness analysis in health care. Growth and composition of the literature. **Med Care** 1980;18(11):1069-84.

17. Ganiats TG, Wong AF. Evaluation of cost-effectiveness research: a survey of recent publications. **Fam Med** 1991;23(6):457-62.

18. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized control trials. **Med Care** 1992;30(3):231-43.

19. Udvarhelyi IS, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost-benefit analyses in the medical literature. Are the methods being used correctly? **Ann Intern Med** 1992;116(3):238-44.

20. Lee JT, Sanchez LA. Interpretation of "cost-effective" and soundness of economic evaluations in the pharmacy literature. **Am J Hosp Pharm** 1991;48(12):2622-7.

21. Gerard K. Cost-utility in practice: a policy maker's guide to the state of the art. **Health Policy** 1992;21(3):249-79.

22. Sacristan JA, Soto J, Galende I. Evaluation of pharmacoeconomic studies: utilization of a checklist. **Ann Pharmacother** 1993;27(9):1126-33.

23. Elixhauser A, Luce BR, Taylor WR, Reblando J. Health care CBA/CEA: an update on the growth and composition of the literature. **Med Care** 1993;31(7 Suppl):JS1-11, JS18-149.

24. Elixhauser A, Halpern M, Schmier J, Luce BR. Health care CBA and CEA from 1991 to 1996: an updated bibliography. **Med Care** 1998;36(5 Suppl):MS1-9, MS18-147.

25. Bradley CA, Iskedjian M, Lanctot KL, Mittmann N, Simone C, St Pierre E, et al. Quality assessment of economic evaluations in selected pharmacy, medical, and health economics journals. **Ann Pharmacother** 1995;29(7-8):681-9.

26. Blackmore CC, Magid DJ. Methodologic evaluation of the radiology cost-effectiveness literature. **Radiology** 1997;203(1):87-91.

27. Iskedjian M, Trakas K, Bradley CA, Addis A, Lanctot K, Kruk D, et al. Quality assessment of economic evaluations published in PharmacoEconomics. The first four years (1992 to 1995). **Pharmacoeconomics** 1997;12(6):685-94.

28. Trakas K, Addis A, Kruk D, Buczek Y, Iskedjian M, Einarson TR. Quality assessment of pharmacoeconomic abstracts of original research articles in selected journals. **Ann Pharmacother** 1997;31(4):423-8.

29. Pritchard C. **Trends in economic evaluation**. London: Office of Health Economics; 1998. OHE briefing no 36.

30. Briggs A, Sculpher M. Sensitivity analysis in economic evaluation: a review of published studies. **Health Econ** 1995;4(5):355-71.

31. Anell A, Norinder A. Health outcome measures used in cost-effectiveness studies: a review of original articles published between 1986 and 1996. **Health Policy** 2000;51(2):87-99.

32. Consensus on Health Economic Criterialist (CHEC). **CHEC-newsletter** 2000;(2). Available: http://www.beoz.unimaas.nl/chec/newsletter2.pdf.

33. Friede A, Taylor WR, Nadelman L. On-line access to a cost-benefit/cost-effectiveness analysis bibliography via CDC WONDER. **Med Care** 1993;31(7 Suppl):JS12-17.

34. **NHS economic evaluation database (NHS EED)** [database online]. York (UK): NHS Centre for Reviews and Dissemination, University of York; 2002. Available: http://agatha.york.ac.uk/nhsdhp.htm (accessed 2000 Aug 4).

35. **Database of abstracts of reviews of effectiveness (DARE)** [database online]. York (UK): NHS Centre for Reviews and Dissemination, University of York; 2002. Available: http://agatha.york.ac.uk/darehp.htm (accessed 2000 Aug 4).

36. **Health technology assessment (HTA) database** [database online]. York (UK): NHS Centre for Reviews and Dissemination, University of York; 2002. Available: http://agatha.york.ac.uk/htahp.htm (accessed 2000 Aug 4).

37. **The Cochrane Library** [database online]. Oxford (UK): Update Software; 2002. Available: http://www.update-software.com/Cochrane/default.HTM (accessed 2000 Aug 28).

38. Office of Health Economics, International Federation of Pharmaceutical Manufacturers' Associations. **Health economic evaluations database** [database online]. London: OHE-IFPMA Database Ltd.; 2002. Available: http://www.ohe-heed.com/internet.htm (accessed 2000 Aug 28).

39. **ICD-9-CM: international classification of diseases, 9th revision: clinical modification, 5th edition**. Los Angeles: Practice Management Information Corporation; 1998.

40. Canadian Coordinating Office for Health Technology Assessment. **Guidelines for the economic evaluation of pharmaceuticals: Canada**. 2nd ed. Ottawa: The Office; 1997. Available: http://www.ccohta.ca/ccohta_production/publications/pubs_e.asp.

41. Drummond MF, O'Brien B, Stoddart GL, Torrance GW, editors. **Methods for the economic evaluation of health care programs**. 2nd ed. Toronto: Oxford University Press; 1997.

42. Evidence-Based Medicine Working Group. **Worksheet for using an article about an economic analysis**. Edmonton: Evidence-Based Medicine Working Group, University of Alberta; 2000. Available: http://www.med.ualberta.ca/ebm/econ.htm.

43. Gerard K, Seymour J, Smoker I. A tool to improve quality of reporting published economic analyses. **Int J Technol Assess Health Care** 2000;16(1):100-10.

44. **Ontario guidelines for economic analysis of pharmaceutical products**. Toronto: Ontario Ministry of Health and Long-Term Care; 1994. Cat no 2227446. Available: www.gov.on.ca/health/english/pub/drugs/drugpro/economic.pdf.

45. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. **JAMA** 1996;276(16):1339-41.

46. **SPSS for Windows** [computer program].Version 10.0.05. Chicago: SPSS Inc.; 1999.

47. **SAS system for Windows** [computer program]. Version 8.02. Cary (NC): SAS Institute; 1999.

Appendix 1: Creation of PEDE Database

Step 1. Development of Search Strategy

The overall search strategy, depicted in Figure 1.1, was comprised of two steps, (1) the comprehensive retrieval of citations followed by (2) reduction, consisting of removal of inappropriate references. Initially, an inclusive search strategy was created to capture a list of all potential pediatric economic evaluations. The search strategy was tested against alternative strategies to ascertain the emergence of additional studies. If further studies were detected, the search strategy was refined accordingly until maximum sensitivity was attained, and a comprehensive list was retrieved. The strategy was designed to achieve high sensitivity (no false negatives). A low specificity (high false positive rate) was acceptable since all citations generated by the search strategy would be manually reviewed. In the reduction phase, unsuitable material (e.g. publications before 1980, editorials, comments, and studies performed in the adult population) was removed. The search strategy was revised as long as relevant material was not eliminated. To determine whether pediatric economic evaluations were eliminated, a quasi-random sample of citations eliminated through the reduction phase of the search strategy was reviewed. If the sample contained pediatric economic evaluations, changes were not made to the search strategy. Thus the validity of the search strategy was confirmed by examining the abstracts or full text of random selections of both accepted and rejected papers.

