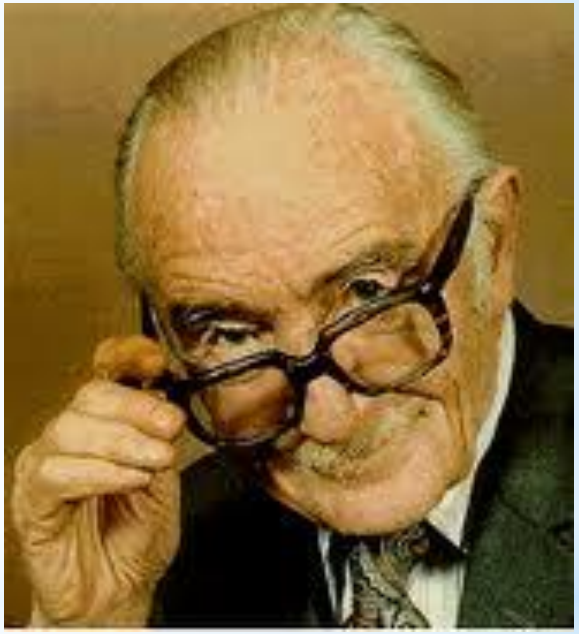
Systematic review

سعید اسلامی Pharm.D., Ph.D., Fellowship
فارماکوتراپی/انفورماتیک پزشکی/ آمار و روش شناسی تحقیق
فلوشیپ فوق تخصصی کیفیت، ایمنی و تصمیم یاری در (دارودرمانی) سالمندان/مراقبتهای ویژه
گروه داروسازی بالینی دانشکده داروسازی
گروه انفورماتیک پزشکی دانشکده پزشکی

Partly based on slides of: Dr. M. Pai, Division of Epidemiology, University of California
BU Alumni Medical Library

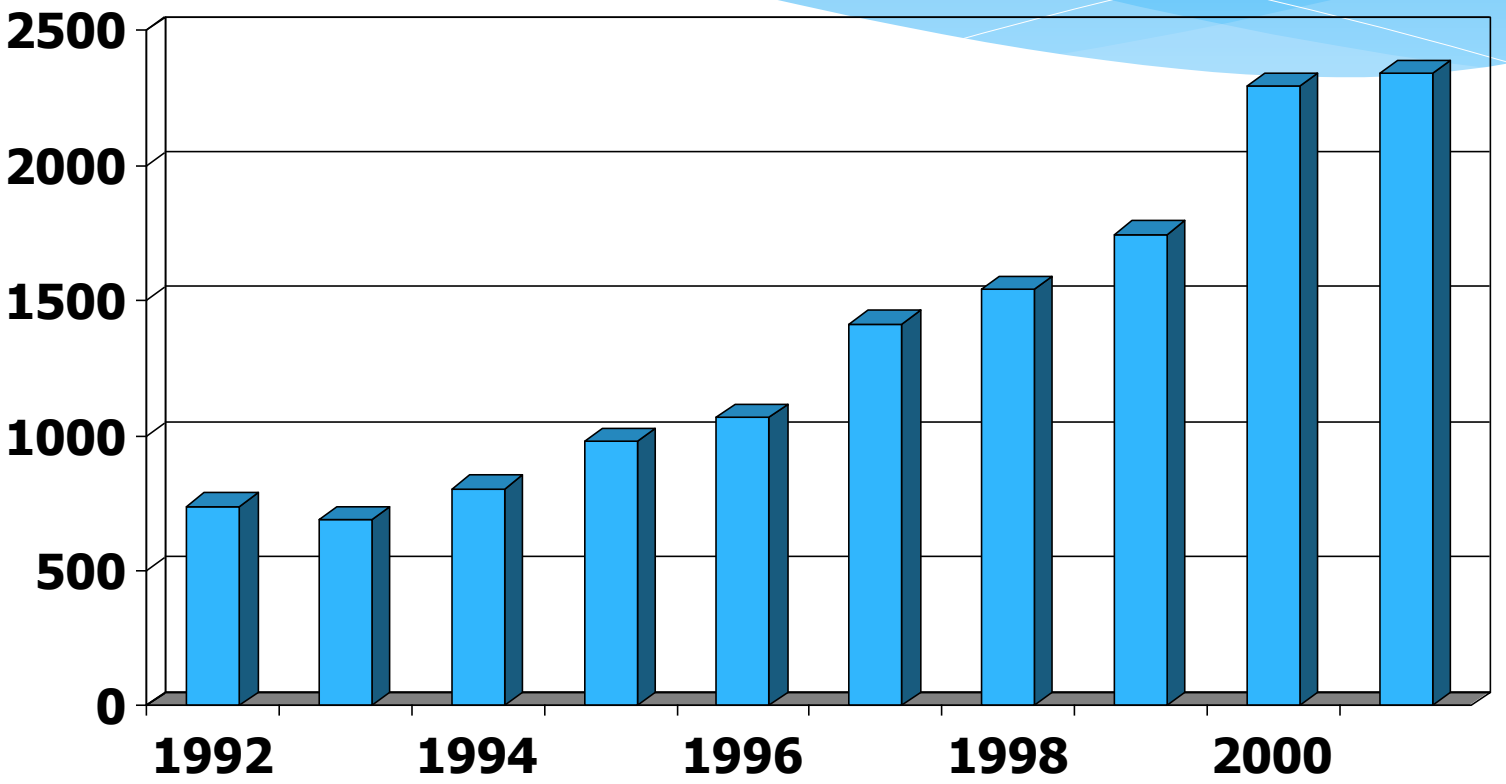
* Prof Archibald Cochrane, CBE (1909 - 1988)

