Nonconformity in the clinical practice guidelines for subclinical Cushing's syndrome: which guidelines are trustworthy?

Jing Shen, Mingfang Sun, Bo Zhou and Juping Yan

Department of Endocrinology, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, People's Republic of China

Correspondence
should be addressed
to B Zhou
Email
zhoubo915@126.com

Abstract

Objective: This study aimed to systematically evaluate the quality of clinical practice guidelines (CPGs) and to compare their recommendations for managing subclinical Cushing's syndrome (SCS) to assist practitioners in making rapid clinical decisions.

Design and methods: SCS management guidelines were retrieved from electronic databases. The Appraisal of Guidelines Research and Evaluation II (AGREE-II) tool and the Institute of Medicine (IOM) criteria were used to evaluate the quality of the selected guidelines. In addition, we further compared recommendations, cited references and levels of evidence between the SCS management guidelines.

Results: We included five guidelines that were obtained through the literature selection process. On the basis of the AGREE-II and IOM criteria, none of the selected guidelines were satisfactory in all aspects. However, the Italian Association of Clinical Endocrinologists (IACE) guidelines demonstrated slightly higher scores than did the other guidelines, so this guideline was recommended (with certain modifications for several domains). Regarding the content of the CPGs, we found considerable differences in the recommendations for managing SCS. These differences were derived from citation selection bias, evidence interpretation bias, differences in the composition of the guidelines' workgroups and the omission of guidelines for updating and externally reviewing the recommendations.

Conclusions: There is generally poor guideline quality among different organisations, and remarkable differences exist in the recommendations for the same clinical subject. Therefore, future guideline development should be performed in strict accordance with the AGREE-II and IOM criteria.

European Journal of
Endocrinology
(2014) 171, 421–431

Introduction

Subclinical Cushing's syndrome (SCS) is used to refer to an adrenal incidentaloma with subtle autonomous cortisol secretion and without typical signs and symptoms of hypercortisolism (1). This subtle cortisol excess has been estimated to be 5–20% in adrenal masses (2). Although no overt manifestation of hypercortisolism is typically present, a large study (3, 4, 5, 6) demonstrated that the continuous subtle cortisol hypersecretion inherent to SCS was associated with metabolic complications, including diabetes mellitus, obesity, hypertension, dyslipidaemia and reduced bone mineral density. In addition, post-adrenalectomy

adrenal insufficiency is potentially fatal for SCS patients if they are not supplemented with hydrocortisone in a timely manner. Therefore, patients with adrenal tumours should be aware of SCS, which is an easily overlooked disease (7). However, because of the absence of the typical clinical features of cortisol excess and the broad variation in autonomous cortisol secretion among patients, the optimal strategy for diagnosing and treating SCS is poorly defined. To resolve this issue, several national and international organisations have published relevant clinical practice guidelines (CPGs) for managing SCS.

CPGs are intended to develop systematic statements to aid practitioners in making appropriate decisions regarding the treatment of specific clinical conditions (8). Trustworthy CPGs should provide recommendations based on a synthesis of all high-quality evidence and on an assessment of the benefits and risks, while minimising conflicts of interest and providing an explicit process for translating the evidence into practice (8). However, guidelines developed by different groups on the same topic are often disparate, creating confusion for practitioners attempting to make clinical decisions based on the appropriate guidelines. Therefore, several organisations have proposed standards to critically appraise guidelines.

The Appraisal of Guidelines Research and Evaluation (AGREE) instrument is used to assess the methodological rigour and transparency of the guideline development process (9). The original AGREE tool was released in 2003, and the updated version (AGREE-II) was published in 2009. AGREE-II is the most widely used appraisal tool (10). Although AGREE-II provides a set of standards to assess the quality of CPGs, it is not comprehensive in certain areas, such as guideline updates, external reviews, conflict of interest management and the process by which evidence is translated into conclusions or recommendations. In addition to the criteria described by AGREE-II, the Institute of Medicine (IOM) crafted related criteria for developing rigorous and trustworthy CPGs in 2011 (8). It has been suggested that the IOM recommendations are more extensive than the AGREE standards in certain respects (11). Therefore, combining the two approaches will allow for a more comprehensive evaluation of CPGs.

To the best of our knowledge, there are no established appraisals of the quality of the CPGs for SCS. Therefore, the purpose of this study was to systematically evaluate the quality of the SCS CPGs using the AGREE-II and IOM criteria and to compare the CPG recommendations for managing SCS to help practitioners to make rapid clinical decisions.

Subjects and methods

Selection of related guidelines

We searched for relevant CPGs regarding the management of SCS in the following electronic databases: the Guidelines International Network (GIN), the National Guidelines Clearinghouse (NGC), the National Institute for Health and Care Excellence (NICE) and PubMed. The following search terms were used: i) 'Cushing's syndrome', ii) 'hypercortisolism', iii) 'adrenal incidentaloma',

iv) 'inapparent adrenal adenomas (mass)', v) 'subclinical Cushing's syndrome', vi) 'subclinical hypercortisolism', vii) 'subclinical autonomous glucocorticoid hypersecretion' and viii) 'preclinical Cushing's syndrome'. In addition, we performed searches in PubMed by combining the above terms with 'guideline' OR 'consensus' OR 'recommendation' OR 'position statement'. We restricted the search to documents that were published in English between January 2003 and December 2013.

