



National Guidelines – How evidence based are they?

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2013-004278
Article Type:	Research
Date Submitted by the Author:	17-Oct-2013
Complete List of Authors:	Baird, Alastair; The White House, Lawrence, James; Dumfries & Galloway Royal Infirmary, Research & Development Support Unit
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	General practice / Family practice
Keywords:	HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical governance < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11 **National Guidelines – How evidence based are they?**
12
13
14
15
16
17
18
19
20
21
22
23
24

25 Dr A Gordon Baird, MB ChB MRCOG FRCGP R&D General
26 Practice Adviser, Research and Development Support Unit,
27 Dumfries and Galloway Royal infirmary, Bankend Road,
28 Dumfries DG1 4AP.
29
30
31

32
33
34
35
36 Dr James R Lawrence, BSc MD FRCP(Glas & Edin) R&D Clinical
37 Lead, Research and Development Support Unit, Dumfries and
38 Galloway Royal infirmary, Bankend Road, Dumfries DG1 4AP.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Background

National guidelines have become an integral part of UK general practice and aim to deliver quality and consistency within practices. As with any intervention, there are negative as well as positive consequences. Guideline effectiveness depends on the quality of evidence used.

Aim

To quantify and analyse the quality of evidence that is presented in national guidelines.

Design and setting

Levels of evidence used in all the current valid recommendations in the Scottish Intercollegiate Guideline Network (SIGN) guidelines were reviewed and statistically analysed.

Method

The data was collected from published guidelines available online to the public. A professional group selected by a national organisation develops each of these guidelines. Statistical analysis of the relationship between the number of guideline recommendations and the quality of evidence used in its recommendations was performed.

Result

A significant correlation between the number of recommendations in a guideline and the use of level D evidence was discovered.

Conclusion

Practice guidelines should be brief and based on scientific evidence. Paradoxically the longest guidelines have the highest proportion of recommendations based on the lowest level of evidence. Guideline developers should be more aware of the need for brevity and a stricter application of evidence-based principles could achieve this. The findings support calls for a review of how evidence is used and presented in guidelines.

Dr A Gordon Baird affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; any discrepancies from the study as planned have been explained.

The data is in the public domain; there were no ethical conflicts; there was no funding.

Article Summary

Article focus

Examines the hypothesis that larger guidelines rely disproportionately on poor evidence.

Key Messages

- One third of current national guidelines is supported only by case reports, case studies and expert opinion.
- Guidelines with large numbers of recommendations are more likely to use weak evidence.
- Guideline development groups appear to vary in their approach to offering recommendations.
- Guideline recommendations should be based on good evidence.
- Without good science scientific evidence, avoiding a recommendation or highlighting the need for research should be considered.

Strengths and limitations of the study

Strengths.

- This is the first objective evidence of inconsistencies in approach by national guideline developers
- This supports commentator suggestion that even without good evidence a group will prefer consensus.
- Adds to the current debate about how guidelines might be developed in the future

Limitations.

- The study is limited to only one set of national guidelines.
- Reasons for the differences in quality of evidence preferred by the guideline development groups is unclear

SIGN Guidelines – How evidence based are they?

INTRODUCTION

The Scottish Intercollegiate Guidelines Network (SIGN) was founded in 1993. It is a national body, professionally led and publicly funded. SIGN's founding principles proposed direct links between evidence and recommendations, offering a brief and succinct quick-reference guide for clinicians¹. Guidelines anticipated presenting brief, evidence based clinical advice. They have developed into long and authoritative texts that are used by managers and politicians to inform policy. SIGN has no responsibility to consider cost-effectiveness and no direct input into the Quality Outcomes Framework (QOF). A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003.

The World Health Organisation (WHO) recognises that current recommendation categories may be ambiguous² and has encouraged guideline developers to use a different form of grading, including a category "Use only in the context of research" where doubt exists.

Guideline developers have conflict of interest policies reported as challenging to apply. Where doubt exists, groups of specialists may feel consensus more defensible than acknowledging uncertainty.³

Even with the best evidence, concerns are expressed about the relevance of guidelines in treating patients with multiple morbidities⁴, and the emergence of the phenomenon of reversal^{5,6}, where established practice, sometimes evidence based, is shown to be sub-optimal or harmful. This study looks at the quality of evidence used for SIGN guidelines, and describes a significant trend for some groups to emphasise poorly evidence-backed recommendations.

