Performance measurements in diabetes care: the complex task of selecting quality indicators

HISKE CALSBEEK, NICOLE A.B.M. KETELAAR, MARJAN J. FABER, MICHEL WENSING AND JOZÉ BRASPENNING

Scientific Institute for Quality of Healthcare (IQ healthcare), Radboud University Medical Center, Nijmegen, The Netherlands
Address reprint requests to: Hiske Calsbeek, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands. Tel: +31-24-3610329; Fax: +31-24-3540166; E-mail: hiske.calsbeek@radboudumc.nl
Accepted for publication 22 September 2013

Abstract

Purpose. To review the literature on the content and development of the sets of quality indicators used in studies on the quality of diabetes care in primary care settings.

Data sources. The MEDLINE (Ovid), PubMed, PsychINFO, Embase and CINAHL databases were searched for relevant articles published up to January 2011.

Study selection and data extraction. We included studies on the quality of adult diabetes care, using quality indicators. We excluded studies focusing on the hospital setting, patient subgroups, specific components of diabetes care and specific outcomes. In total, 102 studies (including 102 sets and 1494 indicators) were analyzed by two independent reviewers, using the criteria of the National Quality Measures Clearinghouse and international guidelines to document the content and selection of the identified indicators.

Results of data synthesis. Sets varied greatly in number, content and definitions of quality indicators. Most of the indicators concerned HbA1C, lipids, blood pressure, eye and foot examination and urinalysis. Few sets included indicators on lifestyle counseling, patient experiences, healthcare structure or access to healthcare providers. Seventy sets did not specify explicit selection criteria, and 19 of these did not report the sources of the indicators.

Conclusions. Sets of quality indicators are diverse in number, content and definitions. This diversity reflects a lack of uniformity in the concept of diabetes care quality and hinders the interpretation of and comparison between quality assessments. Methodology regarding defining constructs such as the quality of diabetes care and indicator selection procedures is available and should be used more rigorously.

Keywords: quality of care, indicators, diabetes mellitus, primary care, systematic review

Introduction

Given the enormous impact that indicator scores may have on contracting, pay for performance, public reporting and quality improvement [1–3], it is crucial that quality indicators are comprehensive and based on sound methods. Regarding the content covered, clinical quality indicators can be grouped into measure domains, such as process, access, outcome, structure and patient experience [4]. In addition, indicator sets for diabetes care should cover the clinical domains of prevention or delay of type 2 diabetes, diagnostics, monitoring, treatment and integrated care [5–7]. The National Quality Measures Clearinghouse (NQMC) [8] provides guidance that addresses the indicator selection considerations that need to be taken into account during the indicator selection process to support the construction of a tailored set of quality indicators.

Sets of quality indicators are often derived from various authoritative sources and used to report on the process or outcomes of diabetes care in order to assess and monitor the quality of care. For instance, the Agency for Healthcare Research and Quality [9] lists 72 diabetes-related indicators, including those from the Healthcare Effectiveness Data & Information Set (HEDIS, 11 indicators) [10], the Veterans Health Administration (11 indicators) [11] and the Quality and Outcomes Framework Indicators (QOF, 8 indicators) [12]. Some indicators are incorporated in all of the sets, e.g. indicators regarding HbA1c levels or blood pressure, while other indicators appear in only a few sets, e.g. screening for depression or neuropathy testing. It is unknown whether all of the different sets represent a different view or a different definition of the quality of diabetes care.

In contrast to the studies that use clinical parameters or measures of health-related quality of life, little seems to have been
reported on the systematic procedures for the selection of quality of diabetes care measures, i.e. quality indicators, in studies measuring the quality of diabetes care. Therefore, we undertook a systematic literature review to provide insight into the extent of diversity in the operationalization of the concept of diabetes care quality. We limited the scope to adult diabetes care in primary care settings, as most diabetic patients are diagnosed, treated and monitored in primary care settings. The key objectives of this review were the following: (i) to examine and compare the content covered by the applied sets of quality indicators in terms of measure domains and clinical domains and (ii) to examine the methods used for defining a set of quality indicators.

