Effects of evidence-based clinical practice guidelines on quality of care: a systematic review

M Lugtenberg, J S Burgers and G P Westert

*Qual Saf Health Care* 2009 18: 385-392
doi: 10.1136/qshc.2008.028043

Updated information and services can be found at:
http://qshc.bmj.com/content/18/5/385.full.html

**References**

This article cites 52 articles, 19 of which can be accessed free at:
http://qshc.bmj.com/content/18/5/385.full.html#ref-list-1

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To order reprints of this article go to:
http://qshc.bmj.com/cgi/reprintform

To subscribe to *Quality and Safety in Health Care* go to:
http://qshc.bmj.com/subscriptions
Effects of evidence-based clinical practice guidelines on quality of care: a systematic review

M Lugtenberg,1 J S Burgers,2 G P Westert1,3

ABSTRACT

Background: Evidence-based clinical guidelines aim to improve the quality of care. In The Netherlands, considerable time and effort have been invested in the development and implementation of evidence-based guidelines since the 1990s. Thus far, no reviews are available on their effectiveness. The primary aim of this article was to assess the evidence for the effectiveness of Dutch evidence-based clinical guidelines in improving the quality of care.

Methods: A systematic review of studies evaluating the effects of Dutch evidence-based guidelines on both the process and structure of care and patient outcomes was conducted. The electronic databases Medline and Embase (1990–2007) and relevant scientific journals were searched. Studies were only selected if they included a controlled trial, an interrupted time series design or a before and after design.

Results: A total of 20 studies were included. In 17 of 19 studies that measured the effects on the process or structure of care, significant improvements were reported. Thirteen of these studies reported improvement with respect to some of the recommendations studied. In addition, the size of the observed effects varied largely across the recommendations within guidelines. Six of nine studies that measured patient health outcomes showed significant but small improvements as a result of the use of clinical guidelines.

Conclusions: This review demonstrates that Dutch evidence-based clinical guidelines can be effective in improving the process and structure of care. The effects of guidelines on patient health outcomes were studied far less and data are less convincing. The high level of variation in effects across recommendations suggests that implementation strategies tailored to individual recommendations within the guideline are needed to establish relevant improvements in healthcare. Moreover, the results highlight the need for well-designed studies focusing on the level of the recommendations to determine which factors influence guideline utilisation and improved patient outcomes.

Increasingly, clinical practice guidelines (CPGs) are being developed in all areas of medicine as a means to improve the quality of care. By translating the best available scientific evidence into specific recommendations, guidelines can serve as useful tools to achieve effective and efficient patient care.1 Whereas guidelines initially were based on consensus among experts, guideline development has been gradually formalised and evidence-based guidelines—linking the individual recommendations with their supporting evidence—are becoming standard practice.2 Developing evidence-based guidelines, however, does not guarantee improved quality of care. Effective implementation should ensure guideline adherence in practice and subsequently lead to improved patient outcomes.

Studies measuring the effects of guidelines on quality of care have predominantly focused on effects on clinical practice. Several international reviews showed that the majority of guideline studies resulted in significant improvements with respect to the process of care.3–5 Fewer studies have focused on the effects of guidelines on patient health outcomes. One review indicated improved patient health outcomes,6 whereas a second review, focusing on primary care, did not find a positive effect.6 However, most of the studies that were included in these reviews used older guidelines that were not developed according to the current standards of evidence-based medicine.7

Guideline utilisation is complex and many factors may influence the impact of guidelines on care. Factors linked to the guideline itself are the strength of the evidence,8 9 the method of development and transparency of the guideline10 and the perceived adoptability, complexity and triability of the recommendations.11 12 Apart from guideline factors, the nature of the implementation strategy can contribute to guideline utilisation. Passive strategies, such as educational material and meetings, generally have a small effect and multifaceted strategies are not necessarily better than single interventions.13 14 Finally, contextual, organisational and cultural factors may impede or limit guideline adoption regardless of how thoroughly they are implemented.14 15