METHODS

SIGN guidelines were accessed online. Guidelines that were “Current” and “Current 3-7 years. Some recommendations may be out of date.” in September 2013 were included. Those that had been “Withdrawn”, “Recommendations being updated”, “Need for update being considered” and those with no recommendations were excluded. Key recommendations and implementation recommendations were excluded.

SIGN guideline 50 describes an established process for developing guidelines.⁷ This process is independent of this study, but is stated to be an objective process. SIGN guidelines have four grades of recommendation outlined in table 1. Table 2 describes the level of evidence supporting the recommendation grading.

The level of evidence used by each guideline was independently recorded by 3 investigators and errors resolved. A statistical analysis of the correlation between the proportion of level D evidence and the total number of recommendations was performed for the 42 guidelines.

RESULTS

The 42 guidelines consisted of 2559 pages, ranging from 26 to 161 (median 59.5) pages. The longest guideline, number 116 was 61 pages longer than the next largest. The number of recommendations per page ranged from 0.2 to 1.8 (median 0.7). The number of recommendations per guideline is presented in table 3.

Of the 1999 recommendations, 480 (24.0%) were level A, 491(24.6 %) were level B, 318 (15.9%) level C, and 710 (35.5%) level D. Thus 40.2% were poorly evidenced (C&D) and over a third (D) developed almost entirely on “expert opinion”. The number of level A recommendations per guideline ranged from 0-57 (median 9), level B from 2-62 (median 8.5) level C ranged from 0-26 (median 6) and D from 0-60 (median 14.5). 4 guidelines had no level A evidence.

1
2
3
4
5 The proportion of level D evidence increases with the number of recommendations
6 made. This correlation is significant with Kendall's Tau=0.22 [approximate 95%
7 confidence interval 0.008-0.45] p value =0.04, and Spearman rho=0.22 [approximate
8 95% CI 0.02-0.57] p value= .04.
9
10

11 12 13 DISCUSSION

14 This study reveals that expert groups who produce long guidelines rely on poor
15 evidence more heavily than others. National guidelines are useful and important and
16 there is a debate about how evidence is best presented. Guidelines define standards
17 of care, help busy clinicians and allow managers and politicians to develop
18 governance. An American study (using 3 not 4 levels of evidence) similarly found that
19 48% were "based on expert opinion, case studies, or standards of care."⁸ Where
20 patients are involved in clinical decisions, honestly declaring uncertainty has merit.
21 In the absence of good scientific evidence, recommending a course of action without
22 understanding the circumstances of the individual to whom it is applied seems both
23 risky and, assuming patient choice, imbalanced.
24
25
26
27
28
29
30
31
32
33

34 This study did not examine why longer guidelines use poorer evidence. It has been
35 postulated that there is security in "just doing what everyone else is doing – even if
36 what everyone else is doing isn't very good."³ Cloistered groups of experts may view
37 their own opinion as more authoritative than science can support. Reliance on
38 expert opinion has a poor track record. Blinded by certainty, expert groups defining
39 established practice have perpetuated radical mastectomy instead of conservative
40 surgery, Class 1C antiarrhythmics⁹, pulmonary artery catheters in heart failure¹⁰,
41 electronic foetal monitoring in low risk pregnancies: even then practice can take a
42 decade to reverse¹¹.
43
44
45
46
47
48
49
50
51

52 Even good evidence is subject to the phenomenon of reversal where new evidence
53 contradicts current practice. Reversal can affect around 13-16% of publications^{5,6}.
54 This may partly explain why the implementation of even the most soundly evidence
55 based national guidelines fails to improve outcome^{12,13,14}. There is potential
56
57
58
59
60

1
2
3 harm^{15,16} from guidelines in real clinical settings, for example increasing radiation
4 dose without benefit¹⁷ or increased risks of anticoagulation¹⁸.
5
6
7

8
9 SIGN 116 (diabetes), is a notable exception. It is more than 50% larger than the next
10 largest, 2.5 times longer than the average and yet uses the 4th lowest level D
11 recommendations. There are a number of hypotheses why this group reports
12 differently. The advisory committee to SIGN's English equivalent (NICE) informs
13 Quality Outcomes Framework (QOF) policy. Diabetes is the largest clinical UK QOF
14 indicator and is associated with substantial payment incentives. The need for
15 objective evaluation of performance drives a use of surrogate outcomes without
16 appropriate clinical endpoints.¹⁹ Diabetes guidelines have suffered several
17 noteworthy reversals. Examples include the recommendation of glycosylated
18 haemoglobin reduction resulting in increased use of rosiglitazone (still mentioned in
19 the current document) both associated with harm²⁰, including mortality.²¹ Aspirin
20 recommendations have also been changed from previous guidelines. Is it possible
21 that the repeated use of surrogate outcomes arises from group dynamics driven by a
22 powerful external agenda?
23
24
25
26
27
28
29
30
31
32
33

