Assessing the Quality of Medical Information Technology Economic Evaluations: Room for Improvement

Eric L. Eisenstein, D.B.A.[1], Maqui Ortiz[1], Kevin J. Anstrom, Ph.D.[1], David R. Crosslin, M.S. [1], and David F. Lobach, M.D., Ph.D., M.S[2].
[1]Duke Clinical Research Institute, Duke University Medical Center, Durham, NC
[2]Department of Community & Family Medicine, Duke University Medical Center, Durham, NC

Abstract
Medical information systems are being recognized for their ability to improve patient outcomes. While standards for the economic evaluation of medical technologies were instituted in the mid-1990s, little is known about their application in medical information technology studies. In a review of medical information technology evaluation studies published between 1982 and 2002, we found that the volume and variety of economic evaluations had increased; however, investigators routinely omitted key cost or effectiveness elements in their designs, resulting in publications with incomplete, and potentially biased, economic findings. Of the studies that made economic claims, 23% did not report any economic data, 40% failed to include any effectiveness measures, and more than 50% used a case study or pre- post- test design. Thus, during a time when health economic study methods in general have experienced significant development, there is little evidence of similar progress in medical information technology economic evaluations.

Introduction
Cochrane and Haynes have proposed four tests for the evaluation of medical technologies.[1, 2]

1. Should it work?
2. Can it work?
3. Does it work?
4. Is it worth it?

The first question examines the *theory* of the medical technology through an explicit statement of the mechanisms by which the technology is expected to alter patient outcomes (both clinical and economic). The answer to this question defines key parameters for subsequent empirical testing. The second question examines the *efficacy* of the medical technology; its answer determines whether the technology works under ideal circumstances (as observed within a clinical trial). The third question addresses the *effectiveness* of the medical technology and determines whether the technology works in usual circumstances (as encountered in actual practice). And, the fourth question addresses the *efficiency* of the medical technology by determining whether its *costs* are appropriate for its level of *effectiveness*.

The Institute of Medicine's report on the computer-based patient record and the recent recommendations from the Leapfrog Group pertaining to computerized provider order entry (CPOE) allow medical information technologies to significantly improve patient care effectiveness.[3, 4] Yet, even the Leapfrog Group admits that, "There exists little published information on the costs of implementing CPOE. Studies to date have been based principally on generic estimates or total costs cited by a handful of organizations."[5, 6] Given that the quality of economic evidence in support of this important medical technology is acknowledged to be inadequate, one could assume that the quality of economic evidence in support of other medical information technologies is no better.

Since the mid-1990s, a number of standards have been proposed for the economic evaluation of medical technologies. While Australian guidelines were specifically targeted to pharmaceuticals, Canadian and United Kingdom guidelines were designed to be used in the evaluation of all types of health care technologies.[7-9] Although there are no official US guidelines, the United States Public Health Service's Panel on Cost-Effectiveness in Health and Medicine presented their recommendations for the economic evaluation of medical technologies in 1995 and these have become the de facto standards in the US.[10] The British Medical Journal's guidelines for authors, with its checklist of items, have become the standard for evaluating publications in this area and form the basis for the Campbell and Cochrane Economic Methods Group's Policy Statement.[11, 12]

In this study, we sought to determine (1) whether efficiency information was being properly incorporated into evaluation studies; and (2) whether the quality of efficiency information included in these studies has changed substantially since the

publication of medical technology economic evaluation guidelines in the mid 1990s.

Methods

Study Sample: Ammenwerth and de Keizer conducted a systematic review of the medical information technology evaluation research accessible through PubMed that was published between 1982 and 2002.[13, 14] These studies were coded and entered into a public, online database.

Data Collection: A search was conducted in Ammenwerth and de Keizer's database for English language articles that made economic claims (e.g., cost-saving or cost-effective) for specific medical information technologies. Because of the database's simplified coding strategy, our search terms were restricted to 'cost' or 'economic' and 'English language'.

Two reviewers independently analyzed a 20% sample of full articles and determined that the economic information reported in abstracts was a plausible representation of the information contained in the full articles. The reviewers then independently coded abstracts of articles meeting our inclusion criteria, and resolved coding differences by discussion.

General characteristics were coded for all papers in our study; whereas, economic study characteristics (cost, effectiveness, or both) were only coded for papers that collected empirical data.

