Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Title

Assessing the quality of dental clinical practice guidelines

Short title

Quality of dental clinical practice guidelines

Authors

Suhaym Mubeen\textsuperscript{a}, Kunal Patel\textsuperscript{b}, Ziba Cunningham\textsuperscript{c}, Niamh O'Rourke\textsuperscript{c}, Nikolaos Pandis\textsuperscript{d}, Martyn.T. Cobourne\textsuperscript{a}, Jadbinder Seehra\textsuperscript{a,*}

Affiliations

\textsuperscript{a} Department of Orthodontics, King’s College London Dental Institute, Floor 22, Guy’s Hospital, Guy’s and St Thomas NHS Foundation Trust, London, SE1 9RT, United Kingdom

\textsuperscript{b} Department of Paediatric Dentistry, King’s College London Dental Institute, Kings College Hospital NHS Foundation Trust, Bessemer Road, London, SE5 9RS, United Kingdom

\textsuperscript{c} Department of Orthodontics, King’s College London Dental Institute, Kings College Hospital NHS Foundation Trust, Bessemer Road, London, SE5 9RS, United Kingdom

\textsuperscript{d} Department of Orthodontics and Dentofacial Orthopedics, Dental School/Medical Faculty, University of Bern, Freiburgstrasse 7 CH-3010, Bern, Switzerland

*Author for correspondence

Mr Jadbinder Seehra
Department of Orthodontics, King’s College London Dental Institute, Floor 22, Guy’s Hospital, London, WC2R 2LS
United Kingdom
Email: jadbinder.seehra@nhs.net
Abstract

Objectives: The primary aim of this study was to evaluate the quality of published dental clinical guidelines using the AGREE II instrument.

Methods: Online searching of a wide range of organisations (national and International) was undertaken to identify dental clinical practice guidelines published between 2000-2014. The quality of each included guideline was assessed in relation to the AGREE II instrument by four assessors independently. Inter-rater agreement was assessed. Descriptive statistics and both univariate and multivariate analyses were conducted.

Results: 162 guidelines were identified. The overall mean quality score was 51.9% (SD 13.3). There was variation in the reporting quality of individual domains with both Applicability (20.4%) and Editorial Independence (34.25%) poorly reported. Variation between the overall quality scores for guidelines produced by different dental specialities was evident. The quality of guidelines improved per publication year (β=0.76, 95% CI: 0.26, 1.26, p=0.003). Guidelines based on formal evidence (β=19.94, 95% CI: 15.25, 24.64, p=0.001) achieved higher quality scores.

Conclusion: Overall, the quality of clinical dental practice guidelines is suboptimal. There is variation in the overall quality, reporting of individual items and domains of the AGREE II instrument between different dental speciality clinical practice guidelines. Guidelines based on formal evidence achieved higher quality scores.

Keywords: Evidence-based dentistry; dentistry; Evidence-Based Medicine; AGREE; guidelines; clinical practice guidelines

Clinical Significance:

Clinicians should be aware of the variation in the quality of dental clinical guidelines in particular related to methodological rigour. The use of formal evidence may be a useful indicator of their quality prior to their implementation.
1. Introduction

Evidence based medicine has been defined as the integration of the best research evidence with clinical expertise and patient values [1]. Clinical practice guidelines can be a means to bridge the gap between research and healthcare provision [2]. The Institute of Medicine defines guidelines as systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances [3]. Whilst clinical practice guidelines have numerous benefits they may also negatively influence patient care or be of questionable applicability in dental practice [2, 4]. It is of paramount importance that guidelines are of sufficient quality to allow the implementation of clear and effective recommendations. whilst numerous instruments have been developed to assess the quality of guidelines [5], the AGREE II (Appraisal of Guidelines, Research and Evaluation II) is an internationally developed, validated, easy-to-use and transparent instrument [6]. Health care organisations which have included and assess their guidelines with the AGREE instrument include National Institute for Clinical Excellence (NICE), National Federation of Cancer Centres (FNCLCC), The Agency for Quality in Medicine in Germany (ÄZQ), Scottish Intercollegiate Guidelines Network (SIGN) and World Health Organisation [7].