34
35 Many doctors whose expertise cross several guidelines^{22,23}, express concerns about
36 guideline development groups. The inappropriate exclusion of disease groups from
37 general population data is common. Smoking cessation advice for testicular cancer
38 survivors is level C, although studies in the general population (without excepting
39 specific disease groups) advises everyone to stop smoking. Overall smoking cessation
40 was level D and C once each and B on three occasions. Using evidence in this way
41 may imply group dysfunction. Differently constituted groups, or greater oversight
42 might avoid problems.
43
44
45
46
47
48
49

50
51 In 1993, SIGN guidelines stated intention was to be evidence based, brief and
52 succinct. Brevity increases value as a quick reference guide. Removing or reducing
53 poorly evidenced recommendations would reduce size by more than a third overall
54 and in some up to two thirds. 2559 pages is longer than the Oxford Textbook of
55 Primary Medical Care. Evidence based medicine is described as "the use of
56
57
58
59
60

1
2
3 mathematical estimates of the risk of benefit and harm, derived from high-quality
4 research on population samples, to inform clinical decision-making in the diagnosis,
5 investigation or management of individual patients"²⁴. Guidelines relevance to daily
6 practice, the reliability of evidence and whether the application of evidence will
7 improve outcomes are important questions.
8
9
10

11
12
13 These results may reflect how professional groups deal with uncertainty. If so, this is
14 not good for individual patients faced with the same uncertainties (whether aware
15 of it or not), nor is it good for scientists who actively seek unanswered questions by
16 challenging established practice, an area in which medicine has a poor record from
17 Semmelweis to the present day.
18
19
20
21

22
23
24 The finding of a significant increase of level D recommendations in larger guidelines
25 has not happened by chance. A wider debate about how guideline groups can create
26 greater clarity about the reliability of evidence used is needed.²⁵ Reducing the use of
27 poorly evidenced recommendations has potential to create a shorter, more reliable
28 and usable clinical support.
29
30
31
32

33
34
35
36
37 Acknowledgements.

38
39 Heather Barrington, Statistical Adviser; Bridget Bird, Administrative Assistant;
40 Research and Development Support Unit, Dumfries and Galloway Royal infirmary,
41 Bankend Road, Dumfries DG1 4AP.
42
43
44
45

46 Conflict of interest; none
47
48
49
50
51
52
53
54

55 Table 1
56
57
58
59
60

GRADES OF RECOMMENDATION

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table 2

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	-Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies
2+	High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

Table 3

Main Recommendations

Number	Name	Pages	A	B	C	D	Total	%age D
133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	Management of diabetes	161	57	62	23	16	158	10.1%
115	Management of Obesity	87	6	11	7	11	35	31.4%
114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	Management of stable angina	59	13	10	3	11	37	29.7%
95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%
87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
		TOTAL	2559	480	491	318	710	1999

Contributorship Statement

Both authors discussed the hypothesis that there was a disproportionate use of poor evidence in longer guidelines; both checked the raw data, and agreed on a statistical approach to discover whether the trend was significant or not. Both have been involved in writing and researching the evidence.

Funding.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing.

Technical appendix, statistical code, and dataset available from the corresponding author, who will provide a permanent, citable and open access home for the dataset.

Competing interests

No competing interests

MeSH Headings

- Practice Guideline N04.761.700.350.650
- General Practice H02.403.340
- Evidence-Based Medicine H02.249.750

How this fits in.

Guidelines should encourage an evidence-based approach to clinical practice. Longer guidelines used significantly higher levels of poor evidence. WHO has proposed a different system of grading evidence. The effect of group behavior altering guideline development has been hypothesised. New research often challenges established clinical practice. Improving the quality of evidence, acknowledging uncertainty and shortening guideline length would make guidelines more clinically relevant and effective.

