Is it Possible to Measure Prescribing Quality using only Prescription Data?

Morten Andersen

Research Unit for General Practice, University of Southern Denmark, J. B. Winsløwsvej 9A, DK-5000 Odense C, Denmark

(Received November 16, 2005; Accepted January 10, 2006)

Abstract: A large number of prescribing quality indicators based on register data have been proposed and many are used routinely in quality management. Often the content and face validity of indicators have been assessed by consensus methods, but studies analysing other validity aspects are scarce. Prescription data are frequently used for indicators, but they do not provide any direct information about disease and patient factors important for judging the quality of prescribing. If register-based proxies for diagnoses, disease severity or risk factors are employed, validation is essential. The concurrent validity of indicators should be assessed by comparing to a "gold standard" quality assessment at the patient level using all available clinical information. The validity of frequently used quality indicators of asthma treatment has been questioned and should be further investigated. NSAID prescribing indicators are currently under evaluation. In the future, detailed clinical information from practice databases and computerised hospital records will be an important data source for indicators and for validation studies. Furthermore, the statistical and epidemiological properties of prescribing quality indicators need more attention.

During the last decades there has been an increased focus on quality development in medical care, and the area of pharmacotherapy and prescribing is no exception from this general tendency. Using a slight modification of an existing well-established definition, a prescribing quality indicator is a measurable element of prescribing for which there is evidence or consensus that it can be used to assess quality, and hence change in the quality, of treatment provided (Lawrence & Olesen 1997). Quality of drug treatment in this context can be considered equivalent to rational pharmacotherapy, covering the aspects of effectiveness (maximising benefit), safety (minimising harm), economy (minimising costs or maximising cost-effectiveness) and patient acceptance of treatment (Barber 1995). Quality indicators can be valuable tools both in the hospital and primary care setting. This review will focus on prescribing quality indicators in general practice.

Usually, the quality of medical care is divided into structure, process and outcome (Donabedian 1988), and indicators can be constructed in all these areas. This also applies to prescribing indicators (Hoven et al. 2005). Structural indicators are aspects of the health system, organisation of care and available resources. In the area of prescribing it may be access to necessary drugs, availability of industry-independent drug information, an updated formulary or prescribing guidelines (World Health Organisation 1993). Process indicators cover the actual performance, the decisions and actions of the physician, for example prescribing the appropriate treatment or choosing a drug according to recommendations. Outcome indicators relate to the benefit or harm to the patient, equivalent to what is measured in clinical trials, but here assessed as consequences of prescribing in a non-experimental setting. Thus, outcome indicators cover all types of drug effects: risk of death or hospitalisation, measures of disease severity or activity, functional impairment, and impact on patients' well-being and quality of life. Prescribing quality indicators are most often process indicators.

Indicators are used for a number of different purposes, covering quality management in a broad sense. At the professional level, physicians use indicators for quality development and educational activities, assisting learning processes (Herbert et al. 2004). They are used by researchers to evaluate interventions, for example in experimental randomised studies testing new methods for changing prescribing behaviour (Søndergaard et al. 2002 & 2003). Finally, indicators are increasingly used by administrators of the health system for monitoring quality, screening for quality problems, benchmarking and providing feedback to physicians. There are sometimes conflicts between the professional, health administrative and political perspectives of indicators. To support accountability and transparency in the health care system, indicators with identification of respon-
sible institutions and professionals are increasingly made available to the public. This has been criticised, because evidence on the value of public disclosure is lacking and because it may have unintended consequences (Marshall & Davies 2001; Werner & Asch 2005). Another development that has caused controversy is the coupling of quality measurements to financial incentives, so-called “pay for performance” (Millenson 2004). There is concern that rewarding quality in selected clinical areas could lead physicians to ignore other diseases not covered by indicators, thus paying most attention to measurable areas of care.

Data sources and orientation of indicators

The information needed for prescribing indicators is most importantly data on drug consumption identified by geographic region and the prescribing physician or practice. Sometimes drug costs are also relevant. With these data, it is possible to calculate the most simple drug-oriented indicators (Haaijer-Ruskamp et al. 2004). Examples are the use of obsolete drugs (the target is zero consumption), or choice of recommended drugs among relevant alternatives. Although simple indicators, based on drug consumption and sometimes a population denominator, e.g. Defined Daily Doses (DDDs) per 1000 patients per year, can be used to compare prescribing patterns, it is doubtful whether they are at all able to measure quality (Blades et al. 1998).

