Influencing physician prescribing in an international context
Sturm, Heidrun Beate

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2007

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Summary and Discussion
Summary and Discussion

Sub optimal uptake of evidence about effective drug therapy in primary care, along with the growing financial burden of drug expenditure causes researchers and policy makers to search for strategies that ensure rational prescribing. This thesis explored factors that potentially influence prescribing on different levels. Besides patient related factors, the emphasis is on the role of national settings and policies in European countries. To do so, two different approaches were used: cross-sectional analyses of heart failure prescribing in multiple European countries and systematic reviews using the methodology of the EPOC/Cochrane collaboration. In this chapter, the main findings are summarized and put in a wider perspective.

Summary of main findings

Prescribing and its determinants

Drug regimes for the treatment of chronic heart failure prescribed in European primary care in 1999, which were in line with first line guideline recommendations varied substantially between the 14 countries, roughly from 30% to 60% (Ch. 4). Taking patient characteristics into account added per country another 7% to 15% and could explain most of this variation in all three of our analyses of the IMPROVEMENT-HF prescribing data.

At the same time also the country, where therapy took place, consistently was a significant determinant. It significantly influenced the number of drugs used as well as prescribing of individual drug regimes.

Our analysis of national European treatment guidelines for chronic heart failure (Ch. 3) suggests that national guidelines only play a minor role in explaining international variation in heart failure therapy: The majority of recommendations were similar. The few differences that were found related to areas where evidence has undergone recent change. Prescribing did not appear to be associated with the national advice.

In the analysis of country-specific primary health care features (Ch. 5), quality of heart failure prescribing appeared to be impaired by strong overall primary care structures along with individual settings that limit care provision. In contrast, factors that facilitate access and treatment intensity such as physician density and fee-for-service payment improved prescribing.
Results from both prescribing policy Cochrane reviews, about pricing policies and financial incentive policies directed at prescribers, were each clustered around a specific setting: pricing policies were mainly about reference pricing in British Columbia based on data from elderly; the majority of the financial incentive studies dealt with British fundholding.

Both policies appear to have the intended effect on drug utilization, namely reducing drug expenditure without evidence of undesirable impacts on the quality of care or health outcomes.

**Practical Implications:**
Taking patient characteristics into account when assessing prescribing quality results in better judgement. Drug regimes for heart failure therapy in line with recommendations increased from 45% to 56% if co-morbidities were taken into account. Another 14% could be explained based on patient characteristics, however were usually not in line with evidence (Ch. 2). Results of this thesis provide information on specific areas of poor prescribing in chronic heart failure that can be specifically targeted in quality improvement measures: for instance patients who are male, have atrial fibrillation or a history of myocardial infarction have an increased risk of under-treatment (Ch. 2).

We found, that analyzed guidelines differed less in content than in comprehensiveness and form. The latter appeared to depend on the target group, suggesting that increased attention to the target group of primary care when issuing recommendations might enhance the uptake of guidelines (Ch. 3). These data also emphasize the relevance of other country specific factors for prescribing at a national level, one of which may be the structure of care provision.

In the light of known primary care’s positive effects on health outcomes our findings of negative effects on the process of prescribing suggest, that best outcomes might be achieved when good primary care structures are effectively combined with specialist involvement. Especially for diseases requiring complex and long-term therapeutic interventions such as heart failure, interdisciplinary approaches like disease management programs appear most promising (Ch. 5).

The policy analyses according to the Cochrane methodology showed that from many countries national evaluations appear to be unavailable. In addition, study designs of available policy evaluations often do not result in strong evidence. This could be improved, if evaluations are
planned in advance. Well-designed studies, including trials, could be applied, which might also be faster and more efficient than observational studies, and could reduce the risk of bias.