We established the following inclusion criteria to select guidelines: i) the target group included patients with SCS; ii) the guidelines focused on the diagnosis, management and/or follow-up of SCS; and iii) the full text of the guidelines was available online. The following exclusion criteria were applied: i) the guidelines concentrated on diagnosing Cushing's syndrome or adrenal incidentaloma but did not address SCS in detail; ii) the guidelines focused entirely on special groups, such as pregnant women; iii) the guidelines did not meet the definition of a guideline; and iv) the guidelines did not present specific or original recommendations. When different guideline versions were released by the same professional team, only the most recent version was selected.

Data extraction from the guidelines

Two reviewers individually extracted details pertaining to the CPGs' characteristics from the selected guidelines. These guidelines primarily included the following content: the development team, country, publication date, guideline panel members, guideline review, update/frequency and funding source. In addition, the reviewers summarised all recommendations of the SCS management guidelines to compare their similarities and differences. For each domain, the decisions of each reviewer were compared, and any disagreements were resolved by consensus or were assessed by a third reviewer.

Quality evaluation of the guidelines

The AGREE-II tool and the IOM standards were used to evaluate the quality of the selected guidelines with respect to SCS management. The AGREE-II instrument contains 23 items grouped into six main domains (i.e. domain 1, scope and purpose; domain 2, stakeholder involvement; domain 3, rigor of development; domain 4, clarity of presentation; domain 5, applicability and domain 6, editorial independence) (9). All guidelines were independently appraised by each assessor. Each of the AGREE-II items was rated on a scale of 1 (strongly disagree) to 7

(strongly agree). Each item with a score discrepancy of more than three between the two reviewers was discussed further and independently reassessed. We obtained the score for each domain by summing all individual item scores. The calculation method for the AGREE-II scores is available in the user manual for the AGREE-II instrument (9). Finally, we performed overall use assessments of the quality of the guidelines based on the appraisal items, as follows: 'strongly recommended', 'recommended for use with certain modification' or 'not recommended'. A guideline was labeled as 'strongly recommended' if most domain scores (pertaining to four or more domains) were above 60%; 'recommended for use with certain modification' if most domain scores were between 30 and 60% or three domain scores were above 60% and 'not recommended' if the majority of items scored low or most domains were scored around or below 30%.

All guidelines were re-evaluated using the IOM criteria, which include eight standards: i) establishment of transparency, ii) management of conflicts of interest, iii) guideline development group composition, iv) systematic review, v) evidence foundations for and rating strength of the recommendations, vi) articulation of the recommendations, vii) external review and viii) updating (8). A CPG was considered to comply with a particular standard only if it fulfilled all standard sub-criteria. Each CPG was independently evaluated by two reviewers and discrepancies were resolved by open discussion.

Evidence grades and a comparison of the citations in the guidelines

All appraisal instruments relied on the quality of the guideline development process, but they did not ensure

excellent recommendations. To further evaluate the recommendations based on the selected guidelines, we analysed the level of evidence for and the cited references in each selected guideline.

The selected guidelines used two evidence grading systems: the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system and the American Association of Clinical Endocrinologists system (Table 1) (12, 13, 14). We merged and synchronised the two grading systems to form high-, moderate- and low-quality evidence groups to compare the levels of evidence for the different guidelines (Table 1). Finally, we calculated the percentages of high-, moderate- and low-quality evidence in the guidelines.

The cited references supporting the recommendations in each guideline were independently extracted by two reviewers and the number of references was quantified for each guideline (15). For each cited reference, we determined the publication date and the place of origin (the USA, Europe or other).

Results

Guideline search

We identified 220 citations in PubMed, as well as 23 guidelines in the GIN, the NICE and the NGC. In total, 219 documents were deemed ineligible based on screening of the titles and abstracts, and 24 documents were identified for possible inclusion. After applying the inclusion and exclusion criteria, 19 documents were excluded for the following reasons: 11 did not meet the definition of a guideline, three did not address SCS in detail, one focused entirely on special groups, one contained a non-original

Table 1 Summary of the grades of the evidence used by the ES, AACE/AAES and IACE guidelines.

Evidence grade
ES and IACE (29)
High, well-performed RCTs or exceptionally strong evidence from unbiased observational studies
Moderate, RCTs with important limitations (inconsistent results, methodological flaws, indirect or imprecise evidence) or unusually strong evidence from unbiased observational studies
Low, at least one critical outcome from observational studies, RCTs with serious flaws or indirect evidence
Very low, at least one of the critical outcomes from unsystematic clinical observations or indirect evidence
AACE/AAES (30)
1, Well-controlled RCTs, well-controlled multi-centre trial, large meta-analysis with quality ratings
2, RCTs with a limited body of data, well-conducted prospective cohort study, well-conducted meta-analysis of cohort studies
3, Methodologically flawed RCTs, observational studies, case series or case reports
4, Expert consensus, expert opinion based on experience
Adjusted grade of evidence
High, data derived from RCTs or a large meta-analysis with quality ratings, well-controlled multi-centre trial
Moderate, data derived from methodologically flawed RCTs, non-randomised studies or observational studies
Low, evidence based on clinical experience, the opinion of respected authorities, case series or case reports

Figure 1

Flowchart for selecting clinical practice guidelines. GIN, Guidelines International Network; NICE, National Institute for Health and Care Excellence; NGC, National Guidelines Clearinghouse; SCS, subclinical Cushing's syndrome.

recommendation and three were overlaps. Accordingly, the five remaining guidelines fulfilled our inclusion criteria (12, 13, 14, 16, 17). The process for selecting the CPGs is shown in Fig. 1.