Because many drugs are used for more than one disease, it is relevant to know the indication. To assess if a particular disease is treated appropriately, a disease-oriented indicator is needed. The picture becomes even more complex when the choice of drug depends on individual patient factors, for example age and sex, disease severity, and comorbidity including risk factors for adverse effects and contraindications. In these cases, a patient-oriented indicator is needed. For such an indicator, very detailed information at the individual patient level is necessary. It involves access to clinical data from patient records, maybe even individual patient assessments, using interview and clinical examination (Hanlon et al. 1992; Pont et al. 2004).

Collecting detailed clinical information at the patient level for quality assessments is expensive and time-consuming. Data on the prescribing, dispensing or reimbursement of drugs have often already been retrieved for administrative purposes either at practices, pharmacies or health insurers, and it is therefore convenient to use these secondary data sources for measuring prescribing quality. To compensate for the lack of clinical information in prescription databases, it is possible to use prescribing patterns at the individual patient level as proxies for diagnoses, disease severity and risk factors, attempting to construct disease or patient oriented indicators.

Validity of indicators

To fulfil their purposes, prescribing quality indicators should be valid, i.e. they should measure what they claim to measure. One important aspect is the validity of data sources. Drugs are often not used as intended, and measuring drug use as prescribing (in practice data), dispensing (at the pharmacy) or ingestion (assessed by patient interview or tablet counting devices) may lead to different results. However, even if there is high validity of all data elements included in an indicator (complete and accurate recording of prescribed drugs and diagnoses), we still cannot be sure that the indicator combining these data really measures quality. Usually, four different aspects of validity is considered.

Face validity is related to the indicator's relevance, credibility and acceptability. It can be assessed using different consensus methods (Campbell & Cantrill 2001). Delphi studies and the RAND appropriateness method have been used for this purpose (Campbell et al. 2000; Fitch et al. 2001; Rasmussen et al. 2005). It is recommended that face validity is assessed in the target groups who are going to use the indicators in practice, including both those who measure quality and those who have their performance measured.

Content validity implies that indicators should be evidence-based, in accordance with updated recommendations and current guidelines. The content validity is related to the initial steps of developing indicators, often involving discussion and consensus among experts. The link between indicators and the relevant evidence should be documented and explicit. This is particularly important when process indicators are used as substitutes for outcome indicators to ensure that meeting the indicator can be expected in reality to improve patients' health (Haaijer-Ruskamp et al. 2004).

Concurrent validity is established by comparing a proposed indicator to other quality measurements. Preferably this should be a “gold standard”, a reference measurement widely accepted as the best available. It is sometimes possible to create such a standard by assessing detailed clinical data (Pont et al. 2004). When quality is measured at the individual patient level, indicators can be evaluated using the same framework as a diagnostic test, comparing the indicator's assessment of suboptimal treatment to that of the “gold standard”. The indicator is characterised by its sensitivity (the proportion of cases of suboptimal treatment found using the reference assessment, also identified by the indicator) and positive predictive value (the proportion of suboptimally treated cases identified by the indicator that is confirmed by the reference assessment). The required sensitivity and positive predictive value depend on the purpose for which the indicator is used. High sensitivity is important when screening individual patients for potential quality problems. For monitoring changes, a selection of patients may be sufficient, and here the positive predictive value is relatively more important than the sensitivity. When assessing prescriber performance, both values should be high (Haaijer-Ruskamp et al. 2004).

If it is difficult to accomplish a reference measurement, however, a comparison of several indicators can instead give some information on their validity. To examine the concur-
rent validity of aggregate indicators at the practice level, correlation analysis can be used (Veninga et al. 2001; Dybdahl et al. 2004). It is also possible to use different indicators to identify physicians who adhere least to guidelines and then assess the level of agreement between indicators, for example with Cohen’s kappa (Veninga et al. 2001).

Construct validity is concerned with the indicators corresponding to our theoretical concepts (constructs) of quality, such as treatment appropriateness, rational therapy and cost-consciousness. This may include a synthesis of several of the validity aspects already mentioned above, concurrent validity being particularly important. Some of the epidemiological and statistical measurement properties of indicators could also be considered to be related to construct validity.

Indicators should be able to measure quality without bias or confounding. For controlling confounding, standardisation procedures can be useful. Standardisation requires data on the age and sex of treated patients and sometimes also demographic data for well-defined patient populations. Adjustment for age and sex is not always sufficient, and the distribution of morbidity (case-mix) may also be relevant, but data of this kind are rarely available.

Indicators should have a high reliability, i.e. when quality does not change, measurements using the indicator should not change either. Indicators should, however, also be able to discriminate between high and low quality and be sensitive to changes in quality over time. This is related to time-windows used for measurement and the statistical limitations of indicators based on rare events in relatively small patient populations (Bird et al. 2003).

