Inside Guidelines

Comparative analysis of recommendations and evidence in diabetes guidelines from 13 countries

JAKO S. BURGERS, MD
JULIA V. BAILEY, MBBS, MRCGP
NIEK S. KLAXINGA, MD, PHD
AKKE K. VAN DER BIJ, MSC

OBJECTIVE — To compare guidelines on diabetes from different countries in order to examine whether differences in recommendations could be explained by use of different research evidence.

RESEARCH DESIGN AND METHODS — We analyzed 15 clinical guidelines on type 2 diabetes from 13 countries using qualitative methods to compare the recommendations and bibliometric methods to measure the extent of overlap in citations used by different guidelines. A further qualitative analysis of recommendations and cited evidence for two specific issues in diabetes care explored the apparent discrepancy between recommendations and evidence.

RESULTS — The recommendations made in the guidelines were in agreement about the general management of type 2 diabetes, with some important differences in treatment details. There was little overlap in evidence cited by the guidelines, with 18% (185/1,033) of citations shared with any other guideline, and only 10 studies (1%) appearing in six or more guidelines. The measurable overlap in evidence between guidelines increases if multiple publications from the same study and the use of reviews are taken into account. Research originating from the U.S. predominated (40% of citations); however, nearly all (11/12) guidelines were significantly more likely to cite evidence originating from their own countries.

CONCLUSIONS — Despite the variation in cited evidence and preferential citation of evidence from a guideline’s country of origin, we found a high degree of international consensus in recommendations made for the clinical care of type 2 diabetes. The influence of professional bodies such as the American Diabetes Association may be an important factor in explaining international consensus. Globalization of recommended management of diabetes is not a simple consequence of the globalization of research evidence.

Diabetes Care 23:1933–1939, 2002

O

ver the past 20 years, clinical guidelines have been developed to bridge the gap between research and practice (1). There has been a concerted effort to base clinical decisions on research evidence (2) and, particularly through the Cochrane collaboration, to make this evidence available globally (3). Guideline development groups aim to use the totality of relevant research evidence to formulate recommendations (4). Since bibliographic databases (for example Medline and Embase) are easily available, one might expect that this would lead to international consensus on the evidence chosen to underpin recommendations for clinical care and a consequent convergence of recommendations made in guidelines.

Nevertheless, recommendations often differ in guidelines on the same topic, particularly when evidence for treatment decisions is weak. For example, Eisinger et al. (5) found substantial differences between recommendations from the U.S. and France about prophylactic mastectomy or oophorectomy in high-risk women. Differences were attributed to cultural variation in ideas about patient autonomy and involvement in health care, differing national views on esthetics of the breast, and fertility. Even where there is good trial evidence, recommendations vary. For instance, analysis of hypertension guidelines from New Zealand, U.S., Canada, U.K., and the World Health Organization showed wide variation in the criteria for blood pressure treatment decisions (6). Differences persisted between more recent editions of national hypertension guidelines, even with more systematic and transparent methods of guideline development (7).

It is evident that there are disparities in recommendations in guidelines for a range of different clinical conditions. Investigators hypothesize that differences are due to insufficient evidence (6,8,9), differing interpretations of evidence (10), unsystematic guideline development methods (11,12), the influence of professional bodies (13), cultural factors such as
differing expectations of apparent risks and benefits (5,6), socioeconomic factors, or characteristics of health care systems (14).

In this study, we compared recommendations among a range of guidelines on the management of type 2 diabetes and analyzed to what extent the variation (or concordance) among recommendations was explained by the evidence cited in the guidelines.

## RESEARCH DESIGN AND METHODS

### Selection of guidelines

We applied the Institute of Medicine’s definition of clinical guidelines: “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances” (15). Systematic reviews and evidence reports that did not contain specific recommendations were not included in this study. Because of the large number of clinical issues related to diabetes, the selection of guidelines was limited to two areas: 1) ambulatory or outpatient care, excluding guidelines exclusively covering type 1 diabetes, complications of diabetes that need specialist care (retinopathy, diabetic foot, nephropathy, and neuropathy), and gestational diabetes; and 2) treatment of diabetes, excluding guidelines on prevention and diagnosis.