Comparison of the characteristics of the guidelines

The features of the included guidelines are provided in Table 2. Three of the selected guidelines were developed in the USA (13, 14, 16), and the rest were developed in Europe (12, 17). The guidelines were published between 2003 and 2011. However, none of the guidelines mentioned updates. Three guidelines were produced by multidisciplinary teams (13, 16, 17), and the remaining guidelines were produced by clinical endocrinologists and/or endocrine surgeons (12, 14). Of the five guidelines, only the Endocrine Society (ES) guideline was externally reviewed before publication (13); the American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons (AACE/AAES) guideline was reviewed by an internal peer (14), the Italian Association of Clinical Endocrinologists (IACE) guidelines were reviewed by experts (12) and the National Institutes of Health (NIH) and French Society of Endocrinology (FSE) guidelines did not disclose any review process (16, 17).

Regarding the funding sources, two guidelines were funded by different organisations (13, 16), one was not funded (12) and the rest did not mention a funding source (14, 17).

Comparison of the SCS management recommendations

The guidelines' recommendations are presented in Table 3. Regarding the diagnostic criteria for SCS, all selected guidelines recommended a 1 mg overnight dexamethasone suppression test (DST) as the first screening test. However, the CPGs recommended various cutoff points. The ES and FSE guidelines dictated that a serum cortisol level $>1.8 \mu\text{g/dl}$ after dexamethasone administration likely indicates SCS (13, 17), the NIH suggested that serum cortisol levels are suppressed to $<5.0 \mu\text{g/dl}$ after dexamethasone administration in most normal individuals (16), and the AACE/AAES recommended that SCS is indicated by serum cortisol levels $>5.0 \mu\text{g/dl}$ after a 1-mg overnight DST (14). In the IACE guideline, cortisol levels $<1.8 \mu\text{g/dl}$ clearly exclude SCS, and levels $>5 \mu\text{g/dl}$ likely indicate SCS after dexamethasone administration, whereas cortisol levels between 1.8 and $5 \mu\text{g/dl}$ might be considered as indeterminate (12). None of the guidelines, with the exception of the ES guideline, recommended late-night salivary cortisol as the first screening test. For the second screening test, the FSE guideline advocated late-night salivary cortisol, late-night serum cortisol and 24-h urine free cortisol (17), whereas the IACE guidelines did not recommend the late-night salivary cortisol test as a screening method for SCS (12). For the confirmatory test, only the AACE/AAES guideline suggested using a low-dose 2-day DST (14). Four of the guidelines recommended measuring plasma adrenocorticotrophic hormone (ACTH) levels to evaluate autonomous cortisol secretion by an adrenal adenoma (12, 13, 14, 17). The ES and AACE/AAES guidelines recommended measuring DHEA sulphate (DHEAS) levels (13, 14), which the IACE and FSE guidelines did not recommend (12, 17). Adrenal scintigraphy was only recommended by the FSE guideline (17).

Regarding surgery in SCS patients, the IACE and AACE/AAES guidelines recommended adrenalectomy, with various indications (12, 14), but the NIH and FSE guidelines did not recommend surgery (16, 17) because of insufficient evidence. Perioperative glucocorticoid therapy was recommended by three of the guidelines (14, 16, 17). For post-operative glucocorticoid replacement, the IACE and AACE/AAES guidelines suggested that patients should be treated with exogenous glucocorticoids after

Clinical Study		J Shen and others		Guidelines for subclinical Cushing's syndrome		171:4		425	
Table 2 Characteristics of the clinical practice guidelines included in this study.									
Guideline title (ref)	Organisation	Country	Publication date	Guideline panel composition	Guideline review	Update/frequency	Funding source		
Management of the clinically inapparent adrenal mass ('incidentaloma') (16)	The NIH	USA	2003	Endocrinologists, surgeons, pathologists, biostatisticians, radiologists, Oncologists and the general public	NM	NM	The National Institute of Child Health and Human Development, the NIH Office of Medical Applications of Research, the National Cancer Institute and the National Institute of Diabetes and Digestive and Kidney Diseases		
The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline (13)	ES	USA	2008	Endocrinologists and methodologist	External review	NM	The Endocrine Society and the European Society of Endocrinology		
	AACE/AAES	USA	2009	Endocrinologists and surgeons	Internal peer review	NM	NM		
American Association of Endocrine Surgeons Medical Guidelines for the management of adrenal incidentalomas (14)	FSE	France	2008	Endocrinologists, surgeons, radiologists and nuclear physicians	NM	NM	NM		
Exploration and management of adrenal incidentalomas (17)	IACE	Italy	2011	Endocrinologists	Expert review	NM	NF		
AME position statement on adrenal incidentaloma (12)									
NIH, National Institutes of Health; ES, Endocrine Society; AACE/AAES, American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons; FSE, French Society of Endocrinology; IACE, Italian Association of Clinical Endocrinologists; NM, not mentioned; NF, no funding source.									
<p>adrenalectomy, but the glucocorticoid treatment durations differed (12, 14).</p> <p>Regarding the recommendations on the follow-up of indeterminate SCS, four guidelines (12, 14, 16, 17) recommended biochemical follow-up, but there was no consensus on the frequency or the time interval. In addition, the IACE and FSE guidelines (12, 17) mentioned the necessity of clinical follow-up.</p>									
<h3>Quality assessment</h3> <p>The results of the domain scores (%) for each guideline determined using the AGREE-II instrument, are presented in Table 4. None of the selected guidelines performed satisfactorily or had a high score in each AGREE-II domain. Among the different guidelines, scores were relatively high for the scope and purpose domain (median, 77.8%). In contrast, all guidelines received poor scores in the applicability domain (median, 20.8%). The IACE guideline (12) received the highest scores in the scope and purpose domain (97.2%), the clarity of presentation domain (91.7%) and the editorial independence domain (100%), but had the lowest score for stakeholder involvement (27.8%). The NIH and ES guidelines (13, 16) scored highest for the stakeholder involvement domain (47.2%) and for the rigour of development domain (51.0%) respectively. In overall assessments, the ES and IACE guidelines (12, 13) were recommended with additional modification in one or more domains. The remaining guidelines (14, 16, 17) were poor in certain domains, based on the AGREE-II criteria, and were not recommended for clinical use.</p> <p>There was generally poor adherence to the IOM standards among the selected CPGs. The IACE guideline received slightly higher scores than the other guidelines did and met half of the 20 sub-criteria. Standards 1.1, 4.1 and 6.1 were met by all guidelines (Table 5). However, eight of the 20 sub-criteria (i.e. standards 3.2, 3.3, 7.2, 7.3, 7.4, 8.1, 8.2 and 8.3) were not met by any of the CPGs; these standards related to the composition of the guideline development group, external review and updating.</p>									
<h3>Comparison of the evidence grades for the guideline</h3> <p>The NIH and FSE guidelines did not provide evidence grades and therefore were not included in the comparison of the evidence grades. For the three remaining guidelines, the ES guideline only involved the diagnosis of SCS, whereas the AACE/AAES and IACE guidelines referred to the diagnosis, treatment and follow-up of SCS. The total numbers of evidence in the ES, AACE/AAES and IAC</p>									