REFERENCES

1 Harbour R, Lowe G, Twaddle S. Scottish Intercollegiate Guidelines Network; the
2 first 15 years (1993-2008). J R Coll Physicians Edinb 2011 Jun;41(2) 163-168

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

2 de Joncheere K, Hill S, Klazinga N, Mäkelä M, Oxman AD. The Clinical Guideline
Programme of the National Institute for Health and Clinical Excellence (NICE) A
review by the World Health Organization May 2006

3 Lenzer J. Why we can't trust clinical guidelines: BMJ 2013;346:f3830

4 Aylett V. Do geriatricians need guidelines? BMJ 2010 Sep;341 :c5340

5 Prasad V, Cifu, A, Ioannidis JP. Reversals of Established Medical Practices: Evidence
to Abandon Ship. JAMA. 2012 Jan4;307(1):37-38,

6 Prasad V, Gall V, Cifu A. The frequency of medical reversal. Arch Int Med. 2011 Oct
10;. 171(18): 1675-1676

7 <http://www.sign.ac.uk/pdf/sign50.pdf>

8 Tricoci P, Allen JM, Kramer JM, Califf RM, Smith Jr. SC. Scientific evidence
underlying the ACC/AHA clinical practice guidelines. JAMA 2009 Feb25; 301 (8): 831-
841

9 Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, Arensberg
D, Baker A, Friedman L, Greene HL, Huther ML, Richardson DW. Mortality and
morbidity in patients receiving encainide, flecainide, or placebo - The Cardiac
Arrhythmia Suppression Trial. New Eng J Med. 1991 Mar21; 324 (12): 781-788

10 Binanay C, Califf RM, Hasselblad V, O'Connor CM, Shah MR, Sopko G, Stevenson
LW, Francis GS, Leier CV, Miller LW Evaluation study of congestive heart failure and
pulmonary artery catheterization effectiveness: the ESCAPE trial. JAMA. 2005 Oct
5;294(13):1625-33.

1
2
3
4 11 Tatsioni A, Bonitsis NG, Ioannidis JA. Persistence of Contradicted Claims in the
5 Literature. JAMA. 2007 Dec5 ;298(21):2517-2526.
6

7
8 12 Anderson JE, McKenzie C, Singh N, Gajree N, Giles L, Sharma P, McDonald A.
9 Compliance with the 62 day target does not improve long-term survival. Association
10 of Coloproctology of Great Britain and Ireland Annual Meeting 2012 Dublin Ireland.
11

12 13 Kerr J, Smith R, Gray S, Beard D, Robertson CE. An audit of clinical practice in the
14 management of head injured patients following the introduction of the Scottish
15 Intercollegiate Guidelines Network (SIGN) recommendations. Emerg Med J.2005
16 Dec; 22(12):850-854
17

18
19 14 Caplan LR. How well does "evidence-based" medicine help neurologists care for
20 individual patients? Rev Neurol Dis.2007;4(2):75-84
21

22
23 15 Hutchison G. Guidelines can harm patients too. BMJ (Clinical research ed.). 2012
24 Apr18;344: e2685.
25

26
27 16 Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines.
28 Potential benefits, limitations, and harms of clinical guidelines. BMJ 1999
29 Feb20;318(7182):527-530
30

31
32 17 Miller L., Kidd A. Are sign guidelines key in the decision to CT? Post-head injury CT
33 scanning within a paediatric population. Academic Emergency Medicine.
34 Conference: 14th International Conference on Emergency Medicine, ICEM 2012
35 Dublin Ireland. Conference Publication: 19 (6) (pp 776), 2012. Date of Publication:
36 June 2012.
37

38
39 18 Thomson R., Eccles M., Wood R., Chinn D.J. A cautionary note on data sources for
40 evidence-based clinical decisions: Warfarin and stroke prevention. MedDecis
41 Making. 2007;27 (4): 438-447
42

43
44 19 Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. BMJ. 2011
45 Dec28;343:d7995
46

47
48 20 Cohen D. Rosiglitazone: what went wrong? BMJ.2010 Sep; 341:c4848
49

50
51 21 Gerstein HC, Miller ME, Byington RP, Goff DC Jr., Bigger JT, Buse JB, Cushman WC,
52 Genuth S, Ismail-Beigi F, Grimm RH Jr. Probstfield JL, Simons-Morton DG, Friedewald
53 WT. Effects of intensive glucose lowering in type 2 diabetes. Action to Control
54 Cardiovascular Risk in Diabetes Study Group. New Engl J Med. 2008
55 Jun12;358(24):2545-59
56
57
58
59
60

22 Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of implementation of clinical guidelines' prescribing recommendations in primary care. Health Policy. 2008 Feb; 85 (2) 148-161

23 Guy S, Wardlaw JM. Who writes guidelines, and who should? ClinRadiol. 2002 Oct ;57 (10): 891-897.

24 Greenhalgh, Trisha. How To Read a Paper: The Basics of Evidence-Based Medicine. Wiley-Blackwell, fourth edition, 2010, p. 1