Assessing prescribing quality in international comparative studies:

The example of heart failure
With this set of studies we sought to single out determinants of prescribing in primary care on the national level as compared to factors on other levels, namely patient characteristics. To tackle this question chronic heart failure was chosen as an example of a condition, which is treated mainly in primary care and its complex therapy is strongly drug-based. Knowledge about factors improving quality of care for chronic conditions like heart failure is highly relevant for society, as its prevalence is high and still increasing while prognosis can be substantially improved under optimal therapy.

We first aimed to assess the impact of patient characteristics, which then allows focusing on national factors while taking patient characteristics into account.

Methodological considerations
This thesis is focussed on quality of care. In order to assess determinants of prescribing, quality of good prescribing has to be defined first and an appropriate indicator has to be chosen. Quality of care, as Donabedian ¹ put it, “may be almost anything anyone wishes it to be, although it is, ordinarily, a reflection of values and goals current in the medical care system and in the larger society of which it is a part”.

The indicator used depends on the objective of the quality assessment and the availability of data². Indicators used for external quality assurance, like benchmarking may be less detailed than indicators used for internal quality improvement activities, used to detect areas of weak performance (as suggested in our study in Ch 2) or for individual feedback to prescribers. The recent trend to include process data in salary and contracting negotiations, for instance in pay for performance from the UK³, underlines the relevance of indicator validity. For complex processes, influencing factors are highly relevant and therefore need to be included in the indicator. Patient characteristics and comorbidities are significantly influencing prescribing decisions for heart failure and therefore need to be incorporated in indicators for measuring quality of heart failure prescribing.
For comparative research, indicators should have the same content validity. Prescribing indicators are frequently defined based on established evidence as reflected in guidelines. It is relevant to realize, that evidence is translated by national or regional expert panels into local practice guidelines. The above definition of quality implies, that perception of gold standards may differ between countries resulting in different guideline recommendations. It has been shown for instance, that thresholds for the introduction of lipid lowering therapy varied significantly between national recommendations. To warrant equal content validity over all countries studied, we assessed the variability of national recommendations for heart failure treatment. As no major differences were found, recommendations of the guideline of the European Society of Cardiology were taken as the gold standard to assess prescribing in Europe.

Applied indicators: complex versus ...
To get detailed insight in prescribing decisions for individual patients, we developed a rather complex indicator based on an algorithm, where guideline recommendations were translated into appropriate drug regimes taking major co-morbidities into account. Subsequent steps of treatment intensity were reflected in the number of drug classes used. At every intensity level some alternative treatment regimes (different to the first line recommendation) were defined to still be in line with recommendations for patients with specific co-morbidities. For instance, patients with atrial fibrillation were also considered to be treated in line, when receiving a combination of ACE-inhibitor and digoxin instead of ACE-inhibitor and beta-blocker.

Using stepwise treatment regimes to assess quality of prescribing, improved insight into the complexity of prescribing for this disease: We found for instance that insufficient treatment was mainly limited to single drug regimes, while multi-drug regimes most often resulted in a recommended therapy. Also, clinical rationale of prescribers became visible: e.g. patients with impaired renal function were more likely not to be treated at all than patients with normal renal function, reflecting most probably uncertainty about renal side-effects of especially ACE-inhibitors. This is in line with Hobbs' findings, that primary care physicians' treatment decisions might be dominated by perceived treatment risk rather than treatment benefit. All this detailed information on patient characteristics promoting or hindering the use of specific drug combinations can be used to tailor interventions in quality improvement activities.

Although we used refined quality indicators, there might be still situations, which are judged inappropriate by our measures, while doctors' individual decisions might have been rational. Care has to be taken when presenting results of clinical performance. These measures can
always only indicate general trends and point to potential areas of sub-optimal performance and should not imply a judgement of individual performance. This is essential especially when data is intended to improve quality of therapy. Information will not be easily accepted, if it implies bad performance of the judged person. Therefore, our findings still should raise awareness of areas with a risk of under-treatment.