Table 2 Characteristics of the clinical practice guidelines included in this study.

adrenalectomy, but the glucocorticoid treatment durations differed (12, 14).

Regarding the recommendations on the follow-up of indeterminate SCS, four guidelines (12, 14, 16, 17) recommended biochemical follow-up, but there was no consensus on the frequency or the time interval. In addition, the IACE and FSE guidelines (12, 17) mentioned the necessity of clinical follow-up.

Quality assessment

The results of the domain scores (%) for each guideline, determined using the AGREE-II instrument, are presented in Table 4. None of the selected guidelines performed satisfactorily or had a high score in each AGREE-II domain. Among the different guidelines, scores were relatively high for the scope and purpose domain (median, 77.8%). In contrast, all guidelines received poor scores in the applicability domain (median, 20.8%). The IACE guideline (12) received the highest scores in the scope and purpose domain (97.2%), the clarity of presentation domain (91.7%) and the editorial independence domain (100%), but had the lowest score for stakeholder involvement (27.8%). The NIH and ES guidelines (13, 16) scored highest for the stakeholder involvement domain (47.2%) and for the rigour of development domain (51.0%) respectively. In overall assessments, the ES and IACE guidelines (12, 13) were recommended with additional modification in one or more domains. The remaining guidelines (14, 16, 17) were poor in certain domains, based on the AGREE-II criteria, and were not recommended for clinical use.

There was generally poor adherence to the IOM standards among the selected CPGs. The IACE guideline received slightly higher scores than the other guidelines did and met half of the 20 sub-criteria. Standards 1.1, 4.1 and 6.1 were met by all guidelines (Table 5). However, eight of the 20 sub-criteria (i.e. standards 3.2, 3.3, 7.2, 7.3, 7.4, 8.1, 8.2 and 8.3) were not met by any of the CPGs; these standards related to the composition of the guideline development group, external review and updating.

Comparison of the evidence grades for the guidelines

The NIH and FSE guidelines did not provide evidence grades and therefore were not included in the comparison of the evidence grades. For the three remaining guidelines, the ES guideline only involved the diagnosis of SCS, whereas the AAACE/AAES and IACE guidelines referred to the diagnosis, treatment and follow-up of SCS. The total numbers of evidence in the ES, AAACE/AAES and IACE

Table 3 Recommendations in the five clinical practice guidelines for SCS.

Recommendation	NIH	ES	AAE/AAS	FSE	IACE
Diagnosis					
The first screening test					
1-mg overnight DST	Recommended ≥5.0 µg/dl	Recommended >1.8 µg/dl	Recommended >5.0 µg/dl	Recommended >1.8 µg/dl	Recommended <1.8 µg/dl: exclude >5.0 µg/dl: consider 1.8–5.0 µg/dl: indeterminate
Cutoff point					NR
Late-night salivary cortisol	NM	Recommended >145 ng/dl	NR	NR	
Cutoff point					
The second screening test					
24-h urine free cortisol	NM	NR	NR	Recommended	Recommended
Late-night serum cortisol	NM	NR	NM	Recommended	Recommended
Late-night salivary cortisol	NM	NR	NR	Recommended	NR
High-dose DST	NM	NM	NM	NM	NR
The confirmatory test					
Low-dose 2-day DST	NM	NR	Recommended NR	NM	NR
Cutoff point					
Evaluation of autonomous cortisol secretion from an adrenal adenoma					
ACTH	NM	Recommended Suppressed	Recommended Low or suppressed	Recommended NM	Recommended Low or suppressed
Cutoff point					NR
DHEAS	NM	Recommended Suppressed	Low NM	Recommended	NR
Cutoff point					
Adrenal scintigraphy	NR	NM	NR	NR	NR
Treatment					
Adrenalectomy	NR	NM	Recommended Patients with worsening hypertension, abnormal glucose tolerance, dyslipidaemia or osteoporosis	NR	Recommended Adequate medical therapy does not achieve the treatment goals for associated diseases potentially linked with hypercortisolism
Indication					NM
Perioperative glucocorticoid therapy	Recommended	NM	Recommended	Recommended	Recommended
Post-operative glucocorticoid replacement	NR	NM	Recommended	NM	NR
Treatment time					
Follow-up of indeterminate SCS			6–18 months		
Clinical follow-up					
Content	NM	NM	NM	Recommended BMI, blood pressure, glycaemia and lipid workup	Recommended Cardiovascular risk factors
Biochemical follow-up					
Content	Recommended 1-mg overnight DST	NM	Recommended NM	Recommended 1-mg overnight DST	Recommended 1-mg overnight DST
Frequency	Annually	NM	Annually	At 6 months and in the 2nd and 5th years	NR
Time interval	4 years	NM	5 years	5 years	NR