Previous quality assessments of clinical practice guidelines in dentistry have been undertaken [8]. The majority of these studies have been limited to specific subspecialties such as cone beam computer tomography [9], orthodontics [10], paediatric dentistry [11], dental management of antithrombotic drug use [12] and common clinical procedures [13]. The results of these studies have identified that the reporting and quality of dental guidelines is lacking and inadequate in relation to the AGREE II instrument.

To date, no assessment of the quality of clinical practice guidelines in dentistry using the AGREE II has been undertaken with four reviewers as recommended by the AGREE collaboration [6,14,15]. In addition, characteristics that may influence quality have not been identified. The primary aim of this study was to evaluate the quality of published speciality dental clinical practice guidelines in relation to the AGREE II instrument. A secondary aim
was to identify factors associated with improved guideline reporting.

2. Materials and Methods

2.1 Information Sources and search strategy

An electronic literature search was undertaken to identify guidelines related to dentistry published between 2000 and 2014. The search was restricted to guidelines published in English and only interventions at the individual/patient level were included. Conference abstracts, non-English guidelines, laboratory based guidelines and those aimed at non-dental healthcare workers were excluded. A MEDLINE (Ovid®) database search was carried out on the 13th February 2015 using the terms described in Table 1. In addition, the TRIP (Turning Research Into Practice) database, National Institute For Health and Clinical Excellence (NICE) Evidence and US National Guideline Clearinghouse were searched using the search term (dent*) limiting results to 2000 to 2014. The TRIP database search was limited to “guidelines only”. The websites of national and international dental organisations were also searched to identify dental guidelines satisfying the inclusion criteria (Appendix I). A single author (SM) initially screened all potentially relevant dental guidelines. All identified guidelines were then independently screened by two authors (SM and JS). Any disagreements were resolved by discussion.

2.2 AGREE II instrument

The quality of the process and reporting of clinical guideline development of each guideline was assessed using the AGREE II instrument which consists of a twenty-three item checklist categorised into six domains (Scope and Purpose, Stakeholder involvement, Rigour of Development, Clarity of Presentation, Applicability and Editorial Independence). Each domain aims to measure a different aspect of guideline quality and identify potential biases [6,7,14,15] (Appendix II). Each of the AGREE II items are rated on a seven-point Likert scale ranging from ‘Strongly Agree’ to ‘Strongly Disagree’. A score is assigned based upon the
reporting of the item in relation to the full criteria or considerations, its level of completeness and quality of reporting.

2.3 Evaluation of guidelines

Four assessors evaluated the guidelines independently. Each assessor was calibrated in the use of the AGREE II instrument by completing the online training tool [15] and by completing a pilot of 5 guidelines. Any discrepancies or clarifications were discussed until a consensus was obtained. In addition, each guideline was assessed by referring directly to the associated explanation of each item as stated in the user manual. Guideline demographic data collected included: the development process classification (whether expert opinion, consensus based or formal evidence based) [9], the dental sub-specialty of each guideline, number of authors, continent of publication, identification as single-or multi-centre guideline development and whether the guideline was an update.

2.4 Statistical analysis

Inter-assessor reliability was assessed using Intra-class correlation coefficient (ICC). Descriptive statistics for individual reporting items for each dental guideline were calculated and converted to a percentage scale with 100% indicating the maximum score for all applicable items. Linear regression modeling was implemented with univariate analysis to identify characteristics associated with mean score; multivariate modeling was used to determine the adjusted effect on reporting quality score. Significant predictors identified during the univariate analysis were entered individually in the multivariate model. The final model was derived by comparing candidate models using the likelihood ratio test. A two-tailed p-value of 0.05 was considered statistically significant. Statistical analyses were performed with STATA® version 14.2 software (Stata Corporation, College Station, Texas, USA).
3. Results

3.1 Inter-assessor reliability

The inter-assessor level of agreement (ICC) between the four assessors was high (0.87; 95% CI: 0.78, 0.92).

3.2 Search Results

A total of 162 dental guidelines were identified Figure 1.