Figure 1.1: Search Strategy Overview

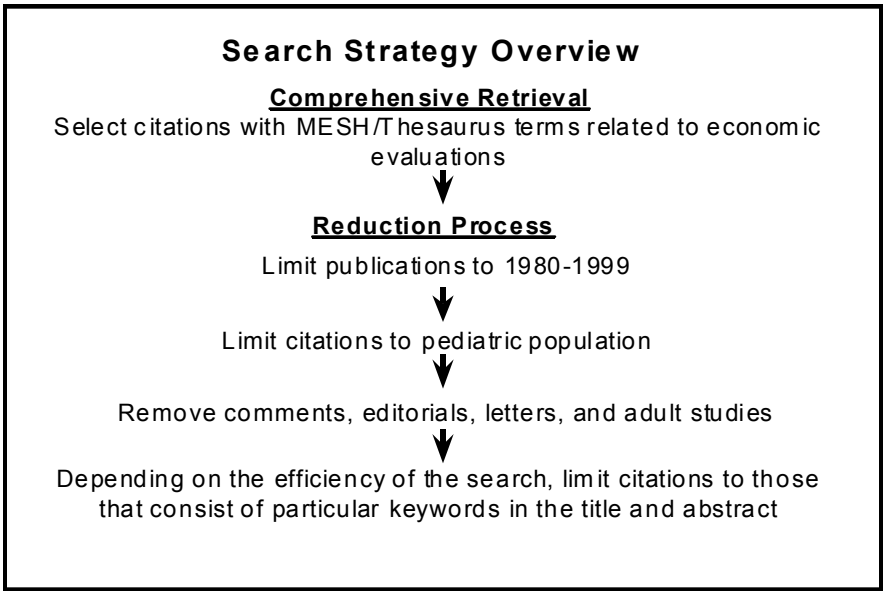


Table 1.1 lists the various sources of information that were searched. These sources included electronic journal databases, other electronic sources such as journals and newsletters, web sites providing access to grey literature such as working papers and government reports and other published material not available on the internet. The journal databases, e-journals, hard copy journals, textbooks and ScanDoc (weekly search of Current Contents) services were accessed through the University of Toronto and University Hospital Library Services.

Table 1.1: Pediatric Economic Evaluation Citation Sources

Journal Databases (published material)	**Websites** (published/grey literature)	**Electronic Sources**
EMBASE®	Government sites	Electronic journals
MEDLINE®	Research institutions	Current Contents
NHS EED	(universities, hospitals)	Listservs/Discussion Groups
HealthSTAR	Non-profit organizations	Electronic Newsletters
IPA	Sites involved with:	
Econlit	•health economics	**Library** (published material)
DARE	•health policy	
HTA	•health services and	Hard copy journals
Cochrane Library	outcomes research	Reports
CINAHL	•health technology	Monographs
CANCERLIT®	assessment	Text books
PubMed(PreMEDLINE)	•research support	

NHS EED=NHS Economic Evaluation Database, IPA=International Pharmaceutical Abstracts, DARE=Database of Abstracts of Reviews of Effectiveness, HTA=Health Technology Assessment Database

The final search strategy that was used to extract citations from MEDLINE® is depicted in Table 1.2. The same search strategy was applied to CINAHL, HealthSTAR, and CANCERLIT®. For these three databases a step was added to remove duplicate citations that would also be found in MEDLINE®. As not all databases use the same search term methods, the basic search strategy was replicated for use with EMBASE®, IPA, EconLit, Mental Health Abstracts, PubMed (PreMEDLINE), Current Contents (ScanDoc) and the Cochrane Library.

Table 1.2: MEDLINE®/CINAHL/HealthSTAR/CANCERLIT® Search Strategy

1	exp economics/ or exp technology assessment, biomedical/ or exp cost control/ or exp cost-benefit analysis/ or exp health care costs/ or exp health expenditures/ or exp "costs and cost analysis"/ or exp economic value of life/ or exp health care surveys/ or exp "health services needs and demand"/ or Health services research/
2	limit 1 to human
3	limit 2 to year=1980 to 1999
4	limit 3 to (newborn infant < birth to 1 month > or infant < 1 to 23 months > or preschool child < 2 to 5 years > or child < 6 to 12 years >)
5	limit 3 to adolescence < 13 to 18 years >
6	limit 5 to (adult < 19 to 44 years > or middle age < 45 to 64 years > or "aged < 65 and over >" or "aged, < 80 and over >")
7	(baby or babies or kid$ or youth$ or child$ or adolescent$ or teen$ or pediatric$ or paediatric$ or infant$ or newborn$ or neonate$ or student$).ab,ti.
8	6 and 7
9	4 or (5 not 6) or 8
10	limit 9 to (addresses or bibliography or biography or comment or congresses or consensus development conference or consensus development conference, nih or dictionary or directory or duplicate publication or editorial or festschrift or historical article or interview or lectures or legal cases or letter or news or overall or periodical index or retracted publication or retraction of publication)
11	9 not 10
12	("economic evaluation$" or "cost evaluation$" or "cost analys#s" or "cost-benefit" or "cost benefit" or "cost effectiveness" or "cost-effectiveness" or "cost utilit$" or "cost consequence$" or "cost$ and benefit$" or "cost$ and consequence$" or "cost$ and effect$" or "economic assessment$" or "cost assessment$" or QALY or cost$ or "quality adjusted life year$" or HYE or "healthy year equivalent" or "willingness to pay" or "willingness-to-pay" or "contingent valuation" or "cost minimi#ation").ab,ti.
13	11 and 12
14	exp fetus/ or exp prenatal care/ or exp prenatal diagnosis/
15	exp Pregnancy outcome/
16	14 or 15
17	1 and 16
18	limit 17 to human
19	limit 18 to (addresses or bibliography or biography or comment or congresses or consensus development conference or consensus development conference, nih or dictionary or directory or duplicate publication or editorial or festschrift or historical article or interview or lectures or legal cases or letter or news or overall or periodical index or retracted publication or retraction of publication)
20	18 not 19
21	20 and 12
22	13 or 21

A manual search through references alluding to economic evaluations was not executed since it was presumed that a majority of pediatric economic studies were captured using the comprehensive search strategy on various databases.

To include literature in the public domain but not published in the medical literature (grey literature), a wide array of health economics, health policy, evidence-based medicine, health services research and health technology assessment web sites were explored (Appendix 2), followed by the examination of relevant links. The web sites that were included in the search are listed in the Technical Report.[15] The University of York maintains three important databases: the National Health Services Economic Evaluation Database (NHS EED),[34] Database of Abstracts of Reviews of Effectiveness (DARE),[35] and the database of Health Technology Assessment (HTA).[36] The majority of their publications were cited by MEDLINE® or EMBASE®, however some were also retrievable through the HTA web site or through direct contact with individual health organizations.

All articles retrieved through the search strategies were imported electronically or input manually into EndNote 3.0.1, a citation management software program. This software enabled the elimination of duplicate citations that were retrieved from multiple electronic and other sources. Over 5,600 citations resulted from the search strategy.

Step 2. Manual Review

According to Drummond et al., an economic evaluation is defined as *the comparative analysis of alternative courses of action in terms of both their costs and consequences*.[41] For the PEDE Project, a publication was accepted as an economic evaluation only when a comparator existed and descriptions of both costs and health outcomes were present. The economic evaluation did not have to be the primary objective of the study to be eligible for the database. Table 1.3 lists the inclusion and exclusion criteria for determining eligibility.

Table 1.3: Inclusion/Exclusion Criteria for the PEDE Database

Inclusion Criteria

1. The study must contain original analyses and include the evaluation of an intervention, e.g. a medical or surgical treatment, a program, a service, or a new process.
2. The intervention must be directed at the pediatric population: neonates, infants, children and adolescents less than 18 years.
3. If outcomes were reported for adults or for the population as a whole, a pediatric group must be distinguished separately.
4. Interventions aimed at pregnant women or mothers were included if outcomes were measured in children (e.g. intervention/outcome - maternal medications/pediatric mortality; maternal smoking cessation/infant birth weight; maternal HIV screening/pediatric HIV cases).
5. A comparator must either be real or implied, e.g. pre- and post-intervention, "do nothing" or "usual care."
6. A health outcome, intermediate or final must be reported.
7. Costs must be measured and reported.
8. Randomized controlled trials and observational studies are eligible. Modeling studies and meta-analyses are eligible if they include novel data aggregation and new analyses.