SCS, subclinical Cushing's syndrome; NIH, National Institutes of Health; ES, Endocrine Society; AACE/AAS, American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons; FSE, French Society of Endocrinology; IACE, Italian Association of Clinical Endocrinologists; DST, dexamethasone suppression test; ACTH, adrenocorticotrophic hormone; DHEAS, DHEA sulphate; NM, not mentioned; NR, not recommended.

Table 4 Domain scores for the clinical practice guidelines for subclinical Cushing's syndrome based on the AGREE-II instrument.

Domain scores (%)	Guideline					Median
	NIH	ES	AACE/ AAES	FSE	IACE	
Scope and purpose	94.4	69.4	72.2	77.8	97.2	77.8
Stakeholder involvement	47.2	36.1	38.9	30.6	27.8	36.1
Rigour of development	14.6	51.0	44.8	8.3	46.9	44.8
Clarity of presentation	63.9	75.0	83.3	22.2	91.7	75.0
Applicability	12.5	29.2	22.9	2.0	20.8	20.8
Editorial independence	45.8	54.2	50.0	4.2	100	50.0
Overall assessment	NR	NM	NR	NR	NM	

NIH, National Institutes of Health; ES, Endocrine Society; AACE/AAES, American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons; FSE, French Society of Endocrinology; IACE, Italian Association of Clinical Endocrinologists; NR, not recommended; NM, recommended with modification.

guidelines were 26, 12 and 52 respectively. The percentages of data sets based on moderate-quality evidence were 76.9, 66.7 and 80.8% respectively, and the percentages based on low-quality evidence were 23.1, 33.3 and 19.2% respectively (Table 6). However, all guidelines lacked high-quality evidence. The evidence for the three aspects of SCS focused on diagnosing and treating SCS, whereas the evidence for the follow-up of indeterminate SCS was relatively insufficient.

Comparison of the references cited by the SCS guidelines

The NIH guideline did not include relevant references; therefore, we compared the cited references in the other four guidelines. The results are presented in Table 7. Of the 26 citations in the ES guideline, nine (34.6%) were for American publications. Of the 12 citations included in the AACE/AAES guideline, seven (58.3%) were for American publications. In contrast, the FSE and IACE guidelines cited more European publications (84.6 and 71.2% respectively).

The cited publications in the four guidelines were predominantly published before 2007 (77.6%); of these, only two references (18, 19) were common to the FSE, AACE/AAES and IACE guidelines. The differences in the cited references likely explain several of the discrepancies between the guidelines.

Discussion

The present study systematically evaluated and compared the quality and content of CPGs for SCS. We identified five guidelines through a systematic literature review, but none of the current CPGs scored well in all domains of the AGREE-II instrument or adhered to each of the IOM standards. In addition, because of a lack of a clear-cut, unambiguous case definition of SCS, clinical research remains poorly guided. Therefore, within the content of the existing CPGs, there were considerable discrepancies in the recommendations, particularly concerning the procedures for diagnosing and treating SCS. Such disagreements occurred for both valid and non-valid reasons (20).

Differences in the recommendations for diagnosing SCS

Regarding SCS diagnosis, all of the guidelines advocated the 1 mg overnight DST as the first screening test, and most of the guidelines suggested measuring plasma ACTH levels as the method to evaluate autonomous cortisol secretion from an adrenal adenoma. However, the main differences in the recommendations were related to the cutoff point for the 1 mg DST, to the second-line investigation, and to the confirmatory testing method. These disagreements among the guidelines required further consideration.

First, the AGREE-II scores and the IOM standards indicated that most of the selected CPGs did not perform well in the systematic review. In addition, in analysing the evidence presented or cited in the guidelines, we found that the AACE/AAES guideline included only four references regarding SCS diagnosis. This finding potentially suggested that the guidelines included diagnostic recommendations based on expert opinions, without clear evidence. The comparison of the references suggested that biased citation selection may have resulted in conflicting conclusions. For example, two American guidelines (ES and AACE/AAES) recommended measuring DHEAS levels to evaluate autonomous cortisol secretion by an adrenal adenoma, whereas the European guidelines (IACE) did not recommend this test, based on certain data that were published by European research groups; these data were not cited in the American guidelines.

Second, the guidelines were developed on different dates; therefore, certain recommendations may have received more support from new evidence. For example, recent experimental data suggest that late-night salivary

Clinical Study	J Shen and others	Guidelines for subclinical Cushing's syndrome	171:4	428
----------------	-------------------	---	-------	-----

Table 5 Compliance of the selected CPGs with the IOM sub-criteria.