* The Cochrane Collaboration is named in honour of Archie Cochrane, a British researcher.
* In 1979: *"It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomized controlled trials"*



Source: http://www.cochrane.org/cochrane/archieco.htm

Systematic reviews/meta-analyses indexed in PubMed – 10 years



2010: ~6500

Search: meta-analysis[MeSH] OR meta-analysis[ti] OR systematic review[ti]

Some definitions

* **Traditional, narrative reviews**, usually written by **experts** in the field, are **qualitative**, **narrative summaries of evidence** on a given topic. Typically, they involve informal and subjective methods to collect and interpret information.
* "**A systematic review** is a review in which there is a **comprehensive search** for relevant studies on a **specific topic**, and those **identified** as then **appraised** and synthesized according to a **predetermined** and explicit method."*

*Klassen et al. Guides for reading and interpreting systematic reviews. Arch Pediatr Adolesc Med 1998;152:700-704.

Some definitions

* "**A meta-analysis** is the **statistical combination** of at least 2 studies to produce a **single estimate of the effect** of the healthcare intervention under consideration."*
* **Individual patient data meta-analyses** (**pooled analyses**) involves obtaining **raw data** on all patients from each of the trials directly and then **re-analyzing** them.

*Klassen et al. Guides for reading and interpreting systematic reviews. Arch Pediatr Adolesc Med 1998;152:700-704.

Types of review articles

Elements of a Systematic Review

* Formulate the review question & write a protocol
* Search for and include primary studies
* Assess study quality
* Extract data
* Analyze data
* Interpret results & write a report

Examples

Review Paper ■

Evaluation of Outpatient Computerized Physician Medication Order Entry Systems: A Systematic Review

SAEID ESLAMI, PHARMD, AMEEN ABU-HANNA, PHD, NICOLETTE F. DE KEIZER, PHD

Abstract This paper provides a systematic literature review of CPOE evaluation stu
setting on: safety; cost and efficiency; adherence to guideline; alerts; time; and satisfaction, u
Thirty articles with original data (randomized clinical trial, non-randomized clinical trial, or
designs) met the inclusion criteria. Only four studies assessed the effect of CPOE on safety.
significant on the number of adverse drug events. Only one study showed a significant redu
of medication errors. Three studies showed significant reductions in medication costs; five c
support this. Most studies on adherence to guidelines showed a significant positive effect. T
number of evaluation studies published to date do not provide adequate evidence that CPC
safety and reduce cost in the outpatient settings. There is however evidence for (a) increasir
guidelines, (b) increasing total prescribing time, and (c) high frequency of ignored alerts.

■ **J Am Med Inform Assoc.** 2007;14:400–406. DOI 10.1197/jamia.M2238.

Introduction

A 1999 Institute of Medicine report estimated that about

trigger a series of actions, ultima
receiving the medication.[7]

Research **Open Access**

A systematic review on quality indicators for tight glycaemic control in critically ill patients: need for an unambiguous indicator reference subset

Saeid Eslami[1], Nicolette F de Keizer[1], Evert de Jonge[2], Marcus J Schultz[2] and Ameen Abu-Hanna[1]

[1]Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Meibergdreef, 1105 AZ Amsterdam, The Netherlands
[2]Department of Intensive Care, Academic Medical Center, University of Amsterdam, Meibergdreef, 1105 AZ Amsterdam, The Netherlands

Corresponding author: Saeid Eslami, s.eslami@amc.uva.nl

Received: 26 Aug 2008 Revisions requested: 6 Oct 2008 Revisions received: 14 Oct 2008 Accepted: 11 Nov 2008 Published: 11 Nov 2008

Critical Care 2008, **12**:R139 (doi:10.1186/cc7114)
This article is online at: http://ccforum.com/content/12/6/R139

Abstract

Introduction The objectives of this study were to systematically identify and summarize quality indicators of tight glycaemic control in critically ill patients, and to inspect the applicability of their definitions.