General Characteristics: General characteristics included in our study were the type of Economic Claim made as well as a Study Rating that differentiated between studies making rhetorical economic claims and those that presented at least some cost data. (Figure 1) We also categorized the type of Economic Finding, assuming that there would be a reporting bias in favor of studies with positive versus negative or mixed findings, and Study Methodology. Studies that used a model to present empirical data collected in the study were coded as empirical.

Economic Study Characteristics: For studies classified as empirical, we coded additional variables describing the economic study design and the economic and effectiveness measures included. With regard to study design, we distinguished between single- and multi-site studies; between prospective and retrospective perspectives; and between designs involving case studies (one study arm), pre and post testing (more than one sequential study arm), observational (more than one simultaneous study arm), and randomized controlled trials (subjects randomized to more than 1 information intervention).

Figure 1: Coded Study Characteristics

General Characteristics
Economic Claim: cost saving, cost effective, other
Study Rating: rhetorical, cost data included
Economic Finding: positive, negative, mixed
Alternatives: one study arm, more than one study arm
Study Method: review, model, empirical
Economic Study Characteristics
Study Framework
Sites: one, multiple
Perspective: prospective, retrospective
Design: case study, pre-post, observational, randomized controlled trial
Economic Measures
Included: yes, no
Reported as: resource use, unit prices
Primary economic data sources: measured in study, secondary data sources
Cost components: information intervention, patient care, other
Patient Effectiveness Measures
Included: yes, no
Reported as: outcome array, primary outcome, patient utilities, monetarized

Applying the Campbell and Cochrane Economic Methods Group (C&CEMG) definitions, we assumed that economic and effectiveness measures could each have two components.[12] The potential economic measures are resource use and unit prices (the valuations of resource use). The potential effectiveness measures are patient clinical outcomes and their valuations as patient utilities. If either of the economic measures were included in the design (even if they were incomplete), we coded economic measures as being present. We also coded the study's primary economic data source. When a design made use of both economic data collected within the study as well as secondary data sources, we coded the source for the study's primary economic analysis. Lastly, we coded the presence or absence of three cost components (medical intervention, patient care, and other) as defined by the US Public Health Service's guidelines for the conduct of economic analyses of medical technologies.[10] Again, we did not attempt to assess whether all costs were included. Rather, we merely noted whether at least one example of each cost component was included.

Following the C&CEMG definitions, we considered four effectiveness measures. The simplest option is to list an array of effectiveness measures with no

attempt made to derive the overall measure of an intervention's effectiveness. During reporting, this outcome array is presented alongside the costs. The second option occurs when the study has a primary outcome. If the outcome is reported without patient utilities, it is combined with the costs in a cost-effectiveness ratio. Third, when the effectiveness measure includes patient utilities, these are combined with costs in a cost-utility ratio. And fourth, when the effectiveness measure is translated into a monetary value, it can be combined with costs in a cost-benefit ratio. In this study only cost-savings and cost-effectiveness were coded individually. Cost of technology, cost-benefit and cost-utility were included as "other". Since our purpose was to code how effectiveness measures were reported, we did not assess whether the studies' reporting was correct.

Analyses: We divided the study period into three eras for analysis. The first era (1982-1988) coincides with the introduction and immediate aftermath of Medicare's Diagnosis Related Group (DRG) prospective payment system. During this time there was an increased interest in health economics in the US. The second era (1989-1995) signals the maturation of health economics as a discipline. The third era (1996-2002) begins with the publication of the US Public Health Service and BMJ guidelines, and continues to the early 21st century.[10,11] Within each era, we calculated the frequency of responses to our coded items.

Results

We begin by reporting general characteristics for all studies and then report economic study characteristics for the subset of empirical studies.

General Characteristics: The Ammenwerth and de Keizer database contained 1036 evaluation studies, 964 of which (93%) were published in English (Table 1). We identified 134 studies (14%) that made economic claims and formed the study sample for our analyses. While the number of economic studies increased throughout the three eras, the percent of English language studies making economic claims (14%) was relatively constant across all eras.

Table 1: Number of Economic Evaluation Studies by Era

	Evaluation Period			
	1982-88	**1989-95**	**1996-02**	**Total**
Total Studies	106	291	639	1036
English language	102	262	600	964
Economic claims	14	33	87	134
Rhetorical	3	8	20	31
Cost data	11	25	67	103

Similarly 23% of all economic studies made rhetorical claims without presenting cost data, and this percentage also did not vary substantially across eras.