The sample consisted of a total of 15 guidelines for the clinical care of type 2 diabetes (Table 1) representing the national guidelines of the Appraisal of Guidelines for Research and Evaluation (AGREE) collaboration. This international group of researchers has investigated variation between guidelines and guideline development models with the aim of advising the European Commission on guideline development, dissemination, and implementation. The east London guideline was chosen because there were no national English guidelines available. Two French guidelines were complementary and were analyzed as one guideline. The guidelines from Australia, New Zealand, Canada, and the U.S. were

### Table 1—Description of selected guidelines

<table>
<thead>
<tr>
<th>Country (ID code)</th>
<th>Organization responsible for guideline development</th>
<th>Title in English</th>
<th>Year of publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia (AU)</td>
<td>NSW (New South Wales) Health Department</td>
<td>Improving diabetes care and outcomes. Principles of care and guidelines for the clinical management of diabetes mellitus</td>
<td>1996</td>
</tr>
<tr>
<td>Canada (CA)</td>
<td>Canadian Medical Association</td>
<td>Clinical practice guidelines for the management of diabetes in Canada</td>
<td>1998</td>
</tr>
<tr>
<td>Denmark (DK)</td>
<td>Danish College of General Practitioners</td>
<td>Non insulin demanding diabetes - NIDDM. A practical guidance for therapists</td>
<td>1998</td>
</tr>
<tr>
<td>Finland (FI)</td>
<td>Finnish Diabetes League</td>
<td>Type II diabetes clinical guideline</td>
<td>1994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Strategy for management of type 2 diabetics, excluding management of complications</td>
<td>2000</td>
</tr>
<tr>
<td>Italy (IT)</td>
<td>Italian Society for Diabetology</td>
<td>Diabetes mellitus. Practical guide for diagnosis and treatment</td>
<td>1997</td>
</tr>
<tr>
<td>The Netherlands (NL2)</td>
<td>Dutch College of General Practitioners (NHG)</td>
<td>Practice guideline diabetes mellitus</td>
<td>1999</td>
</tr>
<tr>
<td>New Zealand (NZ)</td>
<td>New Zealand Guidelines Group</td>
<td>Guidelines for the management of core aspects of diabetes care</td>
<td>1999</td>
</tr>
<tr>
<td>Spain (SP)</td>
<td>Catalan Society of Primary Care</td>
<td>Guideline on treatment of diabetes mellitus type 2 in primary care</td>
<td>1996</td>
</tr>
<tr>
<td>Switzerland (SW)</td>
<td>University Hospital of Geneva</td>
<td>Detection of diabetes mellitus. Guidelines for the outpatient’s clinic</td>
<td>1996</td>
</tr>
<tr>
<td>USA (US1)</td>
<td>American Diabetes Association</td>
<td>Standards of medical care for patients with diabetes mellitus</td>
<td>2000</td>
</tr>
<tr>
<td>USA (US2)</td>
<td>Institute for Clinical System Improvements</td>
<td>Management of type 2 diabetes mellitus</td>
<td>2000</td>
</tr>
</tbody>
</table>
identified through a web-based search and consultation with colleagues. Four guidelines (CA, NL2, US1, and US2) were updated versions of earlier guidelines. Two guidelines (CA and DK) were funded by pharmaceutical companies, and the others were funded by national or regional government agencies (EN, FR, NZ, SC, and SP) or state health care systems (AU and US2), national professional organizations (FI, IT, NL2, and US1), or hospitals (NL1, SW).

Selection of comparable sections
Because the guidelines varied in their scope, we selected sections that covered the treatment and monitoring of hyperglycemia and cardiovascular risk.

The guidelines were in seven different languages. Members of the study team translated those guidelines written in French or Dutch; relevant sections of guidelines in Finnish, Danish, and Spanish were translated by guideline developers in their respective countries. The Italian guideline was excluded from the analysis of recommendations because of its length and lack of structure.

Extraction and comparison of recommendations
We defined recommendations as any statements that promote or advocate a particular course of action in clinical care. Two investigators with medical training, working independently, extracted the recommendations. We resolved discrepancies through discussion within the study team. A panel of four investigators (J.S.B., J.V.B., G.F., and N.S.K.) judged the extent of accordance or discordance of recommendations across guidelines.