Methods We searched in MEDLINE® for all studies evaluating a tight glycaemic control protocol and/or quality of glucose control that reported original data from a clinical trial or observational study on critically ill adult patients.

Results Forty-nine studies met the inclusion criteria; 30 different indicators were extracted and categorized into four nonorthogonal categories: blood glucose zones (for example, 'hypoglycaemia'); blood glucose levels (for example, 'mean blood glucose level'); time intervals (for example, 'time to occurrence of an event'); and protocol characteristics (for example, 'blood glucose sampling frequency'). Hypoglycaemia-related indicators were used in 43 out of 49 studies, acting as a proxy for safety, but they employed many different definitions. Blood glucose level summaries were used in 41 out of 49 studies, reported as means and/or medians during the study period or at a certain time point (for example, the morning blood glucose level or blood glucose level upon starting insulin therapy). Time spent in the predefined blood glucose level range, time needed to reach the defined blood glucose level target, hyperglycaemia-related indicators and protocol-related indicators were other frequently used indicators. Most indicators differ in their definitions even when they are meant to measure the same underlying concept. More importantly, many definitions are not precise, prohibiting their applicability and hence the reproducibility and comparability of research results.

Conclusions An unambiguous indicator reference subset is necessary. The result of this systematic review can be used as a starting point from which to develop a standard list of well defined indicators that are associated with clinical outcomes or that concur with clinicians' subjective views on the quality of the regulatory process.

Introduction

and morbidity [6]. Attempts at achieving TGC, however, are

Hypothesis

"A systematic review should be based on principles of hypothesis testing, and the hypotheses must be conceived a priori."

Margaliot, Zvi, Kevin C. Chung. Systematic Reviews: A Primer for Plastic Surgery Research. PRS Journal. 120/7 (2007) p. 1836

and morbidity [6]. Attempts at achieving TGC, however, are not perfect and carry a risk for hypoglycaemia [4,5].

Several observational studies have reported on the quality of the glucose control process itself. The results and conclusions of these studies are contradictory [7]. Some show that the protocol prescribing the control process improves blood glucose control whereas others do not. Apart from differences in case-mix and in the associated therapy (for example, steroid therapy), two important issues hamper comparability between studies. The first impediment is the existing variability in the intervention's evaluation. The following interpretations, based

Focus of the Question

* The structured question will determine the inclusion and exclusion criteria:
  * What is the population of interest?
  * What are the interventions?
  * What are the outcomes of interest?
  * What study designs are appropriate?

in this paper concerns this second impediment.

The objective of the present systematic review is to identify and summarize quality indicators for glucose control in published studies of TGC in critically ill patients. It also assesses the applicability of definitions of quality indicators and organizes the indicators into categories. This review may form a basis for future developments of a standard list of well defined indicators that may correlate with clinical outcomes or that reflect clinicians' intuition regarding the quality of a given regulatory process.

Materials and methods

We searched for relevant English language articles

although these projections may have been somewhat overestimated.[17,18]

The main objective of this systematic review was to identify and summarize published studies of outpatient CPOE systems that evaluated one of six aspects: safety; cost and efficiency; adherence to guidelines; alerts; time; and satisfaction, usage, and usability.

For the purposes of this review, an outpatient CPOE system is a computer-based system that allows clinicians to enter medication orders directly for outpatients or primary care patients. In this context, a decision support system (DSS) is

Inclusion/Exclusion Criteria

* "Once the study question is formalized, the authors must compose a comprehensive list of inclusion and exclusion criteria."
* "To avoid selection bias, inclusion and exclusion criteria should be agreed upon and formalized before data extraction and analysis."

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1836

Two reviewers independently examined all titles and abstracts. Discrepancies between the two reviewers were resolved by consensus involving a third reviewer. Articles were selected if they reported original data from a clinical trial or observational study conducted in critically ill adult patients, and only if one of their main objectives concerned the evaluation of quality of TGC, with or without implementing an explicitly specified protocol. A study was defined as evaluating a TGC protocol if the (implicit or explicit) protocol implied an upper target range. Adherence to the protocol did not influence whether the study was included. Opinion papers, surveys and letters were excluded. Studies employing glucose-insulin-potassium protocols were excluded because they are not originally designed to achieve TGC.