25 Zuiderent-Jerak T, Forland F, Macbeth F. Guidelines should reflect all knowledge, not just trials
BMJ 2012;345:e6702

TABLE 2

Number	Name	Pages	Main Recommendation	
			A	B
133	Management of hepatitis C	57	20	24
132	Long-term follow-up of survivors of childhood cancer	62	0	7
131	Management of schizophrenia	64	10	19
130	Brain injury rehabilitation in adults	68	0	14
129	Antithrombotics: indication and management	68	25	11
127	Management of perinatal mood disorders	47	0	5
126	Diagnosis and management of colorectal cancer	56	11	19
125	Management of atopic eczema in primary care	34	3	5
124	Management of adult testicular germ cell tumours	63	6	6
123	Management of early rheumatoid arthritis	27	3	7
122	Prevention and management of venous thromboembolism	88	26	15
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16
120	Management of chronic venous leg ulcers	46	5	3
119	Management of patients with stroke: identification and management of dysphagia	42	0	6
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29
117	Management of sore throat and indications for tonsillectomy	37	9	3
116	Management of diabetes	161	57	62
115	Management of Obesity	87	6	11
114	Nonpharmaceutical management of depression	37	5	4
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4
111	Management of hip fracture in old people	49	10	9
110	Early management of patients with a head injury	76	1	7
109	Management of genital Chlamydia trachomatis infection	40	3	6
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27
107	Diagnosis and management of headache in adults	81	17	16
106	Control of pain in adults with cancer	71	5	7
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5
103	Diagnosis and management of chronic kidney disease	50	9	6
102	Management of invasive meningococcal disease in children and young people	46	1	4
99	Management of cervical cancer	73	1	13
97	Risk estimation and the prevention of cardiovascular disease	72	16	12
96	Management of stable angina	59	13	10

95	Management of chronic heart failure	55	9	12	
94	Cardiac arrhythmias and coronary heart disease	42	22	11	
93	Acute coronary syndromes	60	11	14	
91	Bronchiolitis in children	42	4	3	
90	Diagnosis and management of head and neck cancer	92	42	8	
89	Diagnosis and management of peripheral arterial disease	37	11	2	
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	
87	Management of oesophageal and gastric cancer	70	3	26	
61	Investigation of postmenopausal bleeding	26	2	7	
		<u>TOTAL</u>	<u>2559</u>	<u>480</u>	<u>491</u>

TABLE 1.

GRADES OF RECOMMENDATION	
A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

TABLE 2

Levels of Evidence

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	-Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies
2+	High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

TABLE 3

Results -Number of recommendations in each category.

Number	Number	Name	Pages	Main Recommendations				Total	%age D
				A	B	C	D		
133	133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	116	Management of diabetes	161	57	62	23	16	158	10.1%
115	115	Management of Obesity	87	6	11	7	11	35	31.4%
114	114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%
113	113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	96	Management of stable angina	59	13	10	3	11	37	29.7%
95	95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%
87	87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
			<u>2559</u>	<u>480</u>	<u>491</u>	<u>318</u>	<u>710</u>	<u>1999</u>	



Guidelines – is bigger better? A review of SIGN guidelines.

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2013-004278.R1
Article Type:	Research
Date Submitted by the Author:	04-Dec-2013
Complete List of Authors:	Baird, Alastair; The White House, Lawrence, James; Dumfries & Galloway Royal Infirmary, Research & Development Support Unit
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	General practice / Family practice
Keywords:	HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical governance < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

1
2
3
4
5 **SIGN Guidelines –analysis of evidence levels for their**
6 **recommendations.**
7

8
9 *Alternatively*

10
11
12 **Guidelines – is bigger better? A review of SIGN national**
13 **guidelines.**
14

15
16
17
18 Dr A Gordon Baird, MB ChB MRCOG FRCGP R&D General
19 Practice Adviser, Research and Development Support Unit,
20 Dumfries and Galloway Royal infirmary, Bankend Road,
21 Dumfries, Scotland, UK DG1 4AP.
22
23
24

25
26 Address for correspondence; Dr A Gordon Baird, The White
27 House, Sandhead, Wigtownshire UK. DG9 9JA. Email
28 gordon.baird@me.com. Phone/fax 01776830281
29
30
31

32
33 Dr James R Lawrence, BSc MD FRCP(Glas & Edin) R&D Clinical