The Netherlands has been a forerunner in evidence-based guideline development and guideline implementation research, compared with other European countries.16 Since 1982, more than 200 guidelines have been developed by the Dutch Institute for Healthcare Improvement (CBO) and the Dutch College of General Practitioners (NHG), the two most prominent guideline organisations in The Netherlands. Historically, the CBO focused on secondary care and the NHG on primary care, since there is a clear-cut distinction between primary and secondary care in the Dutch healthcare system. In the last decade, other organisations have also become active in guideline development. Partly because of the role of the Centre for Quality of Care Research (since June 2008 Scientific Institute for Quality of Healthcare), many implementation studies have been conducted to measure the effectiveness of the Dutch guidelines.17

In spite of a considerable investment in the area of evidence-based clinical guidelines in The Netherlands, thus far, it is unclear to what extent these activities have been successful in improving...
compliance with guidelines and patient health outcomes. By examining the impact of evidence-based guidelines in a country the size of The Netherlands, which features well-defined organisations responsible for guideline development, unique observations can be made. The primary aim of this study is, therefore, to provide an overview of the effectiveness of Dutch evidence-based guidelines in improving the quality of care. In addition, we want to explore which factors are associated with guideline utilisation and improved patient outcomes.

METHODS
Concepts and definitions
In this review, CPGs were defined as “systematically developed statements to assist practitioner decisions about appropriate healthcare for specific clinical circumstances.” Guidelines that use the results of systematic literature reviews in formulating the recommendations and that link the individual recommendations with their supporting evidence were regarded as evidence-based CPGs. A recommendation was defined as “any statement that promotes or advocates a particular course of action in clinical care.”

Implementation was defined as “a planned process and systematic introduction of innovations or changes of proven value; the aim being that these are given a structural place in professional practice, in the functioning of organisations or in the health care structure.” Dissemination, on the other hand, is regarded as more passive than implementation and involves strategies such as distributing guidelines or publication of guidelines in scientific journals.

To evaluate effects on quality of care, we used Donabedian’s model, which distinguished the structure, processes and outcomes of care. Structure of care refers to “human, physical and financial resources that are needed to provide medical care” (eg, the presence of spirometry in general practice). Process of care refers to “the set of activities that go on within and between practitioners and patient” (eg, prescription of medication), whereas “the change in a patient’s current and future health status that can be attributed to antecedent health care” (eg, blood pressure) is defined as outcome of care.

Search strategy
A systematic literature search was conducted in Medline, Embase and relevant Dutch scientific journals. Searches were performed in Medline and Embase of literature published from 1990 to May 2007 using several combinations of key-words (Appendix A). We did not include studies published before 1990, as evidence-based guideline development in The Netherlands started in the early 1990s. To identify Dutch-language publications we performed a sensitive search in Medline (1990–2007) with the free text word “guideline”, limited to Dutch language. In addition, two relevant Dutch scientific journals, Huisarts & Wetenschap and Nederlands Tijdschrift voor de Geneeskunde, were searched for additional studies.

Two reviewers (ML and JB) independently screened the titles and abstracts of the articles and selected 165 potentially relevant articles. Discrepancies were resolved by discussion and consensus. These articles were further selected according to the following inclusion criteria (fig 1).

(1) The study concerned (a) clinical Dutch guideline(s). Drug formularies, patient guidelines and European guidelines were excluded.

(2) The study addressed the adherence to recommendations related to the process and structure of care and/or the effects of guidelines on patient health outcomes.

(3) The study concerned (an) evidence-based guideline(s). Local or regional protocols and guidelines that were derived from evidence-based national guidelines were also included.

(4) The study included a controlled trial (randomised controlled trial, controlled clinical trial), an interrupted time series or a before and after design. Studies that evaluated the effectiveness of different guideline dissemination and implementation strategies were included as well as studies that measured the effect of a guideline against a non-intervention control group.

In addition, we consulted reference lists from all articles that were retrieved for more detailed information.

Data extraction
Three categories of studies were distinguished based on the target users of the guideline(s): (1) general practitioners; (2) medical specialists and (3) other healthcare providers, such as physiotherapists and midwives. The following data were collected from each study: type of guideline (national or local/regional); clinical area; setting; study design; number of included patients and physicians; type of intervention; process and structure measures; patient outcome measures; effects on process and structure of care; and effects on patient health outcomes.