Literature Search

* “A comprehensive and reproducible literature search is the foundation of a systematic review.”

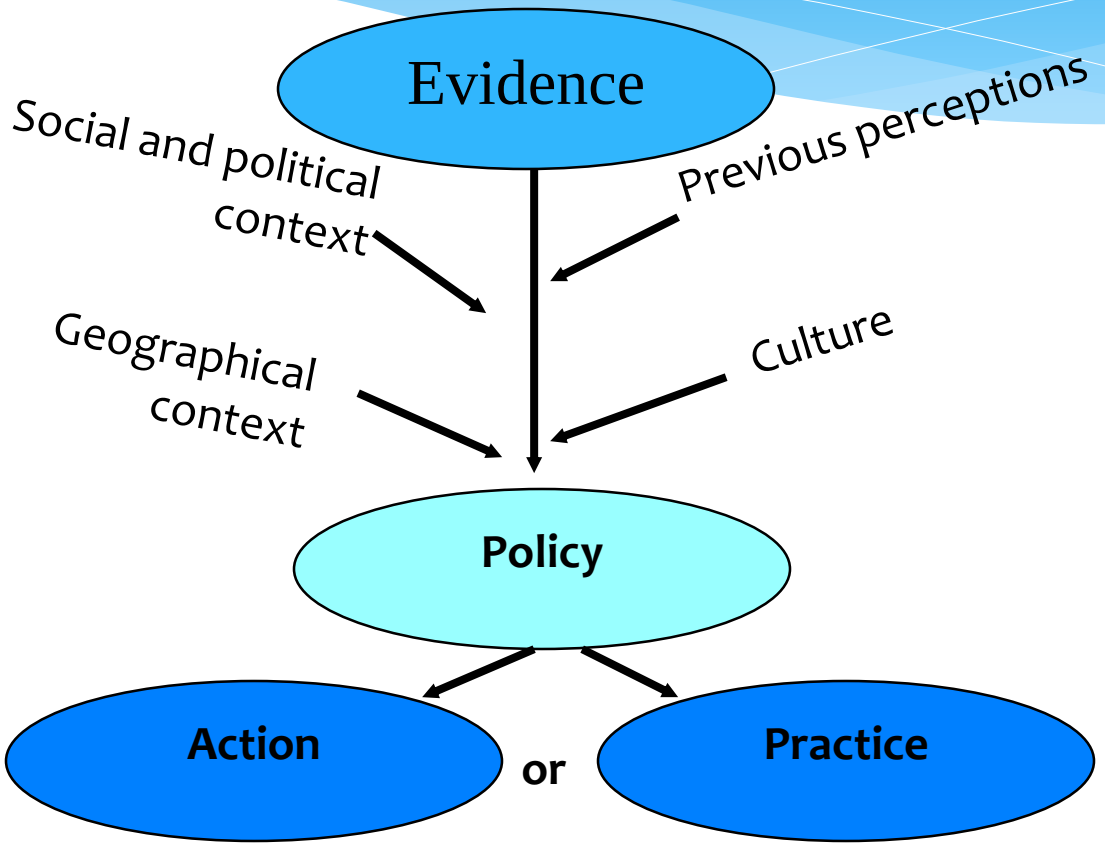
Margaliot, Zvi, Kevin C. Chung. “Systematic Reviews: A Primer for Plastic Surgery Research.” PRS Journal. 120/7 (2007) p.1837

Literature Search Challenges

* **Database Bias** - "No single database is likely to contain all published studies on a given subject."
* **Publication Bias** - selective publication of articles that show positive treatment of effects and statistical significance.
  * Hence, it is important to search for unpublished studies through a manual search of conference proceedings, correspondence with experts, and a search of clinical trials registries.