3.3 Guideline Demographics

Of the 162 guidelines, 33.3% (n=54) did not state the number of authors. The mean number of authors was 5.9 (SD 7.7) (range 1 – 35; n=108). The most frequent number of authors was 3 (13.9%, n=15). The mean number of guidelines published per year 11.6. The most frequent year of publication was 2013 (n=22), whilst the fewest were published in 2005 (n=2). 87 (53.7%) of guidelines were updates. 72 (49.3%) were published in North America, 60 (41.1%) in Europe and 14 (9.6%) in other continents. The remaining 16 were formed by international organisations/groupings. 77 (47.5%) of guidelines were produced in the USA, 46 (28%) in the UK, 9 (5.6%) in Australia and 5 (3.1%) in Ireland. All other source countries authored fewer than 5 guidelines with New Zealand and Norway each producing 3 (1.9%) and Germany, Italy and Singapore each producing 1 (0.6%). The majority of guidelines were multicentre 64.8% (n=90). 106 (65.4%) guidelines were formed by expert opinion, 45 (27.8%) were based on formal evidence and 8 (4.9%) utilised a clearly defined consensus method. The remaining three guidelines used a mixed approach with formal evidence followed by a defined consensus procedure.

The majority of guidelines were produced by the American Association of Paediatric Dentistry (18.8%, n=30) followed by the Royal College of Surgeons (14.4%, n=23) and American Dental Association (8.1%, n=13). Only four other organisations produced more than five guidelines. The European Association of Paediatric Dentistry and Health Partners Dental Group produced 7 (4.3%) whilst NICE and the Academy of Osseointegration both
produced 6 (3.8%). The Irish oral health services guideline initiative produced 4 guidelines (2.5%), the American Association of Periodontology and International Association of Dental Traumatology produced two each (1.88%). A further 8 organisations produced two guidelines (1.25%). 42 organisations or groups produced only one guideline whilst two guidelines were not associated with an organisation or group. 47 guidelines (29%) did not clearly fit into a subspecialty and were classified as “other”. These mainly pertained to preventative dentistry and sedation (Table 2).

3.4 Quality of guideline reporting
The mean quality score of the total sample was 51.9% (SD 13.3) (Table 3). Clarity of Presentation (85.7% SD 20.0) and Scope and Purpose (73.1% SD 15.2) domains were well-reported followed by Rigour of Development (52.4% SD 18.4) and Stakeholder Involvement (49.5% SD 14.3). Applicability (20.4% SD 11.8) and Editorial Independence (34.2% SD 21.4) domains were poorly reported. There was variation between the mean quality scores of guidelines produced by different dental specialities (Table 2). The highest score was for those classified as “other” (57.00% SD 16.41) and was lowest for endodontics (40.22% SD 6.42).

3.5 Regression analysis
Comparisons were made between baseline (reference category) and the following potential predictors: updated guideline, continent of publication, development process, single or multicentred development and publication year (Table 4). In the univariate analyses, lower quality scores were achieved by guidelines produced in Asia and other ($\beta=-7.97$, 95% CI: -15.61, -0.33, $p=0.04$) compared to Europe. Higher quality scores were achieved by guidelines formed from formal evidence ($\beta=20.08$, 95% CI: 16.62, 23.54, $p=0.001$) compared to expert opinion and those produced by multicentre groups ($\beta=7.80$, 95% CI: 3.30, 12.31, $p=0.001$) vs single centre groups. The quality of guidelines improved with publication year ($\beta=0.76$, 95% CI: 0.26, 1.26, $p=0.003$). In the multivariate analysis, the type of evidence remained a
4. Discussion

This study aimed to evaluate the reporting of clinical practice guidelines in dentistry against the domains of the AGREE II instrument and identify characteristics that influence their quality. The overall mean quality score was 51.9% (SD 13.3). Whilst both Scope and Purpose and Clarity of Presentation domains were well-reported, deficiencies were noted in the following domains: Stakeholder Involvement, Rigour of Development, Applicability and Editorial Independence. Similar observations have been previously reported [8, 10, 11, 16, 17]. Clear presentation may obfuscate poor methodological rigour, giving the impression of a well-developed guideline. The low Editorial Independence score does not necessarily indicate a conflict of interest or that guidelines were subject to bias. The poor Applicability score may be due to the resources required to undertake a cost benefit analysis. Smaller organisations or individuals may not have the expertise to undertake such appraisals whilst clinical experts may not consider resource implications or monitoring criteria as part of their remit in guideline development.