Exclusion Criteria

1. Cost of illness or cost of prevention studies, where a specific intervention was not evaluated.
2. Family planning interventions related to birth control that include outcomes such as couple-year of protection or number of averted births. However, if the intervention was related to perinatal screening and the outcome was the number of averted births with a malformation or disease, the study was eligible.
3. Interventions consisting of a guideline, a continuous quality improvement, or a new operating procedure or policy targeted toward improving practice or efficiency.
4. Studies where costs were not quantified. (Merely reporting a reduction in the consumption of a specific resource, such as length of stay, without expressing the savings in monetary units, was inadequate.)
5. Abstracts from conference proceedings, methodological papers, papers without original analyses, policy papers, case studies or reports, letters, editorials or notes.

The citation abstracts that were extracted by the search strategy were manually scanned by a single researcher (MS) for the presence of a pediatric population, a comparator, costs and a health outcome. In certain studies, the comparator was not explicitly stated. These citations were accepted if the comparator was implied, as in a before/after design. Comparators that represented a "do nothing" or a "usual care" approach were accepted. Those abstracts difficult to assess for eligibility were reviewed by a second researcher (WU). The full text article of any abstracts that needed further verification for costs, outcomes, age category, comparator group, or type of publication were retrieved by downloading from the e-journal. If the journal was not available electronically, a hard copy of the publication was reviewed at one of the University of Toronto or Hospital libraries or requested on interlibrary loan. Publications written in languages other than English were accepted and entered into the database if an English abstract was available.

The manual review and the application of the inclusion and exclusion criteria resulted in a final database of 787 eligible publications.

Step 3. Data Extraction

Data entry and data management of eligible publications were performed using Microsoft ACCESS 97. Data were extracted and entered into the PEDE database from the 787 eligible citations. The following data fields were used for data entry.

- Notes
- Citation identification number
- Journal
- Target population
- Population category
- ICD-9-CM disease class
- Age group
- Experimental intervention
- Intervention category
- Primary health outcome
- Analytical technique

These fields are explained in greater detail below:

Notes - A field for comments.

Citation identification number - The unique citation identification number consists of the initial three letters of the first author's surname, the initial letter of the first author's first name, year of publication, a dash, and a unique # in case the author published more than once during the article's year of publication.

Journal - Full journal name.

Target population - A free-text, descriptive column that provides information on the study-specific population at which the intervention is aimed, such as disease, age and setting.

Categorized population - A one or two-word classification of the target population's disease, disorder, condition, or setting.

Disease Class - ICD-9 classification system for diseases, disorders, or medical conditions:[39]

- Infectious and Parasitic Disease
- Neoplasms
- Endocrine, Nutritional and Metabolic Diseases, and Immunity Disorders
- Diseases of the Blood and Blood-Forming Organs
- Mental Disorders
- Diseases of the Nervous System and Sense Organs
- Diseases of the Circulatory System
- Diseases of the Respiratory System
- Diseases of the Digestive System
- Diseases of the Genitourinary System
- Complications of Pregnancy, Childbirth, and the Puerperium
- Diseases of the Musculoskeletal System and Connective Tissue
- Congenital Anomalies
- Certain Conditions Originating in the Perinatal Period
- Symptoms, Signs and Ill-defined Conditions
- Injury and Poisoning

In the case of screening or preventive interventions, the disease variable refers to the condition that was being screened or prevented. With regard to antenatal care or family planning interventions, pregnancy was not considered a state of ill health. Thus a non-disease categories titled "General Health" and "Dental Health" were created for these and other interventions related to basic health services.

Age Group - Refers to the age group associated with the primary health outcome as defined below:

- Perinates: the antenatal period of the fetus or premature newborn, up to seven days of life
- Neonates: newborns, until the first month of age
- Infants one month to one year of age

- Children: one to 12 years of age
- Adolescents: 13 to 18 years of age

Experimental intervention - A free-text, general description of the study-specific experimental interventions.

Categorized intervention - The type of intervention was categorized as follows:
- Prevention (specified as Health or Dental): An intervention for the prevention of illness or disease. There was no distinction made between primary, secondary, or tertiary prevention.
- Treatment (specified as Health or Dental): An intervention administered directly to the patient for the cure or amelioration of a disease or condition.
- Program (specified as Health or Dental): An organization or organizational unit, clinic, department, or health system.
- Surgical: Operative procedures used to correct deformities and defects, repair injuries, diagnose, or cure certain diseases.
- Educational: An educational process or program designed for the improvement and maintenance of health.
- Health care delivery: A process, service, tool, test or treatment pathway.
- Detection: Tests used to screen for disease or the potential of developing disease.
- Diagnosis: The use of clinical tests to confirm the cause of illness.
- Psychological: Interventions used for normal and abnormal mental health conditions.

Primary health outcome - This variable represents a measure of health status or function in the pediatric population. The primary health outcome was directed by the study authors. If multiple health outcomes were listed, the first mentioned, the most clinically relevant, or the outcome most closely linked to the analytic technique (e.g. utility for a cost-utility analysis) was designated as the primary outcome.

Analytic technique - One or more of the analytic techniques listed below:
- Cost-utility analysis
- Cost-benefit analysis
- Cost-effectiveness analysis
- Cost-minimization analysis

To maintain consistency throughout the database, data entry decision rules were constructed for the various data fields of the database and applied to the relevant records.

<u>Categorized Population</u>
- The term "underserviced" was used to denote target populations from developing countries or regions in any country that had minimal access to health care. The term was not used aggressively throughout the database.
- Hearing impairment encompassed hearing loss within the "Categorized population" field.

Disease Category

- Children who were malnourished or receiving parenteral/enteral nutrition were classified under the disease category, "Endocrine, Nutritional and Metabolic Diseases, and Immunity Disorders."
- All dental interventions were listed under the disease category, "Dental Health."
- Interventions applicable to the ear, nose, and throat were inserted into the disease category, "Nervous system and sensory organs."
- Any interventions related to sedation were placed into the disease category, "Nervous system and sensory organs."
- Diseases/conditions as a result of premature birth were listed under the disease category, "Certain conditions originating in the perinatal period."
- The disease category for screening or preventive interventions that were performed on a universal basis is based on the actual condition or disease being screened or detected.
- The disease category for detecting infectious disease among pregnant women was "Complications of pregnancy, childbirth, puerperium."
- The disease category for a pregnant adolescent was "Complications of pregnancy, childbirth, puerperium."
- Studies of infectious disease in children were listed as "Infectious and Parasitic Disease".
- If the intervention involved the detection of a congenital anomaly or a specific disease within the fetus, then the disease category was listed as "Congenital anomaly" or else the specific category for the disease involved.

Age Category

- If the age group for children could not be deciphered from the citation or the full publication, the age group was imputed based on the usual target age group for the specified intervention.

Categorized Intervention

- All screening interventions were categorized as "detection" interventions.
- Early discharge or hospital-based interventions provided on an outpatient basis were categorized under "health program" interventions.
- The categorical intervention for vaccination strategies (e.g. methods of delivering vaccinations) was classified under "prevention", and not "health care delivery".