Drug-oriented indicators are independent of the indication and patient characteristics, and therefore generally of good validity. Establishing content and face validity may be sufficient for these. Disease- and patient-oriented indicators pose other problems. The information collected in databases for patient management or administrative purposes is often incomplete or too inaccurate to allow assessment of diagnoses (indications, comorbidity), disease severity or other patient factors relevant for prescribing. When register-based proxies are used, the assessment of concurrent validity therefore becomes essential (Hoven et al. 2005).

Examples of indicator validity studies

Asthma indicators.

Several indicators have been developed for measuring the quality of asthma treatment, focusing on the use of inhaled corticosteroids. Prescribing an inhaled corticosteroid for patients requiring inhaled β2-agonists more than a couple of times a week is a well-established treatment principle, and the content validity of these indicators rest on evidence-based guidelines (World Health Organisation 1995).

The inhaled corticosteroid/inhaled β2-agonist ratio, calculated from the number of prescriptions or amount in DDDs has been used extensively as an indicator of asthma treatment quality. There are, however, several problems with this indicator. It is measured at the (practice) population level, without knowledge of indications. It is not known whether the patients getting inhaled steroids are those who really need them, and there is a risk of an ecological fallacy in the interpretation. Furthermore, the optimal value is difficult to establish.

Another indicator that has been used for evaluating the quality of asthma treatment at the population level is the proportion of asthma patients using inhaled corticosteroid (Hallas & Hansen 1993; Gaist et al. 1996). A similar indicator has also been used as a measure of intervention effect in a randomised study of feedback to general practitioners (Søndergaard et al. 2002). Asthma patients could conveniently be defined as users of inhaled β2-agonists, but these drugs are not only used for asthma, also for transient respiratory tract infections, most commonly in small children, and for chronic obstructive pulmonary disease. If there is only access to prescription data and not to indications or diagnoses, age restriction has been used to avoid including too many non-asthma patients in the quality measurement. It is also desirable to exclude sporadic use of inhaled β2-agonists. Thus, to create a more disease oriented indicator, asthma patients have been defined as persons 5–44 years old with at least two prescriptions of inhaled β2-agonists during a two-year period (fig. 1) (Søndergaard et al. 2002).

In population-based studies the proportion of inhaled corticosteroid-treated asthma patients has been calculated for one-year periods (Gaist et al. 1996), but for assessing intervention effects repeated measurements over time with shorter intervals are required. In the study of Søndergaard et al. (2002), quality was measured with 3 month intervals. Only current users of inhaled β2-agonists, with their last prescription within 3 months before the point of measurement were included in the quality assessment (fig. 1). For appropriate treatment with inhaled corticosteroid, a one-year time window was considered.

There is, however, considerable doubts about the validity of register-based asthma indicators. In a study comparing...
several asthma-indicators at the practice level, including the inhaled corticosteroid/inhaled β₂-agonists ratio and non-use of corticosteroids, there was a low correlation between indicators, and also low agreement between different indicators’ identification of physicians who adhere least to guidelines (Veninga et al. 2001). When register-based asthma indicators at the patient level were compared to individual assessments using clinical data, the sensitivity and positive predicted value was generally low (Pont et al. 2004). Among the investigated indicators, two were found suitable for screening purposes and one could be used for monitoring, but none were considered valid for the assessment of physician performance. It was a major problem that it was impossible to classify asthma severity accurately based on average daily doses of inhaled β₂-agonists measured in dispensing data (Pont et al. 2004). Another study evaluating computer algorithms for detecting suboptimally treated asthma patients in dispensing data (including underuse of inhaled corticosteroids), concluded that these algorithms could be useful for screening purposes (Stuurman-Bieze et al. 2004).

Some of the validity problems might be caused by a low precision of indicators because there are only few patients in the denominator. This is illustrated with data for 170 practices from the Odense University Pharmacoepidemiologic Database. With the indicator definition from fig. 1, the median (IQR) number of asthma patients per practice identified during a 3-month period is 15.5 (9–26). Fig. 2 shows the indicator values for each of the practices, ranked in ascending order with 95% confidence intervals (CI). The indicator varies considerably among practices with a median of 78.6% and 10th to 90th percentiles 60.0% to 95.1%. There is, however a large overlap of CIs with the population value of 77.9%, and it is thus not possible to distinguish the treatment quality of most practices from the average. Between-practice variation in cross-sections and within-practice variation over time should be further explored using different time windows for the indicator. It should be noted, that the indicator is much more precise when calculated for the entire asthma patient population (95% CI 76.5% to 79.3%) and that it might also still be useful for measuring differences among groups of patients or practices, depending on their size.