IOM standard	NIH	ES	AACE/AAES	FSE	IACE	Number of standards met
Establishing transparency						
Funding and development should be explicitly stated and publicly accessible	Y	Y	P	P	Y	5/5
Conflicts of interest management						
Conflicts of interest should be declared before the guideline development group formation	N	N	Y	N	Y	2/5
All conflicts of interest should be reported and discussed	N	N	N	N	Y	1/5
Guideline development group members should divest conflicts of interest	N	N	N	N	Y	1/5
Members who have conflicts of interest should be a minority of the panel members (except for the chair and co-chairs)	N	N	N	N	Y	1/5
Guideline development group composition						
Guideline development group should be multidisciplinary and balanced	Y	P	N	P	N	3/5
Patients and public should be represented on the guideline development group	N	N	N	N	N	0/5
The representatives should be trained	N	N	N	N	N	0/5
Systematic review						
Systematic reviews should be used	Y	Y	Y	Y	Y	5/5
The guideline development group and systematic review team (if used) should communicate	N	Y	N	N	Y	2/5
Evidence foundations for and rating strength of evidence						
The strength of recommendations and grading of evidence should be explicitly stated	N	Y	Y	N	P	3/5
Articulation of recommendations						
Articulate recommendations in standard form	Y	Y	Y	Y	Y	5/5
Strong recommendations should be worded as such	N	Y	Y	N	Y	3/5
External review						
External review should include full spectrum of stakeholders	N	Y	N	N	N	1/5
Authorship of external review is confidential	N	N	N	N	N	0/5
Guideline development group should consider all external review comments	N	N	N	N	N	0/5
Final draft of CPGs should be available for public comment	N	N	N	N	N	0/5
Updating						
Proposed date of future CPG review should be documented	N	N	N	N	N	0/5
Literature pertaining to CPG should be monitored regularly	N	N	N	N	N	0/5
CPG should be updated if new literature suggests modification	N	N	N	N	N	0/5
Number of standards met	4/20	8/20	6/20	4/20	10/20	

Y, yes; N, no; P, partially; CPG, clinical practice guideline; NIH, National Institutes of Health; ES, Endocrine Society; AACE/AAES, American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons; FSE, French Society of Endocrinology; IACE, Italian Association of Clinical Endocrinologists.

cortisol levels may exhibit poor sensitivity for diagnosing SCS in patients with adrenal incidentalomas (21, 22, 23). Therefore, the IACE guideline, which was published in 2011, did not recommend the late-night salivary cortisol test as a secondary screening method for SCS, whereas this test was recommended in the 2008 FSE guideline. Previous research had found that at least 10% of guidelines' recommendations are no longer valid after 3.6 years and that approximately half of guidelines are no longer valid by 5.8 years (24). It has been recommended that CPGs should be updated at least every 5 years (25). Therefore, guidelines that have not been updated for a long time should be reassessed for validity. In addition, there was

particularly poor adherence to the IOM standards regarding regulation of guideline updates.

Finally, future guidelines must be reviewed externally before they are published. The aim of such a review is to expose biases and to provide suggestions for improving and clarifying the guidelines (8). However, the ES guideline was the only guideline that included external reviewers.

Differences in the SCS treatment recommendations

To treat SCS, the IACE and AACE/AAES guidelines recommended adrenalectomy, with various indications

Table 6 Evidence grades for the conclusions in the ES, AACE/AAES and IACE guidelines, grouped as high, moderate or low.

Adjusted evidence grade	ES	AACE/AAES	IACE
Diagnosis	26	4	40
High	0	0	0
Moderate	20	3	34
Low	6	1	6
Treatment		7	18
High	0	0	0
Moderate	0	4	14
Low	0	3	4
Follow-up		1	3
High	0	0	0
Moderate	0	1	0
Low	0	0	3
Total	26	12	52 ^a
High	0	0	0
Moderate	20 (76.9%)	8 (66.7%)	42 (80.8%)
Low	6 (23.1%)	4 (33.3%)	10 (19.2%)

ES, Endocrine Society; AACE/AAES, American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons; IACE, Italian Association of Clinical Endocrinologists.

^aOf the 40 evidence included in the IACE guidelines for diagnosis, eight were used to justify the treatment and one was cited regarding follow-up.

(12, 14), whereas the NIH and FSE guidelines did not recommend surgery (16, 17) because of insufficient evidence. In addition to lacking high-quality evidence, such discrepant recommendations were likely derived from the different composition of the groups publishing the guidelines. In the process of developing guidelines, multidisciplinary teams should be composed of methodologists or biostatisticians, relevant professional clinicians and affected populations. A guideline's reliability may increase when developed by a multidisciplinary group because of better balancing of individual biases (26). Nevertheless, two of the five identified guidelines were developed only by specialists in clinical endocrinology, and none of the guidelines considered the views

of the affected population. The guidelines' workgroup composition reportedly has substantial influence on the development of the guideline (27). Different panels may propose disparate recommendations on the same topic because they have different values, which may lead to conflicting recommendations in CPGs. The low scores for editorial independence offer another explanation for the differences among the guidelines. The FSE and AACE/AAES guideline development groups did not provide a statement regarding whether the content of the guidelines was influenced by the viewpoints of the funding body. Obviously, the funding bodies, and particularly those with professional associations, may have influenced the development of the recommendations.

Differences in the recommendations for SCS follow-up

Our study demonstrated that most of the recommendations were not evidence based. In particular, there was a lack of high- and moderate-quality research regarding the follow-up of SCS patients. Two American (NIH and ES) guidelines had similar follow-up suggestions, and both European (FSE and IACE) guidelines were similar, implying that the discrepancies in the recommendations were more likely based on consensus or experiences in different regions. Moreover, evidence interpretation bias in different organisations may lead to diverse suggestions. For example, the AACE/AAES and IACE guidelines provided different suggestions for the biochemical follow-up of indeterminate SCS, although the two guidelines cited the same study from Italy (28).