**Methods**

**Data sources and searches**

We performed a systematic review of international quality indicator sets for diabetes care that were used in peer-reviewed and indexed scientific literature. The literature searches used (the related terms and the combination of) the terms diabetes mellitus, quality indicator, report card, public reporting, standards, quality of care, performance measurement and benchmarking in the MEDLINE (Ovid), PubMed, PsychnFO, Embase and CINAHL databases (the complete search strategy is available on request) to identify indicator sets that were used in empirical studies on the quality of diabetes care in primary care settings. All of the publication dates up to January 2011 were accessed on 25 August 2011.

**Study selection**

Our search strategy initially resulted in 3013 hits. The titles and abstracts of all papers identified were screened by H.C. and N.K. using the following inclusion criteria: the study aimed to determine the quality of adult diabetes care in primary care settings using a set of at least two quality indicators and was published in English or Dutch. The exclusion criteria were the following: a hospital setting, the use of patient enrollees of a health-care organization or clinician related to the capacity to provide high-quality health care. We limited the scope to adult diabetes care in primary care settings, as most diabetic patients are diagnosed, treated and monitored in primary care settings. The key objectives of this review were the following: (i) to examine and compare the content covered by the applied sets of quality indicators in terms of measure domains and clinical domains and (ii) to examine the methods used for defining a set of quality indicators.

In the included studies resulted in 16 supplementary papers. In total, we included 102 papers, i.e. 102 sets of quality indicators.

**Data extraction and synthesis**

The included papers were screened by H.C. and N.K. to extract the following characteristics: design, publication year, country of origin, origin of sets and quality indicators, selection methods used (yes/no) and number of quality indicators.

The indicator sets were screened for content based on measure domains and clinical domains (Box 1). To guide this process, we used the definitions of the measure domains of the NQMC [4]. In line with Sidorenkov et al. [13] we distinguished intermediate or surrogate outcomes, such as the HbA1c level or blood pressure and distal or hard outcomes, such as blindness or amputation. To frame the clinical domains of diabetes care, we used evidence-based guidelines for diabetes care [5–7].

Finally, the indicators that measured the same topic were clustered to identify the most frequently used quality indicators. In case of difficulties in identifying and classifying the individual indicators, we discussed the paper with a third reviewer (J.B.).

We used the tutorial of the NQMC [8] to consider the selection procedures used (Box 2). It must be mentioned that the selection criteria in Box 2 show some overlap. For example, if the criterion ‘data availability’ was reported in a study, this was scored as one selection criterion: ‘Availability of data sources’. If apart from ‘data availability’, ‘validity’ was mentioned, we scored two selection criteria, namely ‘Availability of data sources’ and ‘Desirable attributes’.

Descriptive analyses were performed using Access 2007 (Microsoft Office). To examine changes over time, we divided the publication years into three periods: 1993–2000, 2001–8
and 2009–10. The distinction between 2008 and 2009 was based on emerging results of pay-for-performance programs as from 2009.

## Results

### Included studies

Most of the studies were conducted in the USA (n = 46), UK (n = 16) or other European countries (n = 18) (Supplementary data, table). The studies, mostly cross-sectional (31%), retrospective (30%), or clinical trials (14%), increased in number over time: between 1993 and 2000 7 studies, between 2001 and 2009 49 studies and between 2009 and 2010 46 studies. The sets contained 1494 indicators in total; the number of quality indicators per set varied between 3 and 57 (median 14).