Methodological quality
The methodological quality of the eligible studies was assessed by one reviewer (ML) and checked by a second reviewer (JB) using the quality criteria of the Cochrane Effective Practice and Organisation of Care Group (EPOC). The EPOC quality criteria checklist includes seven criteria for randomised controlled trials, seven criteria for controlled before and after studies and seven criteria for interrupted time series. Although the EPOC criteria were not developed to assess the methodological quality of uncontrolled before and after studies, we used them for these studies as well, since there are no high-quality checklists available to measure the quality of these type of studies. The quality criteria, such as concealment of allocation, follow-up of professionals and follow-up of patients or episodes of care, were scored as “done”, “not done”, “not clear” and in some cases as “not applicable.”

Data synthesis
Due to the heterogeneity of the studies, pooling of the results and calculating an overall estimate of the effects were not possible. Instead, we summarised the effectiveness in three categories: mostly effective (if the study demonstrated a significant effect on more than half of the outcome measures), partly effective (if the study reported a significant effect on half or less than half of the outcome measures) and not effective (if no significant effect was demonstrated). If a study evaluated the effectiveness of different guideline dissemination and implementation strategies, effectiveness was determined by the observed significant improvement in either of the study groups, rather than by a significant improvement in the intervention group compared with the control group.

RESULTS
Description of the studies
Setting, study design and type of intervention
Thirty articles referring to 20 different studies fulfilled the inclusion criteria and were included in the selection. Most of the included studies targeted general practitioners (table 1).
Three studies were conducted among medical specialists in the Netherlands and four studies targeted other healthcare providers. The majority of the studies concerned guidelines that had been developed at the national level, while five studies concerned preventive care and three a combination of both. The most common studied medical conditions were cardiovascular diseases, influenza, diabetes mellitus type II, low-back pain, asthma and COPD. Most of the studies were cluster randomised controlled trials. The studies used predominantly multifaceted intervention strategies to implement the guideline(s). The intervention strategies most often used as part of multifaceted interventions were educational meetings, distribution of educational material, and audit and feedback.

Methodological quality

Overall, the quality of the included studies was moderate. A priori calculations of sample size were reported in seven studies. In the majority of the studies that allocated study groups, practices or groups of healthcare providers were the unit of allocation, thereby protecting participants against contamination. In four studies, allocation was by individual provider.

Five studies reported that more than 20% of providers dropped out. In four studies, data from less than 80% of patients were reported for at least one of the outcome measures or points in time. Two studies reported an agreement in (some of the) outcomes between raters of less than 90% (or $\kappa < 0.8$). In most of the studies, however, reliability of outcome measures was not reported at all.

Effects on quality of care

Four studies evaluated the effectiveness of different dissemination or implementation strategies, while 16 studies evaluated the effectiveness of a single implementation strategy (table 2).

Effects on process and structure of care

Of the 19 studies that examined effects of the guideline(s) on the process or structure of care, 17 showed significant improvements (table 2). The majority of these studies reported improvements with respect to some of the recommendations studied. Only four studies showed an effect on all outcomes measuring adherence to the guideline. In six studies, improvements were observed in half or less than half of the assessed process or structure of care measures. Two studies failed to demonstrate any effect on the process or structure of care.

Overall, the size of the effects varied largely across the recommendations within a guideline. Significant improvements in adherence to recommendations ranged from 7.2% to 88% in the 17 studies and varied by 76% across recommendations within one study. In studies including a control group, odds ratios ranged from 0.2 (95% CI 0.1 to 0.6) to 27.15 (95% CI 12.86 to 57.24).
Effects on patient health outcomes

Nine studies assessed the effects of guidelines in terms of patient health outcomes. Six of these studies reported significant improvements in at least some of the outcomes studied (table 2).31 32 34–37 43 44 48–52 Wolters et al31 32 and De Laat et al34–37 reported improvements in all outcome measures, while four studies31 32 34–37 43 44 demonstrated modest improvements in some of the assessed patient outcome measures. In three studies38–40 47 no effect on patient health outcomes was observed.

Characteristics of the studies and effects on quality of care

All studies focusing on preventive care were mostly effective in terms of the process or structure of care (table 3).25–30 42 With respect to type of design, all uncontrolled before and after studies that measured effects on process or structure of care27–

or on patient health outcomes40 41 were categorised as mostly effective. In contrast, both studies that failed to demonstrate an effect on clinical practice were cluster randomised controlled trials.22 41

There were no differences in effects on quality of care with respect to type of healthcare provider, type of guideline or between studies that used multifaceted intervention strategies and studies that used a single intervention.