The majority of dental guidelines were paediatric with surprisingly few guidelines on restorative dentistry, prosthodontics, endodontics and periodontics identified. This may reflect the categorisation of guidelines. Guidelines that spanned more than one speciality were classified according to the affiliation of the producing organisation only. There was variation between the quality scores of guidelines produced by various dental specialities. The overall score was highest for those classified as “other” (57.0% SD 16.4), which consisted primarily of guidelines related to preventative dentistry and sedation. These are often published by public health programmes or national organisations. The former may have greater resources, are more likely to consider cost implications and identify methods to disseminate the guideline [18].

Guidelines produced outside of Europe or North America were more likely to be of
lower quality ($\beta=-7.97$, 95% CI: -15.61, -0.33, $p=0.04$). There was no statistically significant difference between North American and European guidelines. In contrast, American medical guidelines were found to be of poorer quality than European guidelines. This has been attributed to greater involvement and funding from public institutions in the latter, with more guidelines being developed by medical societies in the former [16]. The use of formal evidence in the guideline development process improved quality scores in comparison to expert opinion or consensus based methods ($\beta=19.94$, 95% CI: 15.25, 24.64, $p=0.001$). This is to be expected and to a degree confirms the validity of the AGREE II instrument which recommends that guidelines are evidence based. Despite an improvement in the reporting of guidelines between 2000 and 2014 per year ($\beta=0.76$, 95% CI: 0.26, 1.26, $p=0.003$), overall the quality of guidelines was suboptimal. This mirrors an assessment of medical guidelines, where the quality remained moderate to low despite a significant improvement in reporting between 1980 and 2007 [16]. Despite the improvement in quality over time, being an update of a previous guideline did not lead to an improvement in quality as would be expected. This lack of association between updated guidelines and quality has been found previously [19]. This may be due to guidelines being left unchanged and simply marked as reviewed when no further clinical evidence was published.

Previous studies have been limited to a particular sub-specialty [9-11], utilised the original AGREE instrument [8] and not employed the minimum of four reviewers as recommended by the AGREE collaboration. A review of studies of medical specialities utilising the AGREE II or its predecessor found that only 57% were assessed with three or more appraisers [16]. A high level of inter-assessor agreement was detected between the four reviewers in this study (ICC=0.87, 95% CI: 0.79, 0.85) increasing the external validity. Difficulties identifying dental guidelines have been previously described [9, 11]. To account for this, the search terms were broad and multiple sources searched. Despite this, it is likely guidelines will have been excluded or not accessible due to paywalls. The failure to identify all guidelines may be a source of bias and may be compounded by the exclusion of non-
English guidelines. However, the latter eliminated difficulties with translation and interpretation which may have affected the accurate assessment of guidelines. The International Oral Healthcare Guideline Repository established by the Cochrane Oral Health Group [20] may overcome difficulties in guideline identification in the future. It is prospective and therefore not currently a comprehensive resource of guidelines related to oral health care. Turning Research Into Practice (TRIP) is a clinical search engine that incorporates a publication score based on quality to order results allowing users to identify high quality research evidence [21]. Although content is identified and indexed from PubMed, it is also added manually by the developers based on their own experience and recommendations from their network of users. Whilst the latter may have identified guidelines not otherwise included, this was not the case. All the results of the TRIP search were also found in at least one of the other search methods.

Although other instruments are available to assess the quality of guidelines, a lack of standardisation has been highlighted with variation in the source, number of items and methods of scoring reported [5]. Although the AGREE II instrument incorporates an assessment of whether a systematic search was undertaken and reported, it does not assess clinical content or the quality of the evidence supporting the recommendations. This is similar to other guideline development tools [5]. Guidelines may therefore score highly despite being subject to bias, lack methodological rigour and editorial independence or poorly developed with inappropriate inclusion/exclusion criteria or inadequate critical appraisal [5]. The opposite may be true, as guidelines resulting from well conducted reviews may achieve a low score, if the reporting was inadequate. Both factors may give a false impression to clinicians of the quality of a guideline. Thus, whilst the AGREE II may give an indication of the validity of the recommendations presented, it does not discern this fully and quality scores should be seen for what they are: an indicator of the clarity of reporting [9].