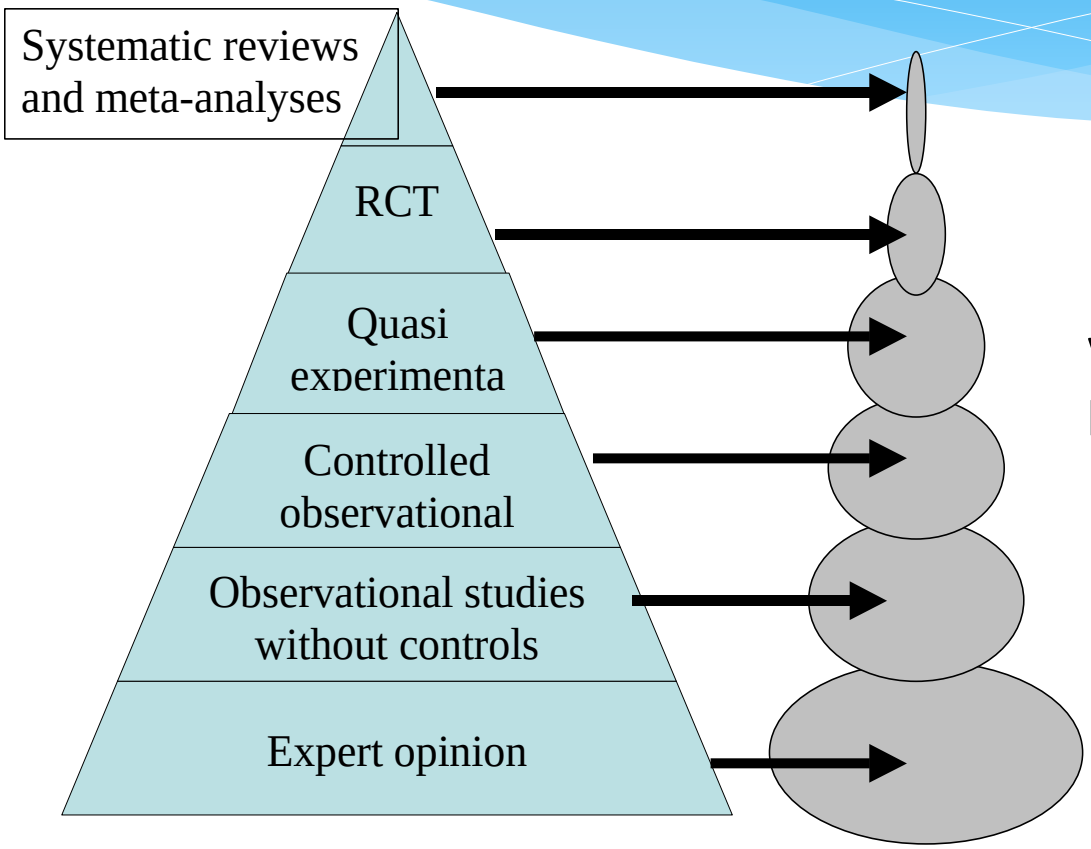
Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1837

Evidence to Policy



Levels of evidence & volume of grey literature

Levels of evidence



Volume of grey literature

Literature Review Challenges (cont.)

* **English-language bias** - occurs when reviewers exclude papers published in languages other than English
* **Citation bias** - occurs when studies with significant or positive results are referenced in other publications, compared with studies with inconclusive or negative findings

Materials and methods

We searched for relevant English language articles based on keywords in title, abstract and MeSH terms, using Ovid MEDLINE® and Ovid MEDLINE® In-Process (1950 to 31 December 2007). The final literature search was performed on 31 December 2007.

The following search strategy was used to identify the relevant articles. In the first stage we searched for 'glucose' and 'insulin'. In the second stage we limited the search using 'critical illness', 'critical care' or 'intensive care'. The results of these two stages were combined using the Boolean operator 'and'. Searching was supplemented by scanning the bibliographies of the identified articles.

A
1. Computerized
2. Computer Assisted
3. Medication Systems
4. Computer Program
5. Software
6. Medical Records Systems, Computerized
7. Electronic Medical Record
8. Ambulatory Care Information Systems
9. Drug Therapy, Computer-Assisted
10. Medical Order Entry Systems
11. Electronic Physician Order Entry
12. Electronic prescribing
13. Electronic prescription
14. Electronic order entry
15. Computerized Physician Order Entry (CPOE)
16. Computerized Prescriber Order Entry
17. Computerized Provider Order Entry
18. Computerized order entry
19. Decision Supports Systems, Clinical
20. 1 or 2 or ~ or 19