Outcome

- The primary outcome for antenatal screening interventions was cases of abnormality detected, not cases of abnormality averted due to abortion.
- Outcomes of mortality, death, or survival were derived strictly from the author's use of terminology within the publication, and not by definition, since mortality and survival may represent statistically distinct analyses.
- Interventions, such as assisted reproductive technology, whereby the primary outcome was cases of live births (and not cases of pregnancies), were included in the database.
- Cases of Haemophilus influenzae type B infection encompassed Haemophilus influenzae type B disease in the "Primary health outcome" field.
- Cases of breastfeeding or cases of cesarean deliveries were not considered pediatric primary health outcomes.

- Cases of adverse events were not equivalent to cases of complications.
- Cases of cures were not equivalent to cases of successful treatment.

Analysis

- A cost-minimization analysis was based on whether the *primary* health outcome was equivalent for the comparison groups.
- In the analysis field, only one technique is entered and is the primary technique used in the analysis. If unable to discern which is primary, the following hierarchy was used: CUA > CEA. CBA>CEA

Step 4. Creation of the PEDE Database

Data from the above fields were entered into an MS ACCESS 97 database for each citation. In addition to the data contained in the ACCESS database, the citation identification number was used to link the record to its full citation, including MeSH terms and abstract, in EndNote. To maintain consistency throughout the database, decision rules were followed for each data field. The PEDE database is thus a linked MS ACCESS – EndNote electronic database of 787 citations.

Step 5. Reliability Assessment of the Citation Selection Process

In addition to establishing the database, the research team completed an inter-rater reliability assessment and a test-retest reliability assessment for the citation selection process. For the inter-rater reliability assessment, an individual not associated with the project used a list of random numbers generated by SAS version 7.0. The list was used to select randomly 36 abstracts that were generated by the search strategy but did not meet the eligibility criteria and 14 abstracts that were acceptable. The 50 abstracts were combined in a random fashion. Two independent researchers (M. Santos and W. Ungar), blind to the classification status, reviewed the 50 citations and classified them as 'accept' or 'reject'. The percent observed agreement was 96% and the kappa coefficient, representing chance-adjusted agreement, was 0.91. This result denotes almost perfect reliability according to the Landis and Koch classification.[16] The test-retest reliability was also calculated for the researcher who performed the initial classification (M. Santos). The kappa coefficient for test-retest was 0.95, representing almost perfect agreement.[16]

Appendix 2: Internet Web Sites for Grey Literature

Health Technology Assessment

Canadian Coordinating Office for Health Technology Assessment (CCOHTA)	http://www.ccohta.ca/ccohta_production/home2_e.htm
Catalan Agency for Health Technology Assessment and Research (Spain)	http://www.aatm.es/
Finnish Office for Health Technology Assessment (FinOHTA)	http://www.stakes.fi/finohta/e/
Health Services/Technology Assessment Text - HSTAT (United States)	http://text.nlm.nih.gov/
Health Technology Assessment – International Links	http://epi-mh-hannover.de/hta-intern-links.html *(no longer operational)*
National Coordinating Centre for Health Technology Assessment (United Kingdom)	http://www.hta.nhsweb.nhs.uk/
New Zealand Health Technology Assessment	http://nzhta.chmeds.ac.nz/
Research & Development Directorate, NHS Executive South West, Publications	http://www.doh.gov.uk/research/swro/rd/publicat/

Health Economics

Canadian Health Economics Research Association (CHERA)	http://www.chera.ca/cgi-bin/WebObjects/chera
Centre for Health Economics and Policy Analysis (CHEPA) -McMaster University	http://www.chepa.org/
Centre for Health Economics Research and Evaluation (CHERE) - Sydney, Australia	http://www.chere.usyd.edu.au/
Databases in Health Economics - Univ. of York	http://www.york.ac.uk/inst/crd/econ4.htm
E-mail discussion lists in Health Economics - University of York	http://www.york.ac.uk/inst/crd/econ9.htm
Health Economic Research Centre (HERC) - University of Oxford	http://www.ihs.ox.ac.uk/herc/
Health Economics - Places to Go-University of Bayreuth	http://www.medecon.de/hec.htm
Health Economics Evaluation Database	http://www.ohe-heed.com
Health Economics Research Group (HERG) -Brunel University	http://www.brunel.ac.uk/depts/herg/
Health Economics Research Unit (HERU) - University of Aberdeen	http://www.abdn.ac.uk/heru/
Health Economics Resource Centre - University of York	http://www.york.ac.uk/res/herc/
Health Services Research Resources-Leonard Davis Institute of Health Economics	http://www.upenn.edu/ldi/hsr.html
HealthEconomics.com	http://www.healtheconomics.com/
Institute of Health Economics (IHE) -Working Paper Series	http://www.ihe.ca/publications/papers/
International Health Economics Association	http://www.healtheconomics.org/
Internet Resources in Health Economics – University of York	http://www.york.ac.uk/inst/crd/econ6.htm
NetEc homepage	http://netec.wustl.edu/NetEc.html

Health Services and Outcomes Research

Agency for Healthcare Research and Quality (AHRQ)	http://www.ahcpr.gov/
Center for Health Program Evaluation (CHPE) - Monash University and the University of Melbourne	http://chpe.buseco.monash.edu.au/
Center for Pharmaceutical Outcomes Research (CePOR)	http://www.unc.edu/~uwolt2/cepor/welcome.htm
Find it on the Internet: University of Toronto	http://www.library.utoronto.ca/gerstein/findanything.html
The Institute for Work & Health (IWH)	http://www.iwh.on.ca/
The Wessex Institute for Health Research and Development	http://www.wihrd.soton.ac.uk/

Research Support

Centre for Disease Control (CDC) WONDER Data Sets	http://wonder.cdc.gov/DataSets.shtml
Centre for Evidence-based Child Health	http://www.ich.bpmf.ac.uk/ebm/framnetw.htm
Centre for Research Support (CeReS)	http://www.ceres.uwcm.ac.uk/
Evidence Based Medicine Tool Kit	http://www.med.ualberta.ca/ebm/ebm.htm
Health Development Agency: Research	http://www.hda-online.org.uk/html/research/index.html
National Association of Children's Hospitals and Related Institutions (NACHRI) - Pediatric Resources	http://www.childrenshospitals.net/nachri/resources/
Unit for Evidence-Based Practice and Policy - UCL Department of Primary Care and Population Sciences	http://www.ucl.ac.uk/openlearning/uebpp/uebpp.htm

Government

Centre for Disease Control and Prevention	http://www.cdc.gov/
Food and Drug Administration	http://www.fda.gov/
Health Canada's Web Site	http://www.hc-sc.gc.ca/
National Institute of Health	http://www.nih.gov/
Ontario Ministry of Health and Long Term Care	http://www.gov.on.ca/health/
World Health Organization (WHO)	http://www.who.int/