Examination of link between recommendations and citations
To explore the discrepancy between disparate citations and largely concordant recommendations, we purposively selected (16) two areas for further analysis: use of metformin in obese patients and self-monitoring of blood glucose. We selected citations that were explicitly linked to the recommendations or listed at the end of relevant sections and compared citations between guidelines. For each citation, we tabulated the type of study, country of origin, study subjects, conclusions, and any recommendations made by the authors. Where secondary citations were used (i.e., meta-analyses, systematic reviews, or other guidelines), we included the evidence cited by these documents. We compared the publication dates of citations and the dates of the latest evidence cited by guidelines (censoring dates). We did not appraise the quality of the studies but examined the consistency between the study conclusions and recommendations made in the guidelines.

RESULTS — Guidelines varied considerably in length (range 3–350 pages), format, and number of references (Table 2). Nine guidelines linked their recommendations to citations; four of these (FR, SC, CA, and US2) also used grading systems to appraise the evidence.

Table 2—Length of guidelines, number of references, and shared references

<table>
<thead>
<tr>
<th>ID code</th>
<th>Number of references</th>
<th>Number of pages</th>
<th>Number of references linked to relevant sections</th>
<th>% (n) of shared references</th>
<th>Weighted shared score (ranking)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR</td>
<td>590</td>
<td>312</td>
<td>422</td>
<td>20.4 (86)</td>
<td>16.0 (9)</td>
</tr>
<tr>
<td>CA</td>
<td>302</td>
<td>29</td>
<td>158</td>
<td>46.2 (73)</td>
<td>19.6 (6)</td>
</tr>
<tr>
<td>NL1</td>
<td>246</td>
<td>164</td>
<td>127</td>
<td>42.5 (54)</td>
<td>18.8 (7)</td>
</tr>
<tr>
<td>US1</td>
<td>233†</td>
<td>93†</td>
<td>171</td>
<td>42.7 (73)</td>
<td>18.1 (8)</td>
</tr>
<tr>
<td>IT</td>
<td>218</td>
<td>350</td>
<td>83</td>
<td>31.3 (26)</td>
<td>15.2 (11)</td>
</tr>
<tr>
<td>NL2</td>
<td>190</td>
<td>18</td>
<td>132</td>
<td>44.7 (59)</td>
<td>20.6 (5)</td>
</tr>
<tr>
<td>SP</td>
<td>95</td>
<td>85</td>
<td>73</td>
<td>39.7 (29)</td>
<td>15.5 (10)</td>
</tr>
<tr>
<td>SC</td>
<td>77</td>
<td>21</td>
<td>56</td>
<td>42.9 (24)</td>
<td>18.1 (8)</td>
</tr>
<tr>
<td>US2</td>
<td>67</td>
<td>92</td>
<td>57</td>
<td>63.2 (36)</td>
<td>35.3 (1)</td>
</tr>
<tr>
<td>AU</td>
<td>65</td>
<td>92</td>
<td>12</td>
<td>66.6 (8)</td>
<td>24.3 (3)</td>
</tr>
<tr>
<td>NZ</td>
<td>44</td>
<td>19</td>
<td>25</td>
<td>56.0 (14)</td>
<td>33.5 (2)</td>
</tr>
<tr>
<td>EN</td>
<td>40</td>
<td>36</td>
<td>30</td>
<td>53.3 (16)</td>
<td>21.8 (4)</td>
</tr>
<tr>
<td>SW</td>
<td>2</td>
<td>3</td>
<td>Excluded from further analysis</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>FI</td>
<td>1</td>
<td>55</td>
<td>Excluded from further analysis</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>DK</td>
<td>0</td>
<td>15</td>
<td>Excluded from further analysis</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Total</td>
<td>2,170</td>
<td>1,344</td>
<td>1,346</td>
<td>37.0 (498)</td>
<td></td>
</tr>
</tbody>
</table>

*Weighted shared score = number of shared references × 100, divided by maximum possible number of shared references according to publication dates of the guideline and its linked references; †for the 11/42 selected ADA position or consensus statements only.
Guidelines varied in their coverage. For example, the Danish and Spanish guidelines allocated >10% of the text to detailed dietary recommendations, whereas the English and New Zealand guidelines only made a few general statements. Guidelines also varied in their scope; for example, the Scottish, Australian, and Dutch guidelines (NL1) did not cover drug treatment of hyperglycemia.