Across all eras, cost-savings was the most frequently cited economic claim (48% of studies), while cost-effectiveness (15% of studies) was the least cited. As anticipated, most studies (66%) reported positive economic results; however, this percent declined after the publication of medical technology economic analysis guidelines in the mid-1990s. While 74% of all studies considered two or more alternatives, this percent declined in the third era. This coincides with a heightened interest in a few, high priced technologies (e.g., picture archiving communication systems (PACS) and telemedicine) and the publication of case studies in these areas. The use of empirical economic study designs remained high through all three eras (average 81%).

Table 2: Economic Study Characteristic Percents by Era

	Evaluation Periods		
	1982-88	**1989-95**	**1996-02**
Number of Studies (n)	14	33	87
Economic Claims (%)			
Cost-saving	50	61	43
Cost-effective	7	18	14
Other	43	21	43
Economic Findings (%)			
Positive	64	88	59
Negative	14	0	9
Other	21	12	32
Alternative Considered (%)			
Treatment only	14	12	21
Two or more arms	86	82	69
Not determined	0	6	10
Study Design (%)			
Review	0	3	13
Model	0	12	11
Empirical	100	85	76

The types of medical information systems evaluated in economic analyses changed considerably during our study period (Table 3). In the first era, ancillary departmental systems (e.g., pharmacy and laboratory medicine) were the most common subjects for economic evaluation; however, by the third era, telecommunication and PAC systems accounted for over half of all economic evaluations. Over the entire study period, telecommunication applications accounted for 28% of all economic studies, PACS 12%, expert, clinical guideline, and reminder systems 11%, and clinical information and documentation systems 10%. Clearly, these temporal shifts in emphasis for economic study applications reflect the changing information system priorities of health care organizations.

Empirical Study Characteristics: Throughout our study period, there was little evidence of evolution in economic study design (Table 4).

Table 3: HIT Economic Study Applications.
Information Systems Studied, by Percent and Era

	Evaluation Period		
Information System	**1982-88**	**1989-95**	**1996-02**
Telecommunication	0	3	41
Picture archiving and communication (PACS)	0	12	14
Expert system, clinical guideline	14	18	8
Clinical Information or documentation	21	18	6
Physician order entry, drug prescription	14	12	8
Pharmacy and laboratory medicine	36	18	2
General practitioner or primary care	14	9	6

Most studies were single center, prospective designs that did not randomized patients to different information interventions. Since many telecommunications

Table 4: Empirical Study Types by Percent and Era

	Evaluation Period		
	1982-88	**1989-95**	**1996-02**
Number of studies (n)	**14**	**28**	**66**
Study Sites (%)			
One	64	86	70
Multi	14	14	27
Unclear	21	0	3
Perspective (%)			
Prospective	79	82	74
Retrospective	0	14	14
Unclear	21	4	12
Study design (%)			
Case study	7	21	20
Pre-post	43	29	30
Observational	21	29	38
Randomized Trial	21	21	12
Unclear	8	0	0

applications involve two sites, the number of multi-center sites in the third era may be deceptively large. Interestingly, 50% of economic studies in each era were case studies or pre-post comparisons. Thus, at a time when randomized controlled trials are becoming common place in medicine, they remain relatively rare in medical information system evaluations.

The economic and effectiveness measures reported in medical information system economic evaluations changed little in our three eras (Table 5). Forty percent of studies in all eras did not include measurements of key resource use. This is important because resource use is more often comparable across sites, whereas costs will differ depending upon local prices. Thus, not reporting key resource use limits