Appendix 3: List of Journal Names

Academic Emergency Medicine
Academic Radiology
Acta Odontologica Scandinavica
Acta Paediatrica International Journal of Paediatrics
Acta Paediatrica Scandinavica
Acta Pediatrica Espanola
Acta Physiologica Scandinavica
Acta Psychiatrica Scandinavica
Acta Tropica
Adolescence
Advances in Oto-Rhino-Laryngology
AIDS
Alaska Medicine
Allergy
Ambulatory Child Health
American Heart Journal
American Journal of Cardiology
American Journal of Critical Care
American Journal of Diseases of Children
American Journal of Emergency Medicine
American Journal of Epidemiology
American Journal of Health-System Pharmacy
American Journal of Hospital Pharmacy
American Journal of Human Genetics
American Journal of Industrial Medicine
American Journal of Infection Control
American Journal of Managed Care
American Journal of Medical Genetics
American Journal of Obstetrics & Gynecology
American Journal of Otology
American Journal of Perinatology
American Journal of Public Health
American Journal of Tropical Medicine & Hygiene
American Psychologist
American Society for Artificial Internal Organs (ASAIO) Journal
American Surgeon
Anales Espanoles de Pediatria
Anesthesia & Analgesia
Anesthesiology
Annales de la Societe Belge de Medecine Tropicale
Annales de Radiologie
Annali dell Istituto Superiore di Sanita
Annali di Igiene
Annals of Emergency Medicine
Annals of Internal Medicine
Annals of Pharmacotherapy
Annals of Surgery
Annals of the New York Academy of Sciences
Annals of Thoracic Surgery
Annals of Tropical Medicine & Parasitology
Annals of Tropical Pediatrics
Archives de Pediatrie
Archives des Maladies du Coeur et des Vaisseaux
Archives of Disease in Childhood
Archives of Disease in Childhood Fetal & Neonatal Edition
Archives of Family Medicine
Archives of Otolaryngology - Head & Neck Surgery
Archives of Pediatrics and Adolescent Medicine
Archivos Latinoamericanos de Nutricion
ASDC Journal of Dentistry for Children
Atencion Primaria
Australian & New Zealand Journal of Public Health
Australian Dental Journal
Australian Health Review
Australian Journal of Public Health
Behavioral Interventions
Biology of the Neonate
Boletin Medico del Hospital Infantil de Mexico
Breastfeeding Review
Bristol Medico-Chirurgical Journal
British Dental Journal
British Journal of Medical Economics
British Journal of General Practice
British Journal of Nutrition
British Journal of Obstetrics and Gynecology
British Journal of Psychiatry
British Medical Journal
British Medical Journal Clinical Research Education
Bulletin of the Pan American Health Organization
Bulletin of the World Health Organization
Canada Diseases Weekly Report
Canadian Dental Association Journal
Canadian Journal of Anaesthesia
Canadian Journal of Cardiology
Canadian Journal of Infectious Disease
Canadian Journal of Nursing Research
Canadian Journal of Public Health
Canadian Journal of Surgery
Canadian Medical Association Journal
Cancer
Caries Research
Central African Journal of Medicine
Centre for Health Program Evaluation
Chemotherapy
Child Abuse & Neglect
Childs Nervous System
Chung-Hua Hu Li Tsa Chih Chinese Journal of Nursing
Chung-Hua Liu Hsing Ping Hsueh Tsa Chih Chinese Journal of Epidemiology
Chung-Hua Min Kuo Hsiao Erh Ko I Hsueh Hui Tsa Chih
Chung-Hua Yu Fang i Hsueh Tsa Chih
Clinical Drug Investigation
Clinical Excellence for Nurse Practitioners
Clinical Genetics
Clinical Infectious Diseases
Clinical Laboratory Management Review
Clinical Orthopaedics & Related Research
Clinical Otolaryngology and Allied Sciences
Clinical Pediatrics

Clinical Pharmacy
Clinical Therapeutics
Clinics in Perinatology
Community Dental Health
Community Dentistry & Oral Epidemiology
Community Medicine
Consult Pharm *
Cost-effectiveness in Health and Medicine (title of book)
Critical Care Medicine
Development and Evaluation Committee Report
Developments in Biological Standardization
Diabetes Care
Diabetes Research and Clinical Practice
Diabetic Medicine
Disease Management Clinical Outcomes
Drug & Alcohol Dependence
Drug Benefit Trends
Drugs Society
Ear and Hearing
Early Human Development
East African Medical Journal
Epidemiology & Infection
Ethiopian Medical Journal
European Journal of Obstetrics, Gynecology, & Reproductive Biology
European Journal of Pediatrics
European Urology
Evaluation & the Health Professions
Family Medicine
Family Process
Gaceta Sanitaria
Gastrointestinal Endoscopy
Gynecologic & Obstetric Investigation
Harefuah
Health Care Financing Review
Health Economics
Health Policy
Health Policy and Planning
Health Services Research
Health Technology Assessment
Helicobacter
Hepatology
HMO Practice
Hospital Formulary
Hospital Pharmacy
Hua-His I Ko Ta Hsueh Hsueh Pao
Indian Journal of Pediatrics
Indian Pediatrics
Infectious Diseases in Clinical Practice
International Journal of Epidemiology
International Journal of Experimental Clinical Chemotherapy
International Journal of Health Planning and Management
International Journal of Pediatric Otorhinolaryngology
International Journal of Technology Assessment in Health Care
Iowa Medicine
Israel Journal of Medical Sciences
J Pharm Clin *
Journal of Tropical Pediatrics
Journal -South Carolina Medical Association
Journal de Gynecologie Obstetrique et Biologie de la Reproduction
Journal de Radiologie
Journal Dentaire du Quebec
Journal of Acquired Immune Deficiency Syndromes
Journal of American Association for Pediatric Ophthalmology & Strabismus
Journal of Antimicrobial Chemotherapy
Journal of Applied Behavior Analysis
Journal of Asthma
Journal of Bone and Joint Surgery, Series B
Journal of Burn Care and Rehabilitation
Journal of Cardiovascular Electrophysiology
Journal of Chemotherapy
Journal of Child Neurology
Journal of Child Psychology & Psychiatry & Allied Disciplines
Journal of Clinical Epidemiology
Journal of Clinical Oncology
Journal of Clinical Pathology
Journal of Communication Disorder
Journal of Critical Care
Journal of Drug Education
Journal of Emergency Nursing
Journal of Epidemiology and Community Health
Journal of Family Practice
Journal of Florida Medical Association
Journal of Hand Surgery
Journal of Health Economics
Journal of Healthcare Education and Training
Journal of Hospital Infection
Journal of Infection
Journal of Infectious Diseases
Journal of Inherited Metabolic Disease
Journal of Intensive Care Medicine
Journal of Internal Medicine
Journal of Intravenous Nursing
Journal of Laparoendoscopic & Advanced Surgical Techniques
Journal of Laparoendoscopic Surgery
Journal of Laryngology and Otology
Journal of Maternal-Fetal Medicine
Journal of Medical Genetics
Journal of Medical Sciences
Journal of Medical Screening
Journal of Nurse-Midwifery
Journal of Obstetric, Gynecologic, & Neonatal Nursing
Journal of Orthopaedic Trauma
Journal of Paediatrics & Child Health
Journal of Parasitology
Journal of Parenteral & Enteral Nutrition