**Comparison of recommendations**

The guidelines largely agreed on general management of patients with type 2 diabetes, which was covered by the following recommendations: 1) all patients should be offered dietary advice and overweight/obese patients should be offered weight management advice; 2) the diet should be low in sugar, fat content, and overall calories and should be combined with exercise; 3) all patients should stop smoking to reduce cardiovascular risk; 4) patient education is necessary to promote good diabetes control; 5) poor glycemic control should be tackled initially with diet alone, followed by oral medication and insulin if necessary, unless the patient is acutely unwell; 6) sulfonylureas or biguanides are recommended in patients with normal BMI, and metformin is recommended in obese patients; 7) a second oral agent should be added to maximum doses of an initial agent in case of poor glycemic control; 8) HbA1c is suitable for long-term monitoring and should be <8%; 9) if on insulin, self-monitoring of blood glucose is recommended; 10) screening and treatment of raised blood pressure, microalbuminuria, and hyperlipidemia is recommended; 11) ACE inhibitors are recommended in patients with hypertension and renal disease; and 12) aspirin is recommended for secondary prevention of cardiovascular disease.

Differences between the recommendations were found in the following areas: 1) length of trial of diet and exercise before oral treatment ranged from 2 to 9 months—some guidelines recommended a longer period in obese than nonobese patients; 2) BMI used to define obesity ranged from 25 to 30 kg/m²; 3) widely varying indications were suggested for the use of α-glucosidase inhibitors; 4) there was no consensus on the value or indications of combination therapy with oral hypoglycemics and insulin; 5) target HbA1c ranged from 6.5 to 7.5%, and target blood pressure ranged from <130/80 to <160/90 mmHg; 6) frequency of monitoring HbA1c and blood pressure ranged from one to four times a year and one to six times a year; 7) there was no consensus on self-monitoring of blood glucose in patients on diet alone or on oral medication; 8) there was no consensus on the first-line drug for raised blood pressure; 9) widely differing opinions were given on the value of aspirin use as primary prevention in high-risk patients; 10) widely differing targets were given for lipid control (e.g., total cholesterol 4.5—6.5 mmol/l)—there was no consensus on the use of absolute cardiovascular risk or isolated lipid levels for treatment decisions; and 11) routine annual electrocardiogram was recommended by half of the guidelines, whereas others recommended electrocardiogram for specific indications or did not mention it.

**Comparison of linked citations**

We selected a total of 1,346 references from 12 guidelines (Table 2); 1,033 of these were different citations. Only 18% (185/1,033) of the unique citations were shared with any of the other 11 guidelines. Considering all of the references made in the guidelines, on average of 37% (498/1,346) of these were shared with any other guideline (range 20—67%). The Diabetes Control and Complications Trial (DCCT) (17) was most frequently cited (in 11 guidelines). A randomized controlled trial addressing intensive insulin therapy with patients with type 2 diabetes was cited by eight guidelines (18). If all 45 publications of the American Diabetes Association (ADA) were analyzed as one document, it would be shared between eight guidelines. Two studies (one randomized controlled trial and one cohort study) were shared among seven guidelines, and six trials were shared among six guidelines. Six guidelines referred to the WHO St. Vincent Declaration. Four of the 12 most frequent citations were from the U.S., 3 were from the U.K. (all three U.K. Prospective Diabetes Study publications), 2 were from Israel, 1 each was from Finland and Japan, and 1 was a WHO document.

The largest proportion of lead authors of papers cited in the guidelines (40%) originated from the U.S. (Table 3). All guidelines, except the Australian, cited a significantly higher proportion of studies from authors of their own countries than