Table 5: Economic Study Measurements.
Empirical Studies, by Percent and Era

	Evaluation Period		
	1982-88	**1989-95**	**1996-02**
Number of Studies (n)	**14**	**28**	**66**
Economic Measurements (%)			
Included in Study	71	82	83
Resource use reported	57	64	58
Cost Components:			
Information	29	18	36
Patient care	29	46	39
Other	29	29	15
Unclear	21	18	24
Effectiveness Measurements (%)			
Patient outcomes included	29	29	39
Effectiveness Reporting (%)			
Outcome array	29	21	26
Primary Outcome	0	7	14
Patient utilities	0	0	0
Monetarized	0	0	0

the ability to generalize study results. Most studies also did not report all three cost elements. Studies typically reported changes in one cost element (e.g., patient care costs) without including other elements (e.g., information intervention costs). If all three elements are not included, it is not possible to make a realistic overall assessment of the costs of an information intervention versus its alternatives. More surprising, less than half of economic evaluation studies reported changes in patient outcomes. Even if it is assumed that the information intervention will have no effect upon patient outcomes, it is still necessary to take effectiveness measurements and demonstrate that this is the case. The majority of studies that did report effectiveness measures used an outcome array, while no studies reported patient utilities (as recommended by various standards) and no studies attempted to place a monetary value on patient benefits.

Discussion:

In 1994, Tierney et. al cautioned that investments in informatics innovations, "must be balanced by attention to studying the costs and benefits."[15] A decade later, we find that despite the advancement of US and other standards for the economic evaluation of medical technologies, medical informatics investigators routinely ignore established economic guidelines in their studies. Thus, despite dramatic increases in the volume of medical information

system economic studies and radical changes in the application areas evaluated, we see little improvement in study design and in the type of economic and effectiveness information that is collected and reported.

Our study was limited in that we did not include publications after 2002, or perform a complete assessment of each full-length article. Nonetheless, our results show a consistent pattern: medical informatics investigators have not followed established standards for the economic evaluation of medical technologies.

The problems we identified are not unique to medical informatics.[16] While the lack of effective interventions and inadequate economic evaluation training are cited as impediments in other fields, efficacy is not as great a problem in medical informatics. We recommend that organizations such as AMIA encourage adherence to existing guidelines, and implement education programs for investigators and journal reviewers which are targeted to increasing the quality of economic information in medical information technology evaluation studies.

References:

1. Cochrane AL. Effectiveness and Efficiency: Random Reflection on Health Services. London: Nuffield Provincial Hospitals Trust; 1972.

2. Haynes B. Can it work? Does it work? Is it worth it? The testing of health care interventions is evolving. BMJ. 1999;319:652-3.

3. Computer physician order entry. The Leapfrog Group.http://www.leapfroggroup.org/for_hospitals/leapfrog_safety_practices/cpoe-Accessed March 12 2006.

4. Institute of Medicine Committee on Improving the Medical Record. The Computer-Based Patient Record: An Essential Technology for Health Care. Washington, D.C.: National Academy Press; 1991.

5. Birkmeyer CM, Lee J, Bates DW, Birkmeyer JD. Will electronic order entry reduce health care costs? Eff Clin Pract. 2002;5:67-74.

6. Computerized Physician Order Entry: Costs, Benefits and Challenges: A Case Study Approach. http://www.hospitalconnect.com/ahapolicyforum/resources/content/AHAFAHCPOEReportFINALrevised.pdf: First Consulting Group; 2003.

7. National Institute for Clinical Excellence. Guide to the Methods of Technology Appraisal. London: 2004.

8. Commonwealth of Australia Department of Health, Housing and Community Services. Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmacy Benefits Advisory Committee. 1995.

9. Canadian Coordinating Office of Health Technology Assessment. Guidelines For Economic Evaluation of Pharmaceuticals: Canada. Ottawa: CCOHTA; 1994.

10. Gold MR, Siegel J, Russell L, Weinstein M. Cost-Effectiveness in Health and Medicine. New York, NY: Oxford University Press; 1996.

11. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313:275-83.

12. Donaldson C, Mugford M, Vale L. Using systematic reviews in economic evaluation. Evidence-Based Health Economics. From Effectiveness to Efficiency in Systematic Review.London: 2002.

13. Ammenwerth E, de Keizer N. An inventory of evaluation studies of information technology in health care trends in evaluation research 1982-2002. Methods Inf Med. 2005;44:44-56.

14. Ammenwerth E, de Keizer N. Inventory of evaluation publications. http://evaldb umit at/index htm. 2006;Accessed March 15, 2006.

15. Tierney WM, Overhage JM, McDonald CJ. A plea for controlled trials in medical informatics. J Am Med Inform Assoc. 1994;1:353-5.

16. Levin HM. Waiting for Godot: cost-effectiveness analysis in education. In: Light RJ, editor. Evaluation Findings That Surprise.San Francisco: Josey-Bass; 2000.