### Origin of sets and indicators

While most of the sets (n = 71) were derived from available sets developed by authoritative agencies, 5 sets were self-developed and 26 sets did not identify a source (Supplementary data, table). In total, 24 different sources were identified. In line with the number of studies, most of the sources originated in the USA from groups such as the American Diabetes Association (ADA) (in 25 studies), the National Diabetes Quality Improvement Alliance or Diabetes Quality Improvement Project (in 11 studies) and HEDIS (in 10 studies), or originated in the UK from groups such as the QOF (in 9 studies). The number of indicators selected from these sources varied widely, e.g. between 4 and 22 indicators from the ADA, while the current Standards of Medical Care in Diabetes include over 100 recommendations (exclusive of diabetes care in specific populations). In other countries, the quality indicators were generally derived from the sets available from the national diabetes organizations, such as the Associazione Medici Diabetologi in Italy. It must be noted that the available agency-developed sets overlap with each other, often using the same guidelines, e.g. the ADA Standards of Medical Care in Diabetes.

### Quality indicator content

The content covered by the quality indicators in terms of measure domain and clinical domain showed little variation (Figs 1 and 2). Most of the indicators (60%) concerned the process of care, included in 95% of the sets (Fig. 1). Thirty-six percent of the indicators were outcomes of care, with primarily intermediate outcomes. Indicators for the structure of care or patient experiences were found in six sets, covering 2 and 1%, respectively, of all indicators. Access to care has hardly been

---

**Box 2: Selection criteria (according to the NQMC [8])**

**Desirable attributes:**

- Importance (relevance to stakeholders, health importance, applicable to measuring the equitable distribution of health care, potential for improvement, susceptibility to being influenced by the health care system)
- Scientific soundness (clinical logic: explicitness of evidence, strength of evidence)
- Scientific soundness (measure properties: reliability, validity, allowance for stratification or case-mix adjustment, comprehensibility)
- Feasibility (explicit specification of numerator and denominator, data availability)
- Availability of data sources
- Application to desired setting or care
- Selection from appropriate domain
- Considerations for comparisons

---

Figure 1 The content of the applied quality indicators stratified by measure domain (102 sets comprising 1494 indicators in total), %. 1The total percentage of sets is >100% because most of the sets comprise different types of indicators. The total percentage of indicators (outcome total, process, structure, access and patient experience) is 100%.

Figure 2 The content of the applied quality indicators stratified by clinical domain (102 sets comprising 1494 indicators in total), %. 1The total percentage of sets is >100% because most of the sets comprise different types of indicators. The total percentage of indicators is 100%.
measured. Figure 2 shows that all of the sets included monitoring indicators. Indicators for treatment were a part of the set in 60% of the studies. Few indicators were found for integrated care, while prevention and diagnostics were not represented in any indicator set.

**Most frequently used quality indicators**

Categorizing the indicators for the same topic resulted in 75 unique indicators or indicator categories. Table 1 shows the 10 most frequently used quality indicators: outcome and process indicators for blood glucose, lipids and blood pressure and process indicators for eye and foot examinations, urinalysis and lifestyle counseling. None of these indicators were used in all of the studies. Most of the sets (83%) contained outcome indicators for blood glucose. Process indicators for lifestyle counseling, which are listed in the 10th position in Table 1, appeared to be used in less than one-third of the studies (29%). In addition, as from 2001, we noticed a shift in the ranking of more outcome indicators upwards to the top of the 10 most frequently used indicators.

The classification of 1494 quality indicators into 75 categories implies much variation among the definitions of quality indicators. Table 1 shows that many sets contain more than one quality indicator for the same topic, a situation that results from different definitions (e.g. regarding measurement periods or choosing negative or positive cut-off points). The following examples illustrate this variation: ‘HbA1c poorly controlled (%)’, ‘HbA1c mean value (%)’, ‘HbA1c attaining target level (%)’ and ‘HbA1c ≤7% (%)’.

**Systematic use of selection criteria**

Most of the studies (n = 70) did not refer to any selection criteria or procedure used to define the composition of the indicator sets (Supplementary data, table). Fifty of these studies mentioned that their indicators were derived from an available set (e.g. HEDIS or QOF) but did not provide justification for why these specific indicators were chosen. The 32 studies that presented selection criteria were published since 2001, with no increase over time. These studies mostly cited one or two ‘desirable attributes’ of the indicators, such as scientific soundness (evidence based), health importance or data availability. The selection criteria and the source of the quality indicators were not reported in 19 studies.