---

**Table 3—Countries of authors of citations (%)**

<table>
<thead>
<tr>
<th>ID code</th>
<th>AU/NZ</th>
<th>CA</th>
<th>UK/IR*</th>
<th>FR</th>
<th>IT</th>
<th>NL</th>
<th>SP</th>
<th>US</th>
<th>SCAN+</th>
<th>Other</th>
<th>Multinational</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU</td>
<td>8</td>
<td>—</td>
<td>17</td>
<td>8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>NZ</td>
<td>16</td>
<td>—</td>
<td>40</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>16</td>
<td>12</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>CA</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>45</td>
<td>17</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>EN</td>
<td>10</td>
<td>—</td>
<td>37</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>27</td>
<td>23</td>
<td>—</td>
<td>3</td>
</tr>
<tr>
<td>SC</td>
<td>2</td>
<td>—</td>
<td>39</td>
<td>4</td>
<td>—</td>
<td>2</td>
<td>—</td>
<td>—</td>
<td>23</td>
<td>23</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>FR</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>11</td>
<td>4</td>
<td>3</td>
<td>—</td>
<td>—</td>
<td>38</td>
<td>11</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>IT</td>
<td>—</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>30</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>41</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>NL1</td>
<td>—</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>3</td>
<td>18</td>
<td>1</td>
<td>32</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>NL2</td>
<td>3</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>36</td>
<td>—</td>
<td>24</td>
<td>12</td>
<td>10</td>
<td>—</td>
<td>0</td>
</tr>
<tr>
<td>SP</td>
<td>3</td>
<td>—</td>
<td>11</td>
<td>1</td>
<td>—</td>
<td>1</td>
<td>11</td>
<td>51</td>
<td>11</td>
<td>7</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>US1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>—</td>
<td>59</td>
<td>18</td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>US2</td>
<td>2</td>
<td>2</td>
<td>14</td>
<td>2</td>
<td>—</td>
<td>—</td>
<td>2</td>
<td>63</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>2.6</td>
<td>2.7</td>
<td>11.8</td>
<td>5.2</td>
<td>4.7</td>
<td>7.1</td>
<td>10</td>
<td>94</td>
<td>11.8</td>
<td>7.8</td>
<td>1.3</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Bold types contain percent of citations of studies from authors of the own country; *IR, Ireland; +SCAN, Scandinavia.
the origin overall of citations in the database ($P < 0.02$). Citations in the English, Scottish, and New Zealand guidelines were predominantly from the U.K., and citations in all other guidelines, except the Dutch general practice guideline (NL2), were predominantly from the U.S.

Sixteen of the total 1,033 citations (2%) were meta-analyses, 89 (9%) were reviews or overviews (including 4 systematic reviews), and 55 (5%) were existing guidelines (including practical guides and clinical practice recommendations) or consensus statements. Twenty of these 160 secondary citations (13%) were ADA publications.

Examination of link between recommendations and citations (case studies)

Use of metformin in obese patients
Eleven guidelines covered the use of oral medication. Nine explicitly recommended metformin as a first choice oral treatment for hyperglycemia in the obese, while the Canadian and US1 guidelines recommended tailoring treatment for the individual. We compared the citations from six guidelines, and the others had no citations linked to their recommendations on use of metformin (online appendix, http://care.diabetesjournals.org). There was little overlap in the 20 citations given: 1 (U.K. Prospective Diabetes Study [UKPDS] 34) was shared by four of five guidelines with a censoring date that would allow use of this paper (19). The UKPDS 13 paper was shared by three of six guidelines (20), and three other citations were shared by two. Over half of the linked citations (11/20) were randomized controlled trials; 1 was a meta-analysis, and the remainder were nonsystematic reviews. All studies concluded that metformin was useful in obese patients. While the choice of citations varied, publications from one trial (UKPDS) predominated and each guideline cited at least one publication that explicitly supported the recommendation.

Self-monitoring of blood glucose
Nine guidelines covered self-monitoring and were unanimous in recommending the self-monitoring of blood glucose in type 2 diabetes treated with insulin. We compared the citations from seven guidelines (online appendix). Only two citations were present in more than one guideline: the DCCT trial (17) was cited in two and the ADA consensus statement (21) was shared by four. However, when we considered the primary studies in systematic reviews, meta-analyses, or guideline and consensus documents, the overlap between citations increased substantially: 17 of 33 references were then shared by at least two guidelines. For example, the Dutch and French guideline had seven citations in common by virtue of a systematic review conducted by Faas et al. (22). Of the seven citations that specifically addressed self-monitoring in type 2 diabetes, five (two randomized controlled trials, one cross-sectional study, one review, and one comment) concluded that there was no evidence to support its use. The two supportive citations were guidelines (an ADA consensus statement and a Canadian guideline).

CONCLUSIONS — This is the first study comparing both guideline recommendations and cited evidence across national guidelines. Our bibliometric analysis included >1,000 citations. We minimized selection and observer bias by prospective choice of inclusion criteria for recommendations and citations and independent extraction by two researchers.