**DISCUSSION**

Overall, the results of this study demonstrate that there is evidence for the effectiveness of Dutch evidence-based guidelines on the process and structure of care in The Netherlands. The majority of the studies reported improvement with respect to some of the recommendations studied. In addition, the size of the effects varied largely across recommendations within the guidelines. The effects of guidelines on patient health outcomes were studied far less and data are less convincing. Two-thirds of the studies that measured patient outcomes reported significant improvements. However, the observed changes in patient outcomes were generally modest and only found for some of the outcomes studied.

Findings from our review in terms of the process of care are comparable with those of previous international reviews which demonstrated small to moderate improvements.5–9 With respect to patient outcomes, results from earlier reviews were inconsistent.5–9 However, measuring patient outcomes is complex due to many factors such as long delays and confounding of many outcomes.24–30 Our review provides some evidence for the effects of guidelines on patient outcomes. It also suggests that guidelines focusing on preventive care are particularly effective in improving the process or structure of care. Nevertheless, the number of studies in our review does not allow us to draw firm conclusions on the effects of guidelines on patient outcomes or on factors that contribute to improved quality of care.

An important finding of our study is that the observed effects varied largely across recommendations. The variation could be explained by barriers related to individual recommendations rather than barriers that apply to the guideline as a whole. For example, although we selected only evidence-based guidelines in our study, the strength of the evidence may vary across recommendations, thereby influencing their impact.5–9 Also, a recommendation may not be performed because of other factors such as healthcare professional issues (eg, lack of motivation) or environmental factors (eg, lack of resources). Future research should focus on barriers related to both the guideline and its specific recommendations when exploring the association with effects on quality of care.

Furthermore, the fact that the effects of guidelines varied largely across recommendations might suggest that guideline implementation should focus more on individual recommendations rather than the guideline as a whole. Whereas the majority of studies included in our review used multifaceted strategies to implement the guideline, these were generally not tailored to individual recommendations. The nature of the implementation strategy is often the same for all recommendations within a guideline. A more focused approach, based on the results of an analysis of barriers of adhering to individual recommendations could improve the use and effectiveness of guidelines in practice.

One of the strengths of the present study is that it focuses on the effectiveness of evidence-based guidelines on quality of care. Previous reviews also considered guidelines that were not developed according to the standards of evidence-based medicine. In addition, in our study the effects of guidelines within one healthcare system were analysed. The guidelines in the studies included in this review were produced by well-known and credible organisations in The Netherlands and these guidelines generally have acceptable quality scores and are adequate tools for healthcare improvement.24–30