The potential benefits that clinical practice guidelines may have for patients, clinicians and healthcare systems cannot be realised if they are of poor quality [2]. They may instead adversely affect patient care and the decision-making process [2]. The latter maybe
further compounded by older guidelines which have not considered the recommendations made by AGREE in their development process. It is thus imperative that clinical practice guidelines in Dentistry continue to improve and adhere to the AGREE recommendations. The adoption of the AGREE tool by numerous organisations such as NICE and SIGN reflect its assessment as a validated, transparent and easy to use instrument [5]. Awareness of its use may be an impetus for guideline developers to improve the reporting of the development process and a shift towards more evidence based as opposed to expert consensus methods. Based on the findings of this study greater emphasis should also be placed on the reporting of Stakeholder Involvement, Rigour of Development, Applicability and Editorial Independence in future dental guidelines. The need for greater adherence to methodological standards during the review, planning and development of guidelines has been previously recommended [22]. An assessment of the quality of the evidence supporting the recommendations would further strengthen the guideline process. Large institutions with greater resources may be able to keep up to date with developing methodology of guideline development [16]. Whilst the need to minimise duplication has been identified [16], guidelines developed only by larger national or international organisations may overlook variations in local needs or may be deemed by some clinicians as a tool to control decision making and reduce costs [4].

5. Conclusion

This study has shown that despite a steady improvement between 2000 and 2014 clinical practice guidelines in dentistry are of suboptimal quality in relation to the AGREE II instrument. There is variation in the overall quality and reporting of individual domains of the AGREE II between different dental speciality clinical practice guidelines. Guidelines developed using formal evidence were of higher quality. The widespread adoption of the AGREE II may help to improve the quality of dental clinical practice guidelines by encouraging improvements to the guideline development process.

Conflicts of interest
Acknowledgements

The authors would like to thank Mr Alan Fricker for his help with the search methodology.

The authors received no financial support and declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

References


Figure 1. Flow diagram of guideline identification and retrieval

- **Identification**
  - MEDLINE Ovid ®
    - Records Screened: n=1097
  - TRIP Database
    - Records Screened: n=1624
  - NICE
    - Records Screened: n=1023
  - Additional Sources
    - Records Screened: n=120

- **Primary Screening**
  - Records Screened: n=3864
    - Records Excluded: n=3214

- **Secondary Screening**
  - Secondary screening: n=650
    - Records Excluded: n=21

- **Removal of duplicates/updates**
  - n=629
    - Guidelines excluded: n=442

- **Eligibility**
  - Full text retrieved reviewed: n=187
    - Guidelines excluded: n=25

- **Included**
  - Guidelines Reviewed: n=162
<table>
<thead>
<tr>
<th>Search term</th>
<th>Number of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. guideline.mp</td>
<td>72298</td>
</tr>
<tr>
<td>2. exp guideline/</td>
<td>27387</td>
</tr>
<tr>
<td>3. Clinical recommendation.mp</td>
<td>126</td>
</tr>
<tr>
<td>4. Position statement.mp</td>
<td>2090</td>
</tr>
<tr>
<td>5. Position paper.mp</td>
<td>2003</td>
</tr>
<tr>
<td>6. 1 or 2 or 3 or 4 or 5</td>
<td>75380</td>
</tr>
<tr>
<td>7. Dent*.mp</td>
<td>468408</td>
</tr>
<tr>
<td>8. 6 and 7</td>
<td>1577</td>
</tr>
<tr>
<td>9. Limit 8 to yr=2000-2014</td>
<td>1097</td>
</tr>
</tbody>
</table>

Table 1. Search terms used for MEDLINE(Ovid®)
<table>
<thead>
<tr>
<th>Speciality</th>
<th>Number of guidelines (%)</th>
<th>Mean overall score (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endodontics</td>
<td>5 (3.1)</td>
<td>40.2 (6.4)</td>
</tr>
<tr>
<td>Implants</td>
<td>8 (4.9)</td>
<td>56.2 (10.2)</td>
</tr>
<tr>
<td>Paediatrics</td>
<td>53 (32.7)</td>
<td>49.9 (7.4)</td>
</tr>
<tr>
<td>Periodontics</td>
<td>8 (4.9)</td>
<td>44.9 (13.3)</td>
</tr>
<tr>
<td>Prosthodontics</td>
<td>2 (1.2)</td>
<td>44.8 (12.2)</td>
</tr>
<tr>
<td>Radiology</td>
<td>3 (1.9)</td>
<td>54.3 (18.7)</td>
</tr>
<tr>
<td>Restorative</td>
<td>8 (4.9)</td>
<td>44.5 (14.1)</td>
</tr>
<tr>
<td>Oral Surgery</td>
<td>20 (12.4)</td>
<td>53.7 (14.1)</td>
</tr>
<tr>
<td>Orthodontics</td>
<td>8 (4.9)</td>
<td>49.6 (15.1)</td>
</tr>
<tr>
<td>Other</td>
<td>47 (29.1)</td>
<td>57 (16.4)</td>
</tr>
</tbody>
</table>