**Discussion**

This review of the applied sets of diabetes quality indicators in the scientific literature revealed variation in the number of indicators, the content covered and the description of indicators. The number of quality indicators per set varied widely (from 3 to 57), exhibiting a median of 14 indicators. Most of the sets included indicators on the measure domains of care processes and outcomes, while patient experiences, the structure of care and access were seldom or not represented in the sets. Regarding the clinical domains, we found that most of the sets included indicators on monitoring and treatment. However, integrated care, prevention and diagnostics were hardly represented. In addition, while most sets covered internationally accepted evidence-based topics, such as blood glucose, lipids and blood pressure, none of these indicators were included in all of the sets. Indicators for lifestyle and self-management counseling were included in less than one-third of the sets. We found substantial differences in the descriptions of the indicators that measured the same topic.

The most obvious explanation for this variation relates to the intended use of the indicators, as the first step in selecting indicators is the identification of the measurement purpose [8]. According to the NQMC, research is one of the three general purposes (next to quality improvement and accountability) that is often conducted to ‘evaluate programs and assess the effect of policy changes on health care quality’. However, when one considers that most of the included studies actually had comparable research goals and that we only searched for sets that were

<table>
<thead>
<tr>
<th>Quality indicator</th>
<th>Used in number of sets (%) (n = 102)</th>
<th>n indicators (%) (n = 1494)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Blood glucose outcome</td>
<td>85 (83)</td>
<td>150 (10)</td>
</tr>
<tr>
<td>2. Cholesterol measurement (process)</td>
<td>84 (82)</td>
<td>111 (7)</td>
</tr>
<tr>
<td>3. Blood glucose measurement (process)</td>
<td>84 (82)</td>
<td>110 (7)</td>
</tr>
<tr>
<td>4. Cholesterol outcome</td>
<td>75 (74)</td>
<td>155 (10)</td>
</tr>
<tr>
<td>5. Blood pressure outcome</td>
<td>73 (72)</td>
<td>148 (10)</td>
</tr>
<tr>
<td>6. Eye examination (process)</td>
<td>72 (71)</td>
<td>81 (5)</td>
</tr>
<tr>
<td>7. Urine test (process)</td>
<td>66 (65)</td>
<td>92 (6)</td>
</tr>
<tr>
<td>8. Foot examination (process)</td>
<td>62 (61)</td>
<td>77 (5)</td>
</tr>
<tr>
<td>9. Blood pressure measurement (process)</td>
<td>53 (52)</td>
<td>58 (4)</td>
</tr>
<tr>
<td>10. Counseling on lifestyle and self-management (process)</td>
<td>30 (29)</td>
<td>71 (5)</td>
</tr>
</tbody>
</table>

*Indicators were categorized into 75 categories because several definitions of the same topic exist.

The number of indicators exceeds the number of sets because sets often contain more than one indicator for the same topic due to different definitions, e.g. regarding measurement periods in process measures or cut-off points in outcome measures.
used for the operationalization of the overall quality of diabetes care, this explanation is unlikely. A more likely explanation is the different sources from which most indicators were obtained (such as the HEDIS, ADA or QOF). Despite some overlap in content, these sources provide different sets of quality indicators, mainly in terms of number and definitions.

Our finding that indicators for lifestyle and self-management counseling were included in only a few sets is in line with the results of Glasgow et al. [14], who stated that major diabetes performance measurement sets (i.e. sources of indicators) do not include items on self-management or psychosocial items. These authors recommend expanding these sets to include such measures because there is broad consensus that patient-centered care and self-management support are essential evidence-based components of good diabetes care. We assume that this reasoning is also valid for our results regarding prevention, diagnostics and integrated care; despite international recommendations, these clinical domains are generally not translated into quality indicators. Most likely because diabetes care in primary care settings currently focuses on monitoring and treatment, what corresponds with our findings. In addition, the lack of knowledge about the validity of many indicators [13] may be an important explanation for the finding that few studies included indicators of structure of healthcare, e.g. for integrated care.