Journal of Parenteral Enteral Nutrition
Journal of Pediatric Gastroenterology Nutrition
Journal of Pediatric Hematology / Oncology
Journal of Pediatric Orthopedics
Journal of Pediatric Psychology
Journal of Pediatric Surgery
Journal of Pediatrics
Journal of Perinatal Education
Journal of Perinatal Medicine
Journal of Perinatology
Journal of Policy Analysis & Management
Journal of Public Health Dentistry
Journal of Public Health Medicine
Journal of Public Health Policy
Journal of Reproductive Medicine
Journal of School Health
Journal of Social Administrative Pharmacy
Journal of Telemedicine & Telecare
Journal of the American Academy of Audiology
Journal of the American College of Surgeons
Journal of the American Dental Association
Journal of the American Dietetic Association
Journal of the American Medical Association
Journal of the Dental Association of South Africa
Journal of the Indian Medical Association
Journal of the Irish Dental Association
Journal of the Medical Association of Thailand
Journal of the National Cancer Institute
Journal of Thoracic & Cardiovascular Surgery
Journal of Toxicology - Clinical Toxicology
Journal of Tropical Medicine & Hygiene
Journal of Tropical Pediatrics
Journal of Urology
Lancet
Medecine et hygiene
Medecine et maladies infectieuses
Medical and Pediatric Oncology
Medical Care
Medical Decision Making
Medical Interface
Medical Journal of Australia
Medicina Clinica
Medizinische Klinik
Military Medicine
National Bureau of Economic Research Working Paper
National Institute of Public Health Annals
Nederlands Tijdschrift voor Geneeskunde
Neonatal Intensive Care
New England Journal of Medicine
New Zealand Medical Journal
Nippon Koshu Eisei Zasshi-Japanese Journal of Public Health
Nursing Case Management
Nursing Clinics of North America
Nursing Research
Obstetrics & Gynecology
Ophthalmologe
Ophthalmology
Orthopedics
Pediatric Asthma Allergy & Immunology
Pediatric Cardiology
Pediatric Dermatology
Pediatric Emergency Care
Pediatric Endosurgery Innovative Techniques
Pediatric Hematology & Oncology
Pediatric Infectious Disease Journal
Pediatric Nephrology
Pediatric Neurology
Pediatric Pulmonology
Pediatric Radiology
Pediatric Surgery International
Pediatrics
Pediatrics International
Ph.D thesis, University of Pennsylvania
PharmacoEconomics
Pharmacotherapy
Pharmacy World & Science
Plastic & Reconstructive Surgery
Pneumologie
Postgraduate Medical Journal
Prenatal Diagnosis
Preventive Medicine
Problemy Tuberkuleza
Public Health
Public Health Nursing
Public Health Reports
Quintessence International
Rand Corporation
Respiratory Care
Revista Espanola De Salud Publica
Revista Portuguesa de Estomatologia e Cirurgia Maxilofacial
Revue d'Epidemiologie et de Sante Publique
Revue de Chirurgie Orthopedique et Reparatrice de l'Appareil Moteur
Ricerca e pratica
Ricerca in Clinica e in Laboratorio
Roentgenology
Royal Society of Tropical Medicine & Hygiene
Salud Publica de Mexico
Sante Mentale au Quebec
Scandinavian Journal of Infectious Diseases
Seattle: WAMI RHRC
Seminars in Hearing
Seminars in Thrombosis and Hemostasis
Skeletal Radiology
Social Biology
Social Science and Medicine
South African Medical Journal
South Carolina Medical Association Journal
Southeast Asian Journal of Tropical Medicine & Public Health

Sozial - und Praventivmedizin
Spine
Stem Cells
Supportive Care in Cancer
Surgery
Surgical Endoscopy Ultrasound and Interventional Techniques
Swedish Dental Journal
Technology Overview
The National Medical Journal of India
Tobacco Control
Transactions of the Royal Society of Tropical Medicine and Hygiene
Tropical Medicine and International Health
Tropical Medicine and Parasitology
U.K. National Health Service, South and West Regional Health
Ugeskrift for Laeger
Ultrasound in Obstetrics & Gynecology
Urology
Vaccine
World Health Forum
World Journal of Surgery
Zeitschrift fur Geburtshilfe und Neonatologie

The symbol * signifies the titles of journals that cannot be identified by the National Library of Medicine Journal Abbreviation Listings as of January 1999.

Appendix 4: List of Categorized Populations

abdominal pain
acute bacillary dysentery
acute renal allograft rejection
adenoidectomy
adolescents
amino acid metabolism
anemia
anemia of prematurity
appendicitis
Ascaris lumbricoides
asthma
attention deficit/hyperactivity disorder
bacterial infections
behavioural problem
bone marrow transplantation
burns
cancer
cardiac abnormality
cardiac arrest
chemotherapy
chickenpox
cholelithiasis
chronic intestinal failure
circumcision
club foot
colitis
congenital anomalies
congenital hip dislocation
cryptorchidism
cystic fibrosis
cystinosis
dental
dental extraction
dermatitis
developmental
diabetes
dialysis
diarrhea
ear, nose, throat
endoscopy
enteral nutrition
epilepsy
esophageal foreign bodies
extracorporeal membrane oxygenation
facial laceration
failure to thrive
febrile
fertile women
fetal lung immaturity
fractures
gastroenteritis
gastroesophageal reflux
gastrointestinal problems
gentamicin therapy
granulocytopenia
group B streptococcal disease
Guillain-Barre syndrome
headaches
healthy newborns
hearing disability
hematologic disorders
hemochromatosis
hemophilia A
hepatitis B
herpes
hip spica casts
Hirschsprung's disease
HMO
human immunodeficiency virus
hyaline membrane disease
hyperactivity
hypercholesterol
hypertension
hypertrophic pyloric stenosis
hypertrophyic pyloric stenosis
hypospadias
hypoxic respiratory failure
imperforated anus
impetigo
infertile
inguinal hernia
intensive care
intravenous therapy
Kawasaki syndrome
kidney transplant
lead poisoning
Legg Perthes disease
liver transplant
low birth weight
low income
lower respiratory tract infection
lysosomal storage disorders
magnetic resonance imaging scans
malaria
malnourished
measles
mechanical ventilation
meningitis
mental health
mentally retarded
multicystic dysplastic kidney
Neisseria gonorrhoeae
neonatal intensive care
nocturnal enuresis
non leukaemic malignant disease
otitis media
parenteral nutrition
pharyngitis
placenta previa
pneumothorax
poisoning
port-wine stains
post-surgery
pregnant
premature
previous cesarean section
primary feeding gastrostomy
psychosocial problems
pulmonary hypertension
pyelonephritis
pyeloplasty
radiography
renal biopsy
respiratory
resonance imaging scans
respiratory distress syndrome
respiratory failure
respiratory syncytial virus infection
retinopathy of prematurity
rhinopharyngitis
risk for neonatal group B streptococcal disease
risk for trisomy 18
risk for trisomy 21
salmonella
scalp laceration
schistosomiasis
schizophrenia
scoliosis
seizures
sepsis
sexually active
sickle cell disease
sleep apnea
smokers
special health care needs
speech disorder
spinal deformity
spinal surgery
splenic injury
strabismus repair
substance abuse
subungual hematomas
susceptible to HIV infection
susceptible to rota virus
thrombocytopenic purpura
thrombosis
tuberculosis
twins
underimmunized

underserviced
universal
upper respiratory tract infection
uretal reimplantation
ureteroneocystostomy
urinary tract infection
urolithiasis
urologic surgery
vaccinated
ventilator assisted
vertical assisted
vertical transmission of HIV
vesicoureteral reflux
Wilms tumor
without adequate prenatal care

Appendix 5: Members of the Expert Panel

Dr. Devidas Menon (Co-investigator) Executive Director and CEO Institute of Health Economics #1200, 10405 Jasper Avenue Edmonton, Alberta T5J 3N4	Dr. Upton Allen Pediatric Infectious Disease Specialist The Hospital for Sick Children 555 University Avenue Toronto, Ontario M5G 1X8
Dr. Ruth Collins-Nakai Pediatric Cardiologist University of Alberta Hospital-SCC/WCM 2C3.86-8440 112 St NW Edmonton, Alberta T6G 2B7	Dr. Cameron Donaldson Svare Professor of Health Economics University of Calgary Department of Community Health Sciences 3330 Hospital Drive NW Calgary, Alberta T2N 4N1
Dr. Thomas Einarson Professor, Faculty of Pharmacy University of Toronto 19 Russell Street Toronto, Ontario M5S 2S2	Dr. David Feeny Professor and Fellow Institute of Health Economics #1200, 10405 Jasper Avenue Edmonton, Alberta T5J 3N4
Dr. George Torrance Professor Emeritus McMaster University Vice-President, Scientific Affairs Innovus Research Inc. 1016-A Sutton Drive Burlington, Ontario L7L 6B8	