Conclusion

The results of this study indicated that the quality of most of the selected CPGs was poor in terms of complying with the AGREE-II and IOM standards. Furthermore,

Table 7 The publication origins and publication dates of the references cited by the four clinical practice guidelines for subclinical Cushing's syndrome expressed as percentages.

Guideline (country)	Publication origin				Publication date					
	USA	Europe	Other	Total	Before 2007	2007	2008	2009	2010–2011	Total
ES (USA)	9 (34.6%)	15 (57.7%)	2 (7.7%)	26 (100%)	19 (73.1%)	6 (23.1%)	1 (3.8%)	0	0	26 (100%)
FSE (France)	3 (11.5%)	22 (84.6%)	1 (3.8%)	26 (100%)	26 (100%)	0	0	0	0	26 (100%)
AACE/AAES (USA)	7 (58.3%)	4 (33.3%)	1 (8.3%)	12 (100%)	9 (75%)	2 (16.7%)	1 (8.3%)	0	0	12 (100%)
IACE (Italy)	9 (17.3%)	37 (71.2%)	6 (11.5%)	52 (100%)	36 (69.2%)	2 (3.8%)	4 (7.7%)	6 (11.5%)	4 (7.7%)	52 (100%)

ES, Endocrine Society; FSE, French Society of Endocrinology; AACE/AAES, American Association of Clinical Endocrinologists/American Association of Endocrine Surgeons; IACE, Italian Association of Clinical Endocrinologists.

there were considerable differences in the content of the CPGs regarding SCS management. There are several plausible reasons for these differences, including citation selection bias, evidence interpretation bias, different composition of the guidelines' workgroups and the omission of guidelines in updating the recommendations and obtaining external review. Therefore, future guideline development should be carried out in strict accordance with the AGREE-II and IOM criteria to produce trustworthy guidelines. In addition, as there is little high-quality research (such as randomised trials) for guiding the appropriate management of SCS patients, there is a pressing need to establish multi-centre studies in different societies and countries. This pooled information will help to guide clinical practice and to achieve optimal management of SCS patients.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

References

- Young WJ. Clinical practice. The incidentally discovered adrenal mass. *New England Journal of Medicine* 2007 **356** 601–610. (doi:10.1056/NEJMcP065470)
- Barzon L, Sonino N, Fallo F, Palu G & Boscaro M. Prevalence and natural history of adrenal incidentalomas. *European Journal of Endocrinology* 2003 **149** 273–285. (doi:10.1530/eje.0.1490273)
- Tauchmanová L, Rossi R, Biondi B, Pulcrano M, Nuzzo V, Palmieri E, Fazio S & Lombardi G. Patients with subclinical Cushing's syndrome due to adrenal adenoma have increased cardiovascular risk. *Journal of Clinical Endocrinology and Metabolism* 2002 **87** 4872–4878. (doi:10.1210/jc.2001-011766)
- Bohdanowicz-Pawlak A, Szymczak J, Waszczuk E, Bolanowski M & Bednarek-Tupikowska G. Subclinical Cushing's syndrome in adrenal incidentalomas – possible metabolic consequences. *Endokrynologia Polska* 2013 **64** 186–190.
- Morelli V, Eller-Vainicher C, Salcuni AS, Coletti F, Iorio L, Muscogiuri G, Della Casa S, Arosio M, Ambrosi B, Beck-Peccoz P *et al.* Risk of new vertebral fractures in patients with adrenal incidentaloma with and without subclinical hypercortisolism: a multicenter longitudinal study. *Journal of Bone and Mineral Research* 2011 **26** 1816–1821. (doi:10.1002/jbmr.398)
- Chiodini I, Morelli V, Masserini B, Salcuni AS, Eller-Vainicher C, Viti R, Coletti F, Guglielmi G, Battista C, Carnevale V *et al.* Bone mineral density, prevalence of vertebral fractures, and bone quality in patients with adrenal incidentalomas with and without subclinical hypercortisolism: an Italian multicenter study. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 3207–3214. (doi:10.1210/jc.2009-0468)
- Eller-Vainicher C, Morelli V, Salcuni AS, Torlontano M, Coletti F, Iorio L, Cuttitta A, Ambrosio A, Vicentini L, Carnevale V *et al.* Post-surgical hypocortisolism after removal of an adrenal incidentaloma: is it predictable by an accurate endocrinological work-up before surgery? *European Journal of Endocrinology* 2010 **162** 91–99. (doi:10.1530/EJE-09-0775)
- Graham R. In *Clinical Practice Guidelines We Can Trust*. Washington, DC: National Academies Press, 2011.
- AGREE Next Steps Consortium. The AGREE II Instrument (electronic version). Retrieved 11, 10, 2013, from <http://www.agreetrust.org>, 2009.
- Sinuff T, Patel RV, Adhikari NK, Meade MO, Schunemann HJ & Cook DJ. Quality of professional society guidelines and consensus conference statements in critical care. *Critical Care Medicine* 2008 **36** 1049–1058. (doi:10.1097/CCM.0b013e31816a01ec)
- Ransohoff DF, Pignone M & Sox HC. How to decide whether a clinical practice guideline is trustworthy. *Journal of the American Medical Association* 2013 **309** 139–140. (doi:10.1001/jama.2012.156703)
- Terzolo M, Stigliano A, Chiodini I, Loli P, Furlani L, Arnaldi G, Reimondo G, Pia A, Toscano V, Zini M *et al.* AME position statement on adrenal incidentaloma. *European Journal of Endocrinology* 2011 **164** 851–870. (doi:10.1530/EJE-10-1147)
- Nieman LK, Biller BM, Findling JW, Newell-Price J, Savage MO, Stewart PM & Montori VM. The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 1526–1540. (doi:10.1210/jc.2008-0125)
- Zeiger MA, Thompson GB, Duh QY, Hamrahian AH, Angelos P, Elaraj D, Fishman E & Kharlip J. The American Association of Clinical Endocrinologists and American Association of Endocrine Surgeons medical guidelines for the management of adrenal incidentalomas. *Endocrine Practice* 2009 **15** (Suppl 1) 1–20. (doi:10.4158/EP.15.S1.1)
- Matthys J, De Meyere M, van Driel ML & De Sutter A. Differences among international pharyngitis guidelines: not just academic. *Annals of Family Medicine* 2007 **5** 436–443. (doi:10.1370/afm.741)
- Grumbach MM, Biller BM, Braunstein GD, Campbell KK, Carney JA, Godley PA, Harris EL, Lee JK, Oertel YC, Posner MC *et al.* Management of the clinically inapparent adrenal mass ("incidentaloma"). *Annals of Internal Medicine* 2003 **138** 424–429. (doi:10.7326/0003-4819-138-5-200303040-00013)
- Tabarin A, Bardet S, Bertherat J, Dupas B, Chabre O, Hamoir E, Laurent F, Tenenbaum F, Cazalda M, Lefebvre H *et al.* Exploration and management of adrenal incidentalomas. French Society of Endocrinology Consensus. *Annales d'Endocrinologie* 2008 **69** 487–500. (doi:10.1016/j.ando.2008.09.003)
- Mantero F, Terzolo M, Arnaldi G, Osella G, Masini AM, Ali A, Giovagnetti M, Opocher G & Angeli A. A survey on adrenal incidentaloma in Italy. Study Group on Adrenal Tumors of the Italian Society of Endocrinology. *Journal of Clinical Endocrinology and Metabolism* 2000 **85** 637–644. (doi:10.1210/jcem.85.2.6372)
- Valli N, Catargi B, Ronci N, Vergnot V, Leccia F, Ferriere JM, Chene G, Grenier N, Laurent F & Tabarin A. Biochemical screening for subclinical cortisol-secreting adenomas amongst adrenal incidentalomas. *European Journal of Endocrinology* 2001 **144** 401–408. (doi:10.1530/eje.0.1440401)
- Oxman AD, Glasziou P & Williams JJ. What should clinicians do when faced with conflicting recommendations? *BMJ* 2008 **337** a2530. (doi:10.1136/bmj.a2530)
- Masserini B, Morelli V, Bergamaschi S, Ermetici F, Eller-Vainicher C, Barbieri AM, Maffini MA, Scillitani A, Ambrosi B, Beck-Peccoz P *et al.* The limited role of midnight salivary cortisol levels in the diagnosis of subclinical hypercortisolism in patients with adrenal incidentaloma. *European Journal of Endocrinology* 2009 **160** 87–92. (doi:10.1530/EJE-08-0485)
- Nunes ML, Vattaut S, Corcuff JB, Rault A, Loiseau H, Gatta B, Valli N, Letenneur L & Tabarin A. Late-night salivary cortisol for diagnosis of overt and subclinical Cushing's syndrome in hospitalized and