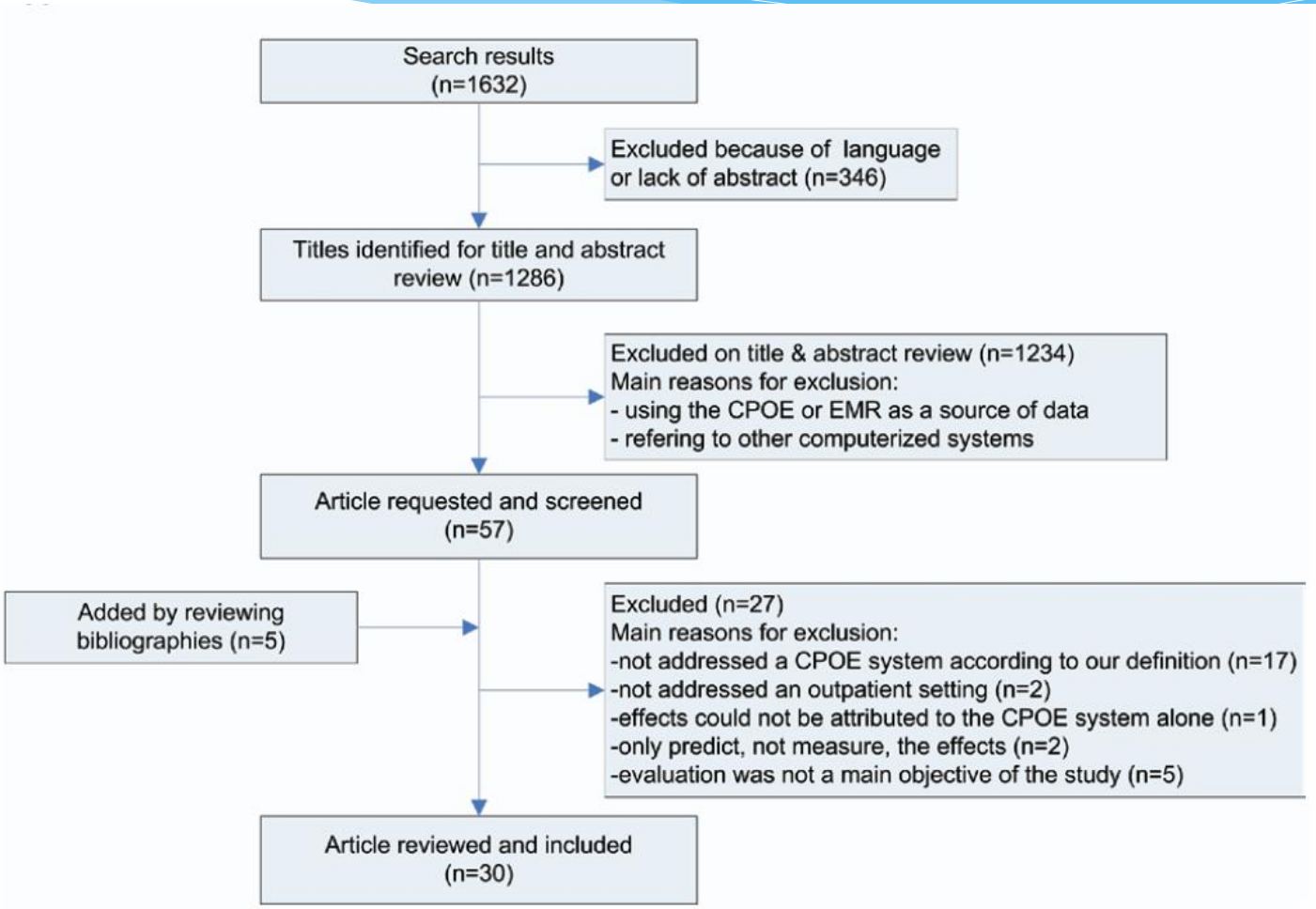
B
1. Drugs
2. Prescribing
3. Prescriptions
4. Medication
5. Drug Therapy
6. Treatment Planning
7. 1 or 2 or 3 or 4 or 5 or 6

C
1. Outpatients
2. Outpatient Care
3. Outpatient Clinics
4. Primary Medical Care
5. Primary Health Care
6. Primary Care
7. Ambulatory Care
8. Physicians, Family
9. Family Practice
10. General Practitioners (GPs)
11. 1 or 2 or ~ or 9

Search query: A.20 and B.7 and C.11

Searching was supplemented by scanning the bibliographies of the identified articles.