As stated before, The Netherlands has been a forerunner in evidence-based guideline development and implementation research in Europe. Based on the current literature, we have no reason to believe that the effectiveness of evidence-based
<table>
<thead>
<tr>
<th>First author; year</th>
<th>Clinical area</th>
<th>Study design</th>
<th>Intervention</th>
<th>Effects on process/structure of care</th>
<th>Effects on patient health outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engers et al, 2005</td>
<td>Management of low-back pain</td>
<td>Cluster RCT; GPs were randomly assigned (167 GPs, 616 consultations; 531 patients).</td>
<td>Multifaceted tailored implementation strategy (distribution of guidelines, educational workshop, a tool for patient education, a tool for reaching agreement with other healthcare providers) vs no intervention.</td>
<td>Fewer referrals to a therapist during follow-up in IG compared with CG (36% vs 76%; OR 0.2; 95% CI 0.1 to 0.6). No sign. differences in 3 other outcome measures (eg, prescription of pain medication on a time-contingent basis)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Frijling et al, 2003 to 2004</td>
<td>Preventive cardiovascular care</td>
<td>Controlled before and after study (617 general practices: IG: 316/C:G: 301).</td>
<td>Multifaceted intervention (conferences, dissemination of manuals and support from trained nonphysicians during outreach visits) vs no intervention.</td>
<td>Improvement in all 8 structure-of-care indicators in IG (varying from 12.5% for reminder for assessment of cardiovascular risk-factor profiles to 39.3% for reminder for BP measurement) and in 2 of 7 process-of-care indicators (varying from 9.7% for smoking cessation to 35.3% for measuring BP) compared with CG. OR from 1.45 (95% CI 1.02 to 2.07) to 27.13 (95% CI 12.66 to 57.24), (mostly effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Hak et al, 2000</td>
<td>Influenza and influenza vaccination</td>
<td>Uncontrolled before and after study (988 general practices).</td>
<td>Multifaceted intervention (employment of facilitators, information-based methods, small-group consensus meetings, individual instructions and supportive computer software).</td>
<td>Improvement in vaccine uptake (7.2%) and all 7 aspects of influenza immunisation practice (varying from 9% for immunisation by practice assistant to 37% for sending personal reminders). (mostly effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Hermens et al, 1999 to 2001</td>
<td>Effective cervical cancer screening</td>
<td>Uncontrolled before and after study (988 general practices).</td>
<td>Multifaceted intervention (educational materials, a computerised module, small group education meetings, consultations, outreach visits).</td>
<td>Improvement in 9 of 10 indicators (varying from 5% for presence of a sex-age register to 33% for sending a reminder to non-compliers). (mostly effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Hulscher et al, 1997</td>
<td>Organisational guidelines for cardiovascular disease prevention</td>
<td>Controlled before and after study (95 general practices: outreach visit: 33; feedback: 31, CG: 31).</td>
<td>Outreach visit method (visiting of practices by trained nurses), a feedback method (provision of feedback report with advice) vs no intervention.</td>
<td>Outreach visit group improved in 6 of 10 indicators (varying from 12% for sex-age register available to 88% for written protocols available). No improvements in feedback group. (mostly effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Jans et al, 2007 to 2011</td>
<td>Management of asthma and COPD</td>
<td>Controlled before and after study (19 practices: IG: 14/C:G: 5)/370 patients (IG: 280/C:G: 90).</td>
<td>Multifaceted intervention (identification of barriers, documentation of the care provided, specific education, feedback and peer review) vs no intervention.</td>
<td>Improvement in 4 of 8 aspects in IG (varying from 36% for monitoring of medication compliance to 74% for measurement of PEFR) compared with CG. (partly effective)</td>