- ambulatory patients. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 456–462. (doi:10.1210/jc.2008-1542)
- 23 Doi M, Sekizawa N, Tani Y, Tsuchiya K, Kouyama R, Tateno T, Izumiyama H, Yoshimoto T & Hirata Y. Late-night salivary cortisol as a screening test for the diagnosis of Cushing's syndrome in Japan. *Endocrine Journal* 2008 **55** 121–126. (doi:10.1507/endocrj.K07E-023)
 - 24 Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM & Woolf SH. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? *Journal of the American Medical Association* 2001 **286** 1461–1467. (doi:10.1001/jama.286.12.1461)
 - 25 Shaneyfelt TM & Centor RM. Reassessment of clinical practice guidelines: go gently into that good night. *Journal of the American Medical Association* 2009 **301** 868–869. (doi:10.1001/jama.2009.225)
 - 26 Shekelle PG, Woolf SH, Eccles M & Grimshaw J. Clinical guidelines: developing guidelines. *BMJ* 1999 **318** 593–596. (doi:10.1136/bmj.318.7183.593)
 - 27 Aarts MC, van der Heijden GJ, Rovers MM & Grolman W. Remarkable differences between three evidence-based guidelines on management of obstructive sleep apnea-hypopnea syndrome. *Laryngoscope* 2013 **123** 283–291. (doi:10.1002/lary.23521)
 - 28 Libe R, Dall'Asta C, Barbetta L, Baccarelli A, Beck-Peccoz P & Ambrosi B. Long-term follow-up study of patients with adrenal incidentalomas. *European Journal of Endocrinology* 2002 **147** 489–494. (doi:10.1530/eje.0.1470489)
 - 29 Swiglo BA, Murad MH, Schunemann HJ, Kunz R, Vigersky RA, Guyatt GH & Montori VM. A case for clarity, consistency, and helpfulness: state-of-the-art clinical practice guidelines in endocrinology using the grading of recommendations, assessment, development, and evaluation system. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 666–673. (doi:10.1210/jc.2007-1907)
 - 30 Mechanick JI, Bergman DA, Braithwaite SS & Palumbo PJ. American Association of Clinical Endocrinologists protocol for standardized production of clinical practice guidelines. *Endocrine Practice* 2004 **10** 353–361. (doi:10.4158/EP.10.4.353)

Received 28 April 2014

Revised version received 18 June 2014

Accepted 1 July 2014