... simple indicators
In the case of heart failure, the frequency of ACE-inhibitor or All antagonist prescribing is the most commonly used performance indicator for administrative quality assessment. It is used for example by RAND, and by governmental organizations such as the US Agency for Healthcare Research and Quality (AHRQ) 7 or the Quality and Outcomes framework in the UK 8,9. Its solid evidence base is reflected since years in major national and international guideline recommendations 10,11. As the first line drug it is applicable to all CHF patients and rather independent of co-morbidities and severity. In addition, its assessment is feasible from most databases and practice based computer systems since it does not involve additional information12.

For our goal in the study of the relevance of the health care settings on prescribing decisions, we chose this simple quality indicator, which is the most feasible indicator. The risk of misclassification of recommended versus non-recommended prescribing was addressed by taking patient characteristics into account.

Improved treatment quality implies improved outcomes. This assumption is based on optimal and defined conditions of randomized controlled trials. However, to what extent trial results can be transferred to other circumstances, which have not been included in the study, yet might influence outcomes in real live situations adversely, is constantly under debate. One example is the effect of beta-blockers in the elderly13. Increasingly however, studies address this issue generating evidence under every-day conditions 14.

The influence of factors on different levels were also the reason when the negative effects of primary care (structure) on prescribing quality (process) appeared to be contradicting its positive effects on mortality (outcome). The differentiation between structure and process components of care helps to clarify this. Both, optimal structure and process are expected to result in better health outcomes. Yet, quality of prescribing is just one component of the care process leading to increased survival. Other factors, which might profit from strong primary care, such life-style or coordination of care, might contribute more to overall survival.
Evidence of policy-effects based on Cochrane Reviews:

By contrast to the first part, where we tried to detect different factors inherent in a health care setting that can modify prescribing decisions in primary care, the second part aims to assess if prescribing policies are effective in achieving the intended effects.

Pharmaceutical policies are intended to promote rational drug use and most of all, increase cost effectiveness of pharmaceutical expenditure. Policy interventions are implemented which aim to influence pharmaceutical expenditures on all levels, however their effects frequently remain unclear.

The EPOC group strives to collect available evidence about interventions published not only in scientific literature but also in governmental reports and alike and rate their methodological quality to improve evidence for decision makers15. The reviews presented here provide the following lessons:

Methodological challenges
Policy evaluations frequently suffer methodological flaws. For instance, simple before-after designs are lacking a control group not allowing conclusions about the true effect of the intervention, since effects of the numerous concomitant regulatory changes that are prevalent in most health care systems cannot be separated from the intervention’s own effect. Often these limitations are inherent in the nature of the policy intervention, as national regulations rarely are implemented only in a part of the organisation or country, which would create a control group within otherwise constant settings. Many studies based on longitudinal data lack necessary data prior to the intervention. The use of enough time-measures before and after the implementation would allow a proper interrupted time series analysis. Also longer follow-up could provide important supplementary evidence, as frequently effects are expected to be transient. At the same time however, the risk for bias related to other confounding interventions would increase with the length of the observation period.

These methodological problems can be overcome as demonstrated by the Canadian province of British Columbia (B.C.). In this case researchers achieved in collaboration with policy makers the sequential introduction of several interventions in pilot-like regional projects, creating quasi control groups for evaluation16. The pioneer role of B.C. in evaluating policies is reflected in our first review where the evaluation of reference pricing is dominated by studies originating
from B.C. Increasingly policy makers indeed require evaluations when implementing new regulations. Resulting evaluations are still of changing methodological quality, which do not allow clear conclusions about achieved effects. This might be improved, if policy makers and health care researchers incorporated the desired policy evaluation already in the planning of the policy.