<td>1 of 4 outcomes improved in IG compared with CG (mean PEFR from 78.5 to 81.0). (partly effective)</td>
</tr>
<tr>
<td>Kasje et al, 2006</td>
<td>Treatment of CHF and DM2</td>
<td>Cluster RCT, balanced incomplete block design, peer groups were randomised (16 peer groups: 10 CHF, 6 T2DM; 65 GPs; 979 patients).</td>
<td>Interactive educational programme for small peer groups (one arm received a programme on treatment of CHF, the other arm on hypertension treatment in DM2).</td>
<td>No effect on both outcome measures (prescribing of ACE inhibitors and antihypertensive treatment) in both groups compared with CG. (not effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Lobo et al, 2002 to 2004</td>
<td>Prevention and treatment of cardiovascular care</td>
<td>Cluster RCT; practices were randomised (124 practices/185 GPs/2288 patients; 537 diabetes/617 cardiovascular disease/1114 hypertension).</td>
<td>Multifaceted intervention (feedback reports and support from facilitators including discussion of feedback reports, selection of clinical issues for improvement, selection of methods for change and evaluation during 15 outreach visits per practice) vs no intervention.</td>
<td>Improvement in all 6 aspects of organising preventive cardiovascular care, such as the nr. of preventive tasks performed by practice assistant in IG compared with CG. Improvement in process of cardiovascular care in 5 of 12 indicators: OR from 1.55 (95% CI 1.35 to 1.77) for risk factors in patients with hypertension to 4.11 (95% CI 2.17 to 7.77) for checking for clinical signs of deterioration in patients with heart failure. Improvement in 2 of 7 indicators of process of diabetes care. OR from 1.52 (95% CI 1.07 to 2.16) for eye examination to 1.68 (95% CI 1.19 to 2.39) for foot examination. (mostly effective)</td>
<td>Improvement in 2 of 8 aspects of HRQOL in diabetes patients compared with CG (mean change from 3.71 (95% CI 0.73 to 6.68; scale 0–100) for mental health to 3.93 (95% CI 1.08 to 6.78) for vitality) and in 3 of 8 aspects in patients with cardiovascular disease (from 3.01 (95% CI 0.72 to 5.30) for vitality to 3.96 (95% CI 0.50 to 7.42) for social functioning). No improvement in patients with hypothyroidism. (partly effective)</td>
</tr>
<tr>
<td>Renders et al, 2001 to 2002</td>
<td>DM2</td>
<td>Controlled before and after study (27 GPs: IG: 22/C:G: 5) and 389 patients (IG: 312/C:G: 77).</td>
<td>Multifaceted intervention (distribution of guidelines, postgraduate education, audit and feedback, templates to register diabetes care; a recall system) vs no intervention.</td>
<td>Improvement in all 9 indicators (varying from 16% for measurement of BP to 44.7% for measurement of HDL cholesterol) compared with CG. OR from 2.43 (95% CI 1.01 to 5.82) to 12.08 (95% CI 4.70 to 31.01), (mostly effective)</td>
<td>The intervention did not improve any of the 14 patient outcomes, such as BP and HbA1c. (not effective)</td>
</tr>
<tr>
<td>Smeel et al, 1999</td>
<td>Treatment of asthma/COPD</td>
<td>Cluster RCT; GPs were randomised (34 GPs: IG: 17/C:G: 17) 433 patients (IG: 210/C:G:223).</td>
<td>Multifaceted intervention (an intensive, interactive group education and peer review programme) vs no intervention.</td>
<td>Improvement in 2 structure-of-care aspects (varying from 16% for skills to 18% for presence of peak flow meters) in IG compared with CG. None of the 6 process-of-care aspects showed sign. changes. (partly effective)</td>
<td>No changes in any of the 3 patient outcomes (symptoms, smoking habit, disease specific quality of life) compared with CG. (not effective)</td>
</tr>
</tbody>
</table>
guidelines, in terms of quality of care, is different in other countries that more robust designs are generally used to assess these effects. Contextual and country-specific factors may, however, influence the effectiveness of evidence-based guidelines in terms of quality of care. In some countries, such as Scotland and New Zealand, guidelines are produced in a similar context as in The Netherlands, making the conclusions applicable to these countries as well.