Date of Completion:	___ /___ / ____ ____ dd mm yyyy Initials
Data Entry Complete:	___ /___ / ____ ____ dd mm yyyy Initials

Appendix 6: Final Quality Appraisal Instrument

Citation Information

ID #: _____

Title: _____

_____

'Consequences' refers to impact on health, health status, well-being, disease incidence, etc. It does not include monetary benefits, such as savings. *The research question should appear in the body of the text (not merely title or abstract).*	**Economic Evaluation** 1. Is the research question posed in terms of costs and consequences? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
CMA, CEA, CBA, CUA or Cost-consequence analysis (CCA). *An analytic technique should be specified, beyond the use of 'cost-effective' or 'cost-benefit' as a generality or adjective.*	2. Is a specific type of economic analysis technique performed? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
(Check all that apply) *Indicate what the authors state is performed, even though they may be incorrect, such as labelling as a CBA what is actually a CEA. (Errors in identifying the technique or valuing the outcomes are addressed in Q38).*	3. What type of analytic technique is performed, according to the authors? ☐ 1. CMA ☐ 2. CEA ☐ 3. CBA ☐ 4. CUA ☐ 5. Cost-consequence analysis ☐ 6. Unknown/can't tell ☐ 7. Other (Specify: _____)
	Comparators 4. Is there a rationale for choosing the intervention(s) being investigated? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No

The rationale for comparators that are 'do nothing' or 'usual care' or pre-intervention practice may not be explicitly stated but may sometimes be inferred.	5. Is there a rationale for choosing the alternative program(s) or intervention(s) used for comparison? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
What are the setting, mode of delivery and timing of the investigational intervention(s) and comparator(s)? What is the dosing or administration regimen or intensity of exposure?	6. Does the report describe the alternatives in adequate detail? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
Are details describing disease and treatment-related events following the intervention(s) provided? For comparisons of screening interventions for detection of cases, the treatment pathway may not be described (not applicable). If screening is followed by treatment, or for prevention studies, an event pathway should be described.	7. Is a description of the event pathway provided? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
Were costs and consequences modeled in a decision tree with chance node probabilities derived from a variety of sources or established directly from a specific patient population? Includes Markov models. A decision analysis may have been performed even if the tree is not depicted.	8. Is a formal decision analysis performed? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated
For whom will the intervention be of value? What is the population for which an allocation decision is required?	**Target Population** 9. Is the target population for the intervention identified? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
Is the study sample similar to the target population with respect to age, sex, severity of disease, co-morbidities, health care system jurisdiction and other characteristics relevant to the intervention and research question? If the subjects are purely hypothetical, then indicate 'Not applicable'.	10. Are the subjects representative of the population to which the intervention is targeted? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated/Can't tell ☐ 5. Not applicable

If only one is explicit and the other is not explicit but can be inferred, then indicate: Yes (inferred from text, tables, or figures). Sometimes the time horizon can be inferred from the result, e.g. costs per episode of care, annual number of cases prevented, etc. It is possible to have costs incurred over a short period (< 1year) with outcomes assessed over a longer period. In this case, the time horizon is long, with costs incurred up front.	**Time Horizon** 11. Is there a time horizon for both costs and outcomes? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
The time frame should be long enough to capture all significant benefits, harms, and costs and relevant periods of child development. If no time horizon is stated, the response is 'no'.	12. Do the authors justify the time horizon selected? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
	Perspective 13. Is a perspective for the analysis given? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
	14. Is a societal perspective taken, either alone or in addition to other perspectives? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated
If only one perspective was studied, then check 'Not applicable'.	15. When there was more than one perspective, are the results of each perspective presented separately? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
Does the range of costs accurately reflect the perspective? When necessary, were fixed, capital costs and overhead costs included? Were costs incurred or averted related to subsequent follow-up and treatment of disease or complications considered?	**Costs and Resource Use** 16. Are all relevant costs for each alternative included? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No

In a societal perspective, these may be included as cost items (in the numerator) or as part of utility assessment (in the denominator). A patient or family perspective may also require these costs, if the lost time results in an income loss to the individual. For other perspectives, indicate 'Not applicable'. These costs may appear in the base case or in a sensitivity analysis.	17. Are opportunity costs of lost time (productivity costs) for parents and informal caregivers measured when required? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated ☐ 5. Not applicable
This would be expected for interventions such as screening or vaccination programs, that are directed at school age children. Interventions related to learning or behaviour would also involve school and/or community resources.	18. Do cost item identification and valuation extend beyond the health care system to include school and community resources when necessary? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated ☐ 5. Not applicable
Either within the stated time horizon of the study or a time horizon that is clearly appropriate for the research question. *This is relevant for papers that state that the prevention or treatment of the condition may impact on lifetime productivity as part of the rationale for doing the study.*	19. Are future salary and productivity changes of the child taken into consideration when appropriate? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated ☐ 5. Not applicable
Examples: questionnaires, interviews, chart abstraction, administrative databases, literature. *If some sources were missing or not mentioned, indicate 'No'.*	20. Were all of the sources for estimating the volume of resource use described? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
Possible sources include fee schedules, formularies, patient self-report, standard cost lists, wholesaler catalogues, case-costing databases, references to other costing papers. If some sources were missing or not mentioned, indicate 'No'.	21. Were all the sources for estimating all of the unit costs described? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No

For example, cases detected, for a screening intervention.	**Outcomes** 22. Is a primary health outcome given? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
	23. Do the authors justify the health outcome(s) selected? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
Effectiveness = routine use in a diverse population; Efficacy = experimentally controlled use in a uniform population *For screening interventions, effectiveness may be expressed as sensitivity and specificity.* *"Efficacy" is sometimes the term used when "effectiveness" is really what is measured.*	24. Is effectiveness, rather than efficacy assessed? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated/Can't tell
(Check all that apply) *Prospective may include clinical trials or observational studies using surveys or questionnaires. Retrospective may include chart reviews, or databases (insurance claims, prior surveys, completed studies). Prospective or retrospective data collection may involve examining data from the study institution.*	25. What approach was used to assess the effectiveness/efficacy? ☐ 1. Prospective data collection ☐ 2. Retrospective data collection ☐ 3. Literature sources ☐ 4. Expert opinion ☐ 5. Other (Specify:_____)
Citation of a reference is inadequate (indicate 'No') as this does not permit evaluation of the quality of the evidence.	26. Are the details of the design of the effectiveness/efficacy study(s) provided? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
This is also relevant for a CMA.	27. Are the results of the efficacy/effectiveness of alternatives reported? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable

Measurement of absences may not be applicable to studies of very young children or studies where the outcome is 'cases detected/prevented'. Measurement of absences may be a cost item to represent parental work absences or a health status/function measure.	28. Are school/day care absences taken into consideration? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
Almost all outcomes are intermediate and include: cases detected, lab or physical function measures, symptom reduction, complication rate. End benefit would be survival or life years gained. There must be evidence to link or extrapolate the intermediate measure to the final value for survival or life years gained. For CMA, check 'not applicable'.	29. If intermediate outcome variables are used, are they linked by evidence or reference to the end benefit? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
(Check all that apply) *If not measured, check 'Not applicable'.* *Standard generic instruments include the SF-36 and Child Health Questionnaire. If quality of life was part of a utility assessment, this is considered generic, preference-based. Direct preference includes standard gamble, time trade-off and visual analog. Indirect preference includes HUI, QWB. EQ-5D, EuroQol.*	**Quality of Life** 30. If quality of life was measured, what type of instrument was used? ☐ 1. Disease-specific ☐ 2. Generic, standard instrument ☐ 3. Generic, direct preference ☐ 4. Generic, indirect preference ☐ 5. Other (Specify:_____) ☐ 6. Not applicable
(Check all that apply) *For studies that include both children and adults, indicate the response relevant to the pediatric evaluation.*	31. Whose quality of life was assessed? ☐ 1. Child ☐ 2. Parent ☐ 3. Caregiver (if OTHER than parent) ☐ 4. Other (Specify: _____) ☐ 5. Not applicable
(Check all that apply)	32. Who performed the quality of life assessment? ☐ 1. Self assessment (child, parent or caregiver) ☐ 2. Parent on behalf of child ☐ 3. Caregiver on behalf or child ☐ 4. Health provider on behalf of child ☐ 5. Other (Specify: _____) ☐ 6. Not applicable

For example, costs should be measured in a currency unit and outcomes should be expressed as a natural unit for CEA, QALY (or equivalent) for CUA and currency unit for CBA. *If only costs or only outcomes are measured in appropriate units, then indicate 'No'.*	**Analysis** 33. Were costs AND outcomes measured in units appropriate for the indicated analytic technique? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated
(Check all that apply) *If data collection was retrospective (e.g. charts, database, literature), then check 'Not applicable'.* *The age limit for direct elicitation of responses may vary from study to study. For infants, indicate 'Not applicable'. For studies of interventions administered to pregnant or breastfeeding women, indicate 'Not applicable'.*	34. For prospective studies that used interviews, questionnaires or surveys, how were data obtained in studies involving young children? ☐ 1. Directly ☐ 2. Parent proxy ☐ 3. Other proxy (specify:_____) ☐ 4. Joint measure (specify:_____) ☐ 5. Not applicable
(Check all that apply)	35. How were direct costs valued? ☐ 1. Opportunity cost ☐ 2. Fixed, overhead, capital or administrative costs ☐ 3. Charges or fees ☐ 4. Deflated charges (using cost to charge ratios) ☐ 5. Market or wholesale prices, replacement costs ☐ 6. Average cost (per patient day, etc.) ☐ 7. Assumption, opinion, expert panel ☐ 8. Unknown/Not stated ☐ 9. Other method (specify:_____)
(Check all that apply) *Two types of Human Capital Approach:* *1. Market value = cost to hire a person, e.g. homemaker, to perform labour* *2. Opportunity cost = earned income if caregiver was in workforce* *Valuation may pertain to parent or caregiver or child's future earnings.*	36. How were productivity costs valued? ☐ 1. Market value approach ☐ 2. Opportunity cost approach ☐ 3. Average statistical wage from population database ☐ 4. Friction cost method ☐ 5. Other method (specify: _____) ☐ 6. Not applicable
When necessary, were adjustments for inflation or currency conversion made? Were costs allocated properly? Where market values were absent (e.g. volunteers), or market values did not reflect actual values, were adjustments made to approximate market values?	37. Were costs valued appropriately? ☐ 1. Yes (explicitly stated) ☐ 2 Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated/Can't tell

For example, in CBA, was the outcome monetarized using willingness-to-pay or other method? Health care savings resulting from disease prevention is insufficient as a monetarized outcome in CBA. For CUA, was utility assessed using a valid instrument or method? If assumptions or expert opinion was used to assign values for utilities or other outcome measures, indicate 'No'.	38. Was the valuation of outcomes appropriate for the type of analysis? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated/Can't tell
Check all that apply. *If the unit costs or an intermediate step express the result per patient but the final result is per population, indicate only 'per population'. If the result is expressed per study sample, then extrapolated to the population, check both. "Per family" refers to costs incurred by multiple members. "Per joint parent-child" refers to a single measure developed to link the parent and child.*	39. What was the unit of analysis used for expressing the **final** results? ☐ 1. Per child or patient or case ☐ 2. Per parent ☐ 3 Per family ☐ 4. Per joint parent-child ☐ 5. Per patient sample or study sample ☐ 6. Per population (e.g. per 10,000 or whole population)
In the case of decision analysis, if probabilities of events are stated explicitly with a description of the consumption of resources associated with each event, indicate 'Yes (explicit)'.	40. Were quantities of resources used reported separately from their unit costs? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
Were cost items summed correctly for each perspective? Was the "math right"? *If the aggregation is not explicit but the results can be reproduced from the data provided, then indicate 'Yes (inferred from text, tables, or figures)'.*	41. Were the costs aggregated correctly? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated/Can't tell ☐ 5. Not applicable
	42. Were details of statistical tests and confidence intervals given for stochastic data? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable

Discounting is not required if the time horizon is less than or equal to 1 year (indicate 'Not applicable'). *If only costs but not benefits were discounted, check 'No'.*	**Discounting** 43. When required, were costs and consequences that occur over more than one year discounted to their present values? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Unknown/Not stated ☐ 5. Not applicable
If the time horizon was less than or equal to 1 year, check 'Not applicable'. If time horizon was greater than 1 year and costs and benefits were discounted, check 'Not applicable'. *If an explanation was not provided for benefits when costs alone were discounted, check 'No'.*	44. If costs or benefits were not discounted when the time horizon exceeded 1 year, was an explanation provided? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
	Incremental Analysis 45. Are incremental estimates of costs and outcomes presented? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
Were the additional cost generated by one alternative over another compared to the additional effects generated? *An incremental ratio is 'not applicable' when one alternative is dominant over the other or when a CMA is conducted. A CBA can be expressed either as a ratio or a linear equation.*	46. Are the incremental estimates summarized as incremental ratios? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable
If incremental estimates for costs and outcomes (separate or in a ratio) were not reported, then check 'No'.	47. Were confidence intervals/limits calculated for incremental ratios or incremental estimates of costs and outcomes? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No

(e.g. discount rates, missing or imprecise resource use, missing prices)	**Sensitivity Analysis** 48. Were all important assumptions given? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
	49. Was a sensitivity analysis performed? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
If a sensitivity analysis was not performed, then check 'No'.	50. Do the authors justify the alternative values or ranges for sensitivity analysis? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
Check all that apply.	51. What methods were used to assess uncertainty? ☐ 1. One-way sensitivity analysis ☐ 2. Two-way sensitivity analysis ☐ 3. Multi-way sensitivity analysis ☐ 4. Bootstrapping or Monte Carlo simulations ☐ 5. Other (Specify: _____) ☐ 6. None performed
For a statement of funding support in the Acknowledgements, Title page or footnote, indicate 'Yes (explicitly stated)'.	**Conflict of Interest** 52. Does the paper present the relationship with the sponsor of the study? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
If the sponsor is a government granting agency or the research was performed independently (self-supported) then response is 'not applicable'. If nothing is indicated, check 'No'.	53. Does the paper indicate that the authors had independent control over the methods and right to publish? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No ☐ 4. Not applicable

	Conclusions 54. Is the answer to the study question provided? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
	55. Are the most important limitations of the study discussed? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
This relates to how the information derived from the study can be used for allocation decisions.	56. Do the authors generalize the conclusions to other settings or patient/client groups? ☐ 1. Yes (explicitly stated) ☐ 2. Yes (inferred from text, tables, or figures) ☐ 3. No
	57. Global impression of the quality of the paper. ☐ 1. Excellent ☐ 2. Very Good ☐ 3. Good ☐ 4. Fair ☐ 5. Poor ☐ 6. Worthless