**Strength of the evidence (GRADING)**

While included studies in the reference pricing review had mostly an interrupted time series (ITS) design, the vast majority of studies about British fundholding used controlled before-after (CBA) design. While it is often difficult to identify a comparable control group, in this specific case this was possible, as fundholding was not mandatory for all practices, creating a “similar” control group of non-fundholders (Eccles, 2003 692 /id). However, even in apparently comparable control and study groups, baseline and other characteristics might still be present and make interpretation difficult. Additionally, developments over longer periods can not be detected. Therefore evidence based on ITS designs is stronger; changes in trends can be detected in timely relationship with the intervention. In grading the strength of the overall available evidence, study design is one criterion 17. Therefore, overall evidence based on a similar number of studies is stronger for reference pricing with ITS studies than for fundholding based mostly on CBA studies.

**Limitations: selective inclusion**

Study design was also the main reason that only a minor part of publications about the interventions met the quality criteria of EPOC and could be included in the reviews.

Furthermore, includable studies seem to be mostly accumulated from few legislations. In the pricing policies review of ten included studies on reference pricing, six were from B.C. in Canada. Only one study about index pricing was included. In the review on financial incentives for prescribers, only studies on budgetary policies could be included, most of which were about British fundholding. In addition, also study designs were similar within countries: while Canadian studies were mostly ITS-designs, most British fundholding studies used a more simple controlled before-after design, although existing data theoretically would have allowed time-series analysis with an additional control group (resulting in a strong controlled ITS design).
Despite ongoing health reforms in most Western countries and growing cross-boarder comparisons, reforms remain vastly national. Consequently also policy evaluations are national. As a result a language bias can probably not be entirely prevented by systematic reviews, despite using search strategies without language restrictions. National reports might not be included in searchable international databases and English-speaking authors might have a lower threshold in publishing results internationally: also both predominant countries in our reviews, Canada and the UK are English-speaking. In addition, reports intended for policy makers might not be authored by scientists and therefore not fulfil requirements of scientific publications in international journals.

Together this indicates that a selection bias of unknown size is very likely in Cochrane policy reviews. Results are limited to specific settings and can therefore not be extrapolated to other settings and legislations.

This demonstrates a general potential flaw of Cochrane reviews recently again debated round a publication in the BMJ comparing the quality of Cochrane versus other (industri funded) reviews: too many Cochrane reviews were stating that the reporting of found trials is poor and therefore more research is required before a clinical recommendation can be made (Tostad). Also, the lack of valid evidence can easily be confused as evidence of a lack of differences (Coyne). The question was raised, whether it is useful to say that several trials do not permit an inference on effectiveness? The same question is of course pressing for policy makers: what is the added benefit for policy makers, when the majority of reports and studies cannot be included in a review? The answer of course has to be cautious: The systematic review allows to get an overview about valid quantitative evidence, helping to separate misleading weak reports from conclusions based on fairly strong studies. Policymakers usually do neither have the knowledge nor the time to judge methodological limitations. However, Cochrane reviews cannot replace policy analysis for the interpretation of evidence for a current situation. Only together it seems, useful information can be achieved.

Role of Cochrane for policy reviews

Cochrane reviews therefore have a very specific role in policy evaluation: they collect and grade available studies about similar interventions, which are designed well enough to allow statistically valid conclusions. This helps to distinguish methodologically valid evaluations from potentially misleading studies with weak statistical design. However, conclusions from such quantitative assessments cannot replace informed qualitative policy analysis, which evaluate policy interventions in the light of specific circumstances.
Chapter 8

Policy makers have to be careful in interpreting Cochrane reviews: simply because problems are similar, one cannot assume that solutions are similar as well. While experiences in other places can be used as inspiration, generalization of effects is only reasonable in the rare case that it holds over multiple diverse settings. If this can be demonstrated by valid studies, Cochrane reviews can add substantially to evidence. However, for all other cases, it has to be emphasized, that structural diversity of different settings cannot be taken into account in this form of analysis. In addition, potential “inclusion bias” has to be taken into account when interpreting evidence of systematic Cochrane reviews.