Several limitations to the present study can be mentioned. First, despite the long tradition of evidence-based guideline development in The Netherlands, the number of studies measuring the effects of guidelines with a robust design was

---

### Table 2 (Continued)

<table>
<thead>
<tr>
<th>First author; year</th>
<th>Clinical area</th>
<th>Study design</th>
<th>Intervention</th>
<th>Effects on process/structure of care</th>
<th>Effects on patient health outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van der Weijden et al, 2005</td>
<td>Cholesterol for screening and management of hypercholesterolemia</td>
<td>Cluster RCT; practices were randomised (32 GPs, 16/CG: 16; 20 general practices; 3950 patient records).</td>
<td>Multifaceted intervention (guideline dissemination, group education, supportive materials, feedback and face-to-face instruction on location) vs guideline dissemination.</td>
<td>No improvement in 2 outcome measures (quality of selective case finding and quality of diagnostic procedures) in both groups. (not effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Van Essen et al, 1997</td>
<td>Influenza vaccination</td>
<td>Controlled before and after study (2 regions; 79 practices. IR: 92/CR: 97): 242 GPs (IR: 118/CG: 124); 550,000 patients.</td>
<td>Multifaceted intervention (distribution of educational materials, educational meetings; distribution of vaccines, information on practice routines, etc) vs no intervention.</td>
<td>Improvement in IR on vaccine rate (21%) and 3 of 5 organisational aspects (varying from 18% for special vaccination hours to 29% for vaccine in stock) compared with CR. (mostly effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Wolters et al, 2005a, 2005b</td>
<td>Management of lower urinary tract symptoms</td>
<td>Cluster RCT; GPs were randomised (42 GPs, IG: 70/CG: 72; 167 patients).</td>
<td>A distance learning programme (evidence-based information, assessment of learning needs, knowledge test, patient education materials) vs written guidelines.</td>
<td>Lower referral rate to an urolgist in distance learning group (OR 0.08; 95% CI 0.02 to 0.40). No effect on other 2 primary outcomes (PSA testing, prescription of medication). (partly effective)</td>
<td>No difference between groups. (not effective)</td>
</tr>
<tr>
<td>Kamphuisen et al, 2002</td>
<td>Diagnosis of pulmonary embolism</td>
<td>Uncontrolled before and after study (117 patients before and 119 patients after).</td>
<td>Physicians were asked to strictly follow the diagnostic protocol after a non-high-probability ventilation- scan.</td>
<td>Improvement of 26% in adherence to the guideline (20% before and 46% after the implementation of the guideline). (mostly effective)</td>
<td>Not measured</td>
</tr>
<tr>
<td>Schouten et al, 2007</td>
<td>Antibiotic treatment of lower respiratory tract infections</td>
<td>Cluster RCT; multicentre; hospitals were randomised (6 hospitals; 1906 patients).</td>
<td>Multifaceted intervention (feedback on baseline performance and selection of interventions on the basis of analysing barriers) vs no intervention.</td>
<td>Improvement in 2 of 5 primary outcomes in the intervention group vs the control group (OR 2.63 (95% CI 1.57 to 4.42) to 15.7% for adaptation of antibiotic dose; OR 7.32 (95% CI 2.09 to 25.7). (partly effective)</td>
<td>Not reported</td>
</tr>
<tr>
<td>Van Kasteren et al, 2005</td>
<td>Optimising antibiotics policy</td>
<td>Interrupted time series design (13 hospitals; 1763 procedures before/2050 after).</td>
<td>Multifaceted intervention (performance feedback and implementation of national clinical practice guidelines).</td>
<td>Improvement in all 4 outcome measures (costs excluded) (varying from 12.4% for timing to 56% for antibiotic choice). (mostly effective)</td>
<td>No effect on overall SSI rates (not effective)</td>
</tr>
<tr>
<td>Bakker et al, 2008</td>
<td>Treatment of DM2</td>
<td>Uncontrolled before and after study (70 patients).</td>
<td>Medical doctors were instructed to strictly adhere to the guideline.</td>
<td>Not measured</td>
<td>Improvement in 6 of 7 outcome measures (eg, lowering HbA1c; decrease 1.7% and body weight (decrease 3.8 kg). (mostly effective)</td>
</tr>
<tr>
<td>Bekkering et al, 2005a, 2005b</td>
<td>Management of low-back pain</td>
<td>Cluster RCT; practices were randomised, block-randomisation (113 physiotherapists; 88 practices).</td>
<td>A multifaceted active strategy (dissemination of guideline and active training strategy consisting of education, discussion, role-playing, feedback, reminders) vs standard dissemination.</td>
<td>Improvement in all 4 outcome measures in active strategy group compared with standard dissemination group. OR from 1.99 (95% CI 1.06 to 3.72) to 12.7 for setting functional treatment goals to 3.95 (95% CI 1.35 to 9.55) for giving adequate patient information. Adherence to all four criteria also improved in active strategy group (42% vs 30%; OR 2.05; 95% CI 1.15 to 3.65). (mostly effective)</td>
<td>No sign. difference between groups. Improvement in 2 of 3 primary outcome variables (physical functioning from 38 to 20 (scale 0–100) in active strategy group and from 40.5 to 17.5 in standard group and pain from 7.0 to 2.0 (scale 0–10) in both groups) in first 12 weeks. (mostly effective)</td>
</tr>
<tr>
<td>De Laat et al, 2006a, 2007a</td>
<td>Prevention and treatment of pressure ulcers</td>
<td>Uncontrolled before and after study (process of care: T0: 657; T1: 735; T2: 755 patients and patient outcomes: 399 patients).</td>
<td>Guideline was introduced in staff meeting, announcement in several newspaper and the introduction of pressure reducing viscoelastic foam mattresses.</td>
<td>Improvement in inadequate prevention (from 19% to 4% after 4 months and to 6% after 11 months) and in inadequate treatment (from 60% to 31%). (most effective)</td>
<td>Improvement in both patient outcome measures (incidence of pressure ulcers decreased from 54 to 32 per 1000 patient days; pressure ulcer free time increased from 12 to 19 days). (mostly effective)</td>
</tr>
<tr>
<td>Van der Sanden et al, 2005a</td>
<td>Management of asymptomatic impacted lower third molars</td>
<td>Cluster RCT; GPs were randomised (92 GPs: 46/C. 46).</td>
<td>A multifaceted intervention (ie, feedback, reminders and an interactive meeting) vs no intervention.</td>
<td>Increased knowledge of dentists in IG compared with CG. No improvement in other outcome measure (referral rates). (partly effective)</td>
<td>Not measured</td>
</tr>
</tbody>
</table>
limited. Therefore, we decided to include uncontrolled before and after studies as well, while taking into account the weaknesses of these designs. However, our results demonstrate that these studies tended to have more positive results than studies using more robust designs. Because of the relatively small number of uncontrolled before and after studies in our review, our conclusions regarding the effectiveness of guidelines were not much affected by it. Nevertheless, this finding highlights the need for well-designed studies measuring the effectiveness of Dutch evidence-based guidelines.