Two reviewers independently examined all titles and abstracts. Discrepancies between the two reviewers were resolved by consensus involving a third reviewer. Articles were selected if they reported original data from a clinical trial or observational study conducted in critically ill adult patients, and only if one of


Search results
(n=1632)
Excluded because of language
or lack of abstract (n=346)
Titles identified for title and abstract
review (n=1286)
Excluded on title & abstract review (n=1234)
Main reasons for exclusion:
- using the CPOE or EMR as a source of data
- refering to other computerized systems
Article requested and screened
(n=57)
Added by reviewing
bibliographies (n=5)
Excluded (n=27)
Main reasons for exclusion:
-not addressed a CPOE system according to our definition (n=17)
-not addressed an outpatient setting (n=2)
-effects could not be attributed to the CPOE system alone (n=1)
-only predict, not measure, the effects (n=2)
-evaluation was not a main objective of the study (n=5)
Article reviewed and included
(n=30)

Data Collection

* “The list of data to be extracted should be agreed upon a priori consensus during the design stage of the study.”

Margaliot, Zvi, Kevin C. Chung. “Systematic Reviews: A Primer for Plastic Surgery Research.” PRS Journal. 120/7 (2007) p.1839

Data Collection (cont.)

* Collected data includes:
  * Study characteristics
  * Sample demographics
  * Outcome data

Data Collection (cont.)

* "It is necessary to design a review-specific data **extraction form,** so that the same data are extracted from each study and **missing data** are clearly apparent."

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1839

Data Collection (cont.)

* "To ensure that data extraction is accurate and reproducible, it should be performed by at least **two independent readers**."

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1839

From the selected papers, the same two reviewers extracted data on TGC quality indicators (their definition and applicability). A quality indicator was defined as a measurable quantity of the TGC process that may, alone or in combination with other quantities, indicate some aspect of its quality. This includes, for example, mean (or median) BGLs as well as interpretations thereof in terms of counts of hyperglycaemic events. Discrepancies between the two reviewers were again resolved by consensus after involving the same third reviewer. We then attempted to categorize the quality indicators into coherent categories that capture their essence.

Results

Searching the online databases yielded 486 articles. Initial

Quality Assessment

* "The validity of a systematic review ultimately depends on the scientific method of the retrieved studies and the reporting of data."

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1839

Quality Assessment (cont.)

* Randomized Controlled Trials (RCT):
  * RCT are considered to be more rigorous than observational studies
  * A review based on well-designed RCT will likely be more valid and accurate than a review based on observational studies or case reports
  * However … !

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1839

Quality Assessment (cont.)

* "The most common way to assess and report study quality has been using a composite, numerical scoring instrument."

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1839

Quality Assessment (cont.)

* "More than 35 different quality assessment instruments have been published in the literature, and most are designed for randomized clinical trials."

Margaliot, Zvi, Kevin C. Chung. "Systematic Reviews: A Primer for Plastic Surgery Research." PRS Journal. 120/7 (2007) p.1839

Jadad score

* **Randomization (2 points possible)**
  * 1 point if study described as randomized
  * Add 1 point if randomization method described and appropriate (e.g. random numbers generated)
  * Deduct 1 point randomization described and inappropriate
* **Double-blinding (2 points possible)**
  * 1 point if study described as double-blinded
  * Add 1 point if method of double-blinding described and appropriate
  * Deduct 1 point if double-blinding described and inappropriate
* **Withdrawals (1 point possible)**
  * Give 1 point for a description of withdrawals and drop-outs

Linda N. Meurer, MD, MPH Department of Family and Community Medicine. "Systematic Synthesis of the Literature: Introduction to Meta-analysis". Power Point Presentation.

Jadad Score Example

Study	Randomization	Blinding	Drop-out
1	++	+	++
2	+	++	0
3	++	0	+
4	+	++	++
5	0	++	+

Data Synthesis

* “Once the data have been extracted and their quality and validity assessed, the outcomes of individual studies within a systematic review may be pooled and presented as summary outcome or effect”

Margaliot, Zvi, Kevin C. Chung. “Systematic Reviews: A Primer for Plastic Surgery Research.” PRS Journal. 120/7 (2007) p.1840

Data Synthesis (cont.)

* The authors summarize heterogeneous data qualitatively
  * “Data that are very conflicting and widely variable should not, under most circumstances, be combined numerically.”