This review also revealed that most studies have applied quality indicators without reporting any systematic selection considerations. Scientific literature requires the underpinning of choices to operationalize research variables in order to guarantee measure properties such as validity and reliability. Although quality indicators are being derived from guidelines, we still see a variety of operationalizations of the concept of quality of diabetes care. As in other constructs, methodology can support the resolution of this problem. This process, however, is not yet commonly used for indicator measurements. The use of available indicator sets provides some basis for justification. But why specific indicators have been selected or dropped is unclear. This lack of systematic procedures in selecting quality indicators also may explain the variance in the number, content and descriptions of indicators.

These results raise the question of whether differences in the number, content and description of quality indicators influence the conclusions about the quality of diabetes care. In other words, what is the impact of using more, less or different quality indicators? This question addresses a validity issue. If the quality of diabetes care is considered as a construct that needs to be modeled by reflective indicators, the construct validity does not change by choosing other indicators: any two measures that are equally reliable are interchangeable [15]. Adding more indicators would then increase the precision of the aggregated scores. However, it is most likely and more appropriate to model the construct for diabetes care using formatives indicators because the diabetes guidelines describe all of the steps necessary for quality diabetes care. Therefore, the separate indicators are not interchangeable, and in that case, conclusions regarding the quality of care are affected by the selection of the indicators. This line of reasoning is applicable for only the process indicators. The outcome indicators all have an informative value of their own.

It must be mentioned that our results combine all studies from 1993 onwards. This is a long span of time and there have been considerable developments in the fields of diabetes care and quality indicators. For example, we noticed that the outcome indicators for blood glucose, lipids and blood pressure became more relevant in evaluating the quality of diabetes care in research published as from 2001. Apart from this, one might expect progress to be indicated by greater variation in the earlier years, with indicator sets becoming more standardized as time goes on. However, the classification in three periods did not show more consistency in the content covered or in the selection process.

Our study has some limitations. First, we looked at only the scientific literature, i.e. peer-reviewed journals, thus leaving relevant reports on developments in the quality of diabetes care, e.g. from websites, out of consideration. However, systematic selection procedures and justifications for the application of a measurement instrument are expected more often in scientific literature. In addition, by including the term diabetes in our search strategy, we might have missed some relevant studies on assessing the quality of care in primary care settings, including the condition diabetes mellitus. We do not have any reason, however, to assume that our findings would be different by including more indicator sets. Thirdly, we have been lenient in screening the literature for the use of selection methods. For example, if only one criterion for desirable attributes was mentioned, then the entire category of desirable attributes was scored positively, thus classifying the study as one that used this selection criterion. Therefore, our results for the use of systematic selection procedures are most likely overestimated.

We conclude that the sets of quality indicators show much variation in the content covered. Some important clinical domains of diabetes care in primary care settings, such as prevention, diagnostics and integrated care, are minimally or not at all represented in the indicator sets. Much variety exists in the number and descriptions of the selected quality indicators. This huge lack of consistency and uniformity in the operationalization of the concept of quality of diabetes care not only hinders the interpretation of the quality of care that has been measured and reported on but also hinders comparisons between quality assessments. Methodology for defining constructs such as quality of diabetes care is available but rarely used. More rigorous methods for the selection of quality indicators, including considerations of the concept of quality of diabetes care, are needed to address this important problem.

**Supplementary material**

Supplementary material is available at INTQHC Journal online.

**Acknowledgements**

The authors wish to thank Marc Padros-Goossens and Janine Liefers for their assistance with the data extraction using Access and Ellen Keizer for help in preparing the manuscript.
Funding

Funding was supported by The Netherlands Organisation for Health Research and Development (ZonMw).

References