Table 2. Number of guidelines and Mean Domain scores (%) by speciality (n=162)
<table>
<thead>
<tr>
<th>Domain</th>
<th>Mean Domain Score (SD)</th>
<th>Mean Domain Score (%) (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 1: Scope and Purpose</td>
<td>15.3 (3.2)</td>
<td>73.1 (15.2)</td>
</tr>
<tr>
<td>Domain 2: Stakeholder Involvement</td>
<td>10.4 (3.0)</td>
<td>49.5 (14.3)</td>
</tr>
<tr>
<td>Domain 3: Rigour of Development</td>
<td>29.4 (10.3)</td>
<td>52.4 (18.4)</td>
</tr>
<tr>
<td>Domain 4: Clarity of Presentation</td>
<td>18.0 (4.2)</td>
<td>85.7 (20.0)</td>
</tr>
<tr>
<td>Domain 5: Applicability</td>
<td>5.7 (3.3)</td>
<td>20.4 (11.8)</td>
</tr>
<tr>
<td>Domain 6: Editorial Independence</td>
<td>4.8 (3.0)</td>
<td>34.2 (21.4)</td>
</tr>
<tr>
<td>Overall Score</td>
<td>83.6 (21.4)</td>
<td>51.9 (13.3)</td>
</tr>
</tbody>
</table>

Table 3. Quality of dental guideline reporting by domain (n=162)
| Predictor Variables | Univariable | | | Multivariable | | |
|---------------------|-------------|----------------|----------------|----------------|----------------|
|                     | Variable    | Category       | β-coefficient  | 95% CI         | p-value        | β-coefficient  | 95% CI         | p-value        |
| Update              | Yes         | No             | -2.18          | -6.37, 2.01    | 0.31           | -2.73, 5.94    | 0.46           |
|                     |             | Baseline       |                |                |                |                |                |                |
|                     | Yes         | Baseline       |                |                |                |                |                |                |
|                     | Yes         | Multi          | 7.80           | 3.30, 12.31    | 0.001*         | -0.23          | -5.06, 4.60    | 0.93           |
|                     |             | Baseline       |                |                |                |                |                |                |
| Continent           | Yes         | Europe         | -2.17          | -6.67, 2.33    | 0.34           | -0.73          | -4.92, 3.47    | 0.73           |
|                     | Yes         | Americas       |                |                |                |                |                |                |
|                     | Yes         | Asia and other | -7.97          | -15.61, -0.33  | 0.04*          | 2.72           | -4.48, 10.23   | 0.47           |
| Type                | Yes         | Expert Opinion | 5.81           | -1.31, 12.94   | 0.10           | 5.79           | -2.24, 13.82   | 0.16           |
|                     | Yes         | Consensus based|                |                |                |                |                |                |
|                     | Yes         | Formal evidence| 20.08          | 16.62, 23.54   | 0.00*          | 19.94          | 15.25, 24.64   | 0.001*         |
| Number of centers   | Yes         | Single         |                |                |                |                |                |                |
|                     | Yes         | Multi          | 7.80           | 3.30, 12.31    | 0.001*         | -0.23          | -5.06, 4.60    | 0.93           |
| Publication year    | Yes         | Baseline       |                |                |                |                |                |                |
|                     | Yes         | Multi          | 7.80           | 3.30, 12.31    | 0.001*         | -0.23          | -5.06, 4.60    | 0.93           |

Table 4. Univariate and multivariate linear regression derived coefficients ($\beta$) and 95% confidence intervals (CI) for quality assessment as dependent variable for the 162 guidelines. * p-value <0.05, ** p-value <0.01