We found a high degree of international consensus on the clinical care of people with type 2 diabetes, despite differences in detailed recommendations. This was in contrast to what we expected, considering the range of influences on the guideline development process and the variation in organization of care and health care system among countries (23). The citations linked to and presumably justifying the guideline recommendations were widely disparate. The influence of large pragmatic treatment trials (e.g. DCCT [17]) and UKPDS studies (19,24,25) was nevertheless visible in most of the guidelines and apparent even in guidelines without references.

Little use was made of systematic reviews (for example Cochrane reviews), which is consistent with the findings of Silagy et al. (26). National guidelines were significantly more likely to cite research from investigators from the same country, explaining some of the variation in citations between guidelines. Others have found that local sources of evidence are overrepresented in guidelines (27) and that the results of trials conducted in the same country may be given more prominence (28).

We used the case studies to generate hypotheses to explain the small degree of overlap in citations between guidelines. Recommendations for the use of metformin in obese patients drew on supportive trial and review evidence. The different studies linked to these concordant recommendations often had similar conclusions. We also observed a consensus in recommendations for the use of self-monitoring of blood glucose, despite citation of evidence that did not support this position. The overlap in evidence would have been larger if we had aggregated citations from the same study (e.g., UKPDS) and if we had included the primary citations made within reviews and meta-analyses. Even taking this into account, the evidence cited in type 2 diabetes guidelines largely does not overlap. Therefore, we hypothesize that there are other potential influences on guideline developers. For example, the recommendations of the ADA strongly influenced the other guidelines on diabetes, with the exception of the English and Scottish. Similarly, Littlejohns et al. (13) found that professional opinion expressed in a consensus statement from the Royal College of General Practitioners and the Royal College of Physicians influenced the recommendations made in nine U.K. guidelines for the treatment of depression in primary care.

Guideline development is a social as well as technical process that is affected by access to and choice of research evidence and decisions about the interpretation of evidence and formulation of recommendations (29–31). Our study suggests that research evidence is not necessarily the most powerful influence on the content of recommendations in the current generation of guidelines on the management of type 2 diabetes. Guideline developers might first aim to achieve consensus about recommendations and then switch to the evidence as a rhetorical device to support decisions post hoc. Thus, the relationship between choice and interpretation of research evidence and the formulation of guideline recommendations is neither necessarily linear nor unidirectional. However, we are not suggesting a complete epistemological divide between evidence as represented by research papers and guidelines recommendations. As Greenhalgh and McCor-
high-quality systematic reviews may sup-
understood. The growing availability of
ADA, on national guidelines is not well
conferences, pharmaceutical companies,
evidence and the role of international
nants of guidelines require further inves-
tions and evidence in diabetes guidelines.
Two case studies exploring in more detail
large trials should be given more weight
relationship between guideline recommen-
dations and the cited evidence.
1998. In our analysis of the citations, we
dealt with this confounding factor by cor-
r recting for publication dates of the guide-
lines and the cited evidence.
Finally, analysis of shared references is a blunt instrument for exploring the re-
lationship between guideline recommenda-
tions and evidence. High-quality and
large trials should be given more weight
in the analysis. That is why we included
two case studies exploring in more detail the relationship between recommenda-
tions and evidence in diabetes guidelines.
Other clinical issues will need this kind of
analysis to test the generalizability of our
findings.
The process of formulating guideline
recommendations and the social determin-
ants of guidelines require further inves-
tigation. Decisions about choice of
 evidence and the role of international
conferences, pharmaceutical companies,
and opinion-forming bodies, such as the
ADA, on national guidelines is not well
understood. The growing availability of
high-quality systematic reviews may sup-
port more uniformity in the use of re-
search evidence in guidelines (33). Neverthe-
less, guidelines go beyond simple
reviews of available evidence and nec-
essarily reflect value judgements in
considering all the issues relevant to clini-
cal decision making. Transparency by
guideline developers about how their
judgements have been made would allow
clinicians to evaluate the applicability of
guideline recommendations to their own
health care context and to individual pa-

Acknowledgments — The research was
funded by a grant From the EU BIOMED2 Pro-
gramme (BMH4-98-3669).
We thank Jean Ramsay, Rachel Kerr, and
Emily Hallgarten for entering data into the re-
ference manager database, Sudip Nandy for pi-
loting the study methodology, and Rob
Dijkstra for contributing to the data collection
of the Dutch recommendations.