**NSAID indicators.**

The consumption of selective cyclooxygenase-2 inhibitors increased rapidly after their marketing. Because they were more expensive than traditional non-steroid antiinflammatory drugs (NSAIDs), it was recommended by the Danish Institute for Rational Pharmacotherapy to use them only for patients with an increased risk of gastrointestinal adverse effects to NSAIDs (Glintborg et al. 2002; Rasmussen et al. 2005). Risk factors included a history of peptic ulcer disease, age above 70 years, anticoagulant treatment, long-term treatment with oral corticosteroids, or serious illness. If an NSAID and low-dose acetylsalicylic acid was indicated, it was recommended to use a non-selective NSAID supplemented with a proton pump inhibitor protecting against the ulcer risk of both these drugs.

Risk factors related to drug treatment (anticoagulants, oral corticosteroids) could easily be identified in prescription data, and it was also possible that prescribing of ulcer drugs could be used as a proxy for previous ulcer. A series of indicators were constructed using prescription data, ranging from simple drug-oriented indicators to complex patient-oriented indicators focusing on the risk of adverse effects and using information on prescribed drugs as proxies for risk factors (Rasmussen et al. 2005). The content validity of these indicators was based on the aforementioned Danish recommendations and the evidence at that time, which was before cardiovascular adverse effects of cyclooxygenase inhibitors came into focus.

To assess the face validity of the developed indicators, a Delphi study was performed among Danish general practitioners (Rasmussen et al. 2005). They were chosen as target group because indicators are often used for feedback and learning processes in general practice. In this study, indicators related to the cost of treatment and simple volume indicators were either not considered valid or consensus could not be reached. Examples of such indicators are expenditures for NSAIDs or volume of NSAID use in DDDs per 1000 patients per year, standardised to the average practice population.

Two complex indicators with clinical content related to the treatment of patients with risk factors were, however, assessed as face valid. One of these were the percentage of patients treated with either a selective cyclooxygenase inhibitor or a non-selective NSAID + proton pump inhibitor among NSAID-treated patients at high risk for adverse effects (see above for risk factors). Physicians apparently prefer indicators with clinical content, even if they are fairly complex.

To qualify the choice of indicators, they have further been explored at the practice level using factor analysis.

---

**Fig. 2. Quality of asthma treatment in each of 170 practices from County of Funen, Denmark, measured during a 3-month period.** Percentage of asthma patients treated with inhaled corticosteroids (ICS), ranked in ascending order and with 95% (one-sided 97.5%) confidence intervals. Horizontal line indicates population value (77.9% of asthma patients treated with ICS).
(Rasmussen et al., unpublished results). This is a statistical technique useful for condensing the information contained in a number of original variables into a smaller set of new composite dimensions (factors) with a minimum loss of information (Mainous 1993). Preliminary results have shown that the method provides useful information when exploring relations between a large set of indicators. Furthermore, frequent use of cyclooxygenase inhibitors (judged by prescribing volume or preference percentage) was apparently associated both with their correct use among patients with risk factors and with inappropriate use among patients without increased risk of complications. Thus, simple indicators alone are not able to distinguish between high and low quality prescribing.

As a final step the validity of some of the advanced indicators, where individual patients are categorised as appropriately treated or not, is currently being investigated (Rasmussen et al., unpublished results). An expert panel contributed to create a “gold standard” for appropriate treatment based on information that can be obtained from patient records in general practice and implemented as a decision matrix. Data has been collected for 856 patients in 43 practices, and the agreement of register-based indicators with evaluation of treatment appropriateness based on the gold standard is currently assessed using an approach similar to that previously applied to asthma drugs (Pont et al. 2004).

Conclusions and perspectives

A large number of candidate indicators of prescribing quality based on register data have been constructed and validated by consensus methods. Few studies have systematically analysed other validity aspects of these indicators. Results for asthma drugs, the most thoroughly investigated drug group, have so far been disappointing. NSAID indicators are currently being validated. Validity studies may, however not be generalisable across countries because of differences between disease patterns, treatment recommendations and the organisation of health systems. To take into account these factors, indicators should ideally be validated in the setting in which they are used.

Prescription data only provide limited information on disease and patient factors important for judging the quality of prescribing. While register-based drug-oriented indicators only have few validity problems, disease- and patient-oriented indicators created using register-based proxies for diagnoses, disease severity or risk factors should always be validated. Most importantly, the concurrent validity of indicators should be assessed by comparing to a “gold standard” quality assessment at the patient level using all available clinical information. To get more accurate clinical data for patients in a larger scale, the most important future development will be to establish comprehensive practice databases and computerised hospital records with easily accessible information on diagnoses and clinical details. These data can either be used directly for indicators or for validation purposes. Although statistical and epidemiological properties of indicators in general are well-known, little research in this area has so far been done in the primary care setting and in particular for prescribing indicators.

References

Pont, L. G., P. Denig, T. Van Der Molen, W. Jan Van Der Veen &
MiniReview

MEASURING PRESCRIBING QUALITY