Margaliot, Zvi, Kevin C. Chung. “Systematic Reviews: A Primer for Plastic Surgery Research.” PRS Journal. 120/7 (2007) p.1840

Table 3 ■ Frequency of All Selected Articles According to Outcome Categories and Study Design

Outcome Category	Specific Study Design	Study Designs with Control: Total	Positive Effect: Demonstrated	Positive Effect: Statistically	Positive Effect: Mix	No Effect	Negative Effect: Demonstrated	Negative Effect: Statistically	Negative Effect: Mix	Mixed Effect	ODWC*
Safety	RCT	1				1[21]					
	non-RCT	1			1[22]						
	OWC	2	1[19]			1[20]					
Cost and Efficiency	RCT	4				3[21,28,33]			1[29]		
	non-RCT	5		2[31,25]		1[23]	1[24]			1[30]	
	OWC	3	1[32]	2[26,27]							
Adherence to Guideline	RCT	4			1[37]	3[28,29,33]					
	non-RCT	4		3[25,35,38]		1[30]					
	OWC	3	2[34,36]			1[32]					
Alert	RCT										6[19,39–43]
	non-RCT										
	OWC										
Time	RCT	1						1[44]			1[46]
	non-RCT	2						2[31,45]			
	OWC										
Satisfaction, Usage and Usability	RCT	3				2[28,29]		1[21]			9[25,30,32,33,36,43,44,46,48]
	non-RCT	2					1[23]	1[35]			
	OWC	1	1[47]								
Total	RCT				1	9		2	1		16
	non-RCT			5	1	2	2	3		1	
	OWC		5	2		2					

ODWC = Observation design without control; OWC = observational with controls.
Controlled Studies are also described by the effects of CPOE.

of the TGC process that may, alone or in combination with other quantities, indicate some aspect of its quality. This includes, for example, mean (or median) BGLs as well as interpretations thereof in terms of counts of hyperglycaemic events. Discrepancies between the two reviewers were again resolved by consensus after involving the same third reviewer. We then attempted to categorize the quality indicators into coherent categories that capture their essence.

Results

Searching the online databases yielded 486 articles. Initial screening of titles and abstracts resulted in 50 articles eligible

Table 3

Hypoglycaemia quality indicator subgroups

Indicators	Description	Reference
Hypoglycaemic events	Reported thresholds[a] for defining a BGL as hypoglycaemic event: • <40 [8,10,17-19,30,35,43]; • ≤ 40 [26]; • <45 [44,54]; • <50 [31,49,50]; • <54 [34]; • <57 [32]; • <60 [8,11,45,52,55]; • ≤ 60 [16,21]; • <63 [23]; • <65 [25]; • <70 [9,17,22,27,28,57]; • <72 [20]; or • threshold was not reported [48] Represented as: • percentage and number of measurements and/or patients with hypoglycaemic event during the trial, or normalized for duration of therapy [16]; • mean or median of events per patient-day [57]; • mean of events per patient [27]; • patients with at least one event per day [22]; or • in three studies, hypoglycaemic events did not occur and therefore were not reported [32,48,54]	31 articles [8-11,16-23,25-28,30-32,34,35,43-45,48-50,52,54,55,57]
Severe or marked hypoglycaemic events	Reported threshold for defining a BGL as severe hypoglycaemic event: • <40 [15,20,25,29,33,42,46,47,51,56]; • ≤ 40 [12,13,16];	15 articles [8,12,13,15,16,20,24,25,29,33,42,46,47,51,56]

Discussion

We have identified, listed and categorized TGC quality indicators, as used in 49 studies. In our search for studies pertaining to TGC, we allowed any synonym, without limiting the search *a priori*. A limitation of our search is that we addressed only studies in which evaluation and quality measurement formed a main objective; we might therefore have missed some studies with a limited evaluation and quality measurement focus. In addition, frequency was used as the ordering principle for presenting and describing indicators. Although this approach provides a good overview of the popular indicators used, it may overlook less frequent but useful indicators. Finally, although indicator categories are useful in terms of managing and understanding indicators, their induction is subjective. One may for example also consider the complexity of calculation of indicators (for instance, calculating mean BGL is simpler and faster than time-weighted mean BGL).

To our knowledge this is the first review dedicated exclusively to quality indicators for TGC in critically ill patients. Existing reviews on TGC have focused on its effects [7,37]; evidence of its utility and its advantages were reported, and ways to implement TGC protocols successfully discussed.

Discussion and Recommendations

We have identified and described the results of 30 papers on evaluation of CPOE systems in outpatients. The number of such evaluation studies has clearly increased since 2002 (only 10 articles out of 30 before 2002). We used extensive search criteria in order to capture the different ways a CPOE system is referred to in the published literature. However, the following are two limitations of our search. First, because we only addressed studies in which evaluation formed a main objective, we could have missed some studies with a limited evaluation focus. Second, we may have missed some studies

Systematic reviews => meta-analyses

* Heterogeneity
* Mixing high and poor quality studies
* When reporting biases are a problem
  * Publication bias
  * Time lag bias
  * Duplicate publication bias
  * Language bias
  * Outcome reporting bias