APPENDIX
The AGREE Collaboration
The following individuals participated in
The AGREE Collaboration: José Asua, MD,
PhD, Basque Office for Health Technol-
ology Assessment, Spain; Anne Bataillard,
MD, Fédération Nationale des Centers de
Lutte contre le Cancer, Paris, France;
Melissa Brouwers, PhD, McMaster Uni-
versity and Cancer Care Ontario, Hamil-
ton, ON, Canada; George Brownaw, MD,
Hamilton Regional Cancer Center, Hamil-
ton, Canada; Jako Burgers, MD, Center
for Quality of Care Research, University
Medical Center Nijmegen, the Nether-
lands; Bernard Burnand, MD, MPH, Insti-
tut Universitaire de Médecine Sociale et
Préventive, Lausanne, Switzerland; Fran-
çois Cluzeau, MSc, PhD, St George’s
Hospital Medical School, London, U.K.;
Isabelle Durand-Zaleski, PhD, Hôpital
Henri Mondor, Cedez, France; Pierre Du-
rieux, MD, Hôpital Européen Georges
Pompidou, Paris, France; Cindy Far-
quhar, MD, PhD, New Zealand Guide-
lines Group, Auckland, New Zealand;
Gene Feder, MD, FRCG, Barts and The
London, Queen Mary’s School of Medi-
cine and Dentistry, University of London,
U.K.; Béatrice Fervers, MD, Fédération
Nationale des Centers de Lutte contre le
Cancer, Paris, France; Roberto Grilli, MD,
Agenzia Sanitaria Regionale, Bologna, It-
aly; Jeremy Grimshaw, MB, PhD, Ottawa
Health Services Research Institute, Ot-
tawa, Canada; Richard Grol, PhD, Center
for Quality of Care Research, University
Medical Center Nijmegen, the Nether-
lands; Steven Hanna, PhD, McMaster
University, Hamilton, ON, Canada; Pieter
ten Have, MD, Dutch Institute for Health-
care Improvement CBO, Utrecht, the
Netherlands; Rod Jackson, PhD, Effective
Practice Institute, University of Auckland,
New Zealand; Albert Jovell, MD, PhD,
Fundació Biblioteca Josep Laporte, Barce-
lona, Spain; Niek Klazinga, MD, PhD, Ac-
ademic Medical Center, University of
Amsterdam, the Netherlands; Finn Krist-
tensen, MD, PhD, Danish Institute for
Health Technology Assessment, Copen-
hagen, Denmark; Peter Littlejohns,
MBBS, MD, National Institute for Clinical
Excellence, London, U.K.; Fia Bruun
Madsen, Danish Institute for Health
Technology Assessment, Copenhagen,
Denmark; Marjukka Makela, MD, PhD,
MSc, Finnish Office for Health Care Tech-
nology Assessment, Helsinki, Finland;
Juliet Miller, MA, MBA, Scottish Intercol-
legiate Guidelines Network (SIGN), Edin-
burgh, U.K.; Günter Ollenschläger, MD,
PhD, Agency for Quality in Medicine,
Cologne, Germany; Camilla Palmhøj-
Nielsen, Danish Institute for Health
Technology Assessment, Copenhagen,
Denmark; Loes Pijnenburg, MD, PhD,
Dutch College of General Practitioners,
Utrecht, the Netherlands; Safia Qureshi,
PhD, Scottish Intercollegiate Guidelines
Network (SIGN), Edinburgh, U.K.; Rosa
Rico-Irurruz, MD, MSc, Basque Office for
Health Technology Assessment, Spain;
Kitty Rosenbrand, MD, Dutch Institute for
Healthcare Improvement CBO, Utrecht,
the Netherlands; Jean Slutsky, Agency for
Healthcare Research and Quality, Rock-
ville, MD; John-Paul Vadder, MD, MPH,
Institut Universitaire de Médecine Sociale
et Préventive, Lausanne, Switzerland;
and Joost Zaai, MD, PhD, Center for Quality
of Care Research, University Medical
Center Nijmegen, the Netherlands.

References
1. Bero LA, Grilli R, Grimshaw JM, Harvey E,
Osman AD, Thomson MA: Closing the
gap between research and practice: an
overview of systematic reviews of inter-
ventions to promote the implementation
of research findings: the Cochrane Effec-
tive Practice and Organization of Care Re-
2. Sackett DL, Rosenberg WM, Gray JAM,
Haynes RB, Richardson WS: Evidence-