Conclusions

In this thesis the following determinants were found relevant for prescribing:

1. As expected, patient characteristics explained most of the variation in prescribing. This implies that quality improvement efforts on a professional level as opposed to the policy level are essential to achieve optimal therapy.

2. In all cross-sectional heart failure studies in this thesis, on top of clinical characteristics the country where the treatment took place was a statistically significant determinant of prescribing.

3. Prescribing policies evaluated in the Cochrane reviews, appeared to affect prescribing and decrease drug expenditure. However, extrapolation to other settings is difficult.

4. National guidelines for heart failure therapy did not appear to play a significant role in explaining international differences in heart failure prescribing, mainly because recommendations were similar.

5. Strong primary care structures appear to have a negative effect on prescribing quality for heart failure.
Implications

Improving quality of prescribing can be addressed on different levels: physician centred measures or regulatory interventions of care provision and structure. This thesis provides evidence that both approaches are relevant and indicates therefore that both should be considered together.

Implications for health care planners

Since patient characteristics determine prescribing decisions for the biggest part (at least in the case of chronic heart failure), effective prescriber-centred measures are important. Information about specific situations where poor prescribing is found to be more likely (our results provide some concrete examples for chronic heart failure) can enhance continuous medical education (CME) or quality improvement activities. Results of the guideline analysis suggest further, that the uptake of guideline recommendations might profit from increased attention to the target group – in our case primary care – already when issuing recommendations.

Within the national setting, free access to care seems to enhance prescribing quality for chronic heart failure, while strong primary care (PC) structures (including a gatekeeper role of the primary care physician) had a negative impact on prescribing quality. In the light of known positive effects of PC on overall health outcomes (including improved cardiovascular mortality)\(^3\), our findings suggest, that in the case of chronic diseases such as heart failure, combining PC with interdisciplinary approaches like disease management programs promise optimal clinical outcomes.

Implications for policy makers

The relevance of national settings for prescribing decisions was shown in the cross sectional studies as well as in the longitudinal studies included in the Cochrane review. This means, that the national health care environment (determined by factors such as organisation of and access to care, availability of diagnostics and therapeutics, financial incentives for physicians and patients) influences individual medical decisions significantly. Better knowledge about these relations can be used to actively adjust settings in order to facilitate best performance or to prevent potential adverse effects when planning reforms. Policy makers need to be made aware of these factors in order to integrate them in the policy planning process.
However, care has to be taken with respect to unreflected transfer of findings from one place to another. Generalizing conclusions so far can rarely be based on evidence.

Because pharmaceutical policies might cause harm as well as benefits, also for policies proper evaluation is crucial. Well-designed studies will result in stronger evidence and might even be performed more efficiently. Planning, including funding of evaluations therefore should be a routine part of the policy process and take place well ahead of the policies’ implementation. Routine assessment of performance, costs etc can serve as baseline measures and safe extra efforts. Also here careful planning in cooperation with researchers is indispensable in order to facilitate systematic and standardised data collection and adjust collected data to evaluation needs.

**Implications for researchers**

International comparative health care research as applied in this thesis offers the possibility to detect factors within the health care setting that modify health care provision. The systematic Cochrane policy reviews collected experiences from various legislations. In theory, similar results for a given intervention in all settings indicate fairly strong evidence. However, if results are based on few settings as in this case, these results cannot easily be transferred to other settings without detailed analysis of modifying factors. Cross sectional research can complement that knowledge by detecting relevant modifying factors.

The purpose of the assessment should guide the choice of the indicator. If clinical feedback is planned and therefore more detailed data is needed, data requirements should be carefully planned ahead of collection and where possible included into routine data collection.

Relatively few studies meet the strict methodological inclusion criteria of the Cochrane Collaboration. To obtain more high quality evaluations, research design should be integrated into the policy development process. Then available data and financial resources can be adjusted to optimal methodological requirements.
References

Chapter 8