Second, the studies included in our review were very heterogeneous, not allowing pooling of the results. On the other hand, heterogeneity of studies may increase the generalisability of findings as a wider range of different settings, study populations and behaviours are included. Third, to determine effectiveness of guidelines on quality of care, we counted the number of measures in each study that showed a significant result. In this assessment, we did not take into account the effect sizes of the individual measures (equal weights are given to improvements of 1 or 70%). As an alternative, we categorised the effectiveness of an intervention in mostly, partly and not effective, which may provide more insight than a dichotomy.

In conclusion, there is a huge misbalance between the number of guidelines developed and the number of high-quality studies that assess their effectiveness. Despite this, our review demonstrates that Dutch evidence-based guidelines can be effective in improving the process and structure of care. Evidence on the effectiveness of guidelines on patient outcomes is less convincing. The variation in effects across recommendations suggests that it is useful to focus on recommendations when analysing barriers to guideline adherence and to design implementation strategies tailored to individual recommendations instead of to the guideline as a whole. Further research is needed to determine which factors linked to the guideline and its specific recommendations are important in predicting guideline utilisation and improved patient outcomes.

Acknowledgements: The authors wish to thank Rikke Deurenberg, information specialist at the Dutch Institute for Healthcare Improvement (CBO), for developing and performing the search strategy.

Competing interests: None.

REFERENCES

Table 3 Characteristics of included studies (n = 20) and effects on quality of care

<table>
<thead>
<tr>
<th>Process/structure of care</th>
<th>Patient outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not effective</td>
</tr>
<tr>
<td></td>
<td>Not effective</td>
</tr>
<tr>
<td>Type of healthcare provider</td>
<td>General practitioners</td>
</tr>
<tr>
<td>Medical specialists</td>
<td>0</td>
</tr>
<tr>
<td>Other healthcare providers</td>
<td>0</td>
</tr>
<tr>
<td>Type of guideline</td>
<td>National</td>
</tr>
<tr>
<td>Regional/local</td>
<td>1</td>
</tr>
<tr>
<td>Subject of study</td>
<td>Preventive care</td>
</tr>
<tr>
<td>Diagnosis and treatment</td>
<td>1</td>
</tr>
<tr>
<td>Prevention/diagnosis and treatment</td>
<td>1</td>
</tr>
<tr>
<td>Study design</td>
<td>Cluster RCT</td>
</tr>
<tr>
<td>Controlled before and after study</td>
<td>0</td>
</tr>
<tr>
<td>Interrupted time series design</td>
<td>0</td>
</tr>
<tr>
<td>Uncontrolled before and after study</td>
<td>0</td>
</tr>
<tr>
<td>Implementation strategy</td>
<td>Single</td>
</tr>
<tr>
<td>Multifaceted</td>
<td>2</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial.

Mostly effective, significant effect on more than half of the outcome measures was reported; partly effective, no significant effect on any of the outcome measures was reported; partly effective, significant effect on half or less than half of the outcome measures was reported.
Original research

45. Kampwesin PW, Jacobs EMG, Mol JJ, et al. Resultaten van het invoeren van diagnostiek naar longembolie volgens de CBDO-richtlijn in een algemeen opzichtersziekenhuis [Results from a general training hospital for the implementation of a diagnostic workup for pulmonary embolism according to the Dutch Institute for Practice Health Care Improvement (In Dutch)]. Ned Tijdschr Geneeskd 2002;146:2083–6.

APPENDIX A: SEARCH STRATEGY

Description: A detailed description of the search strategy used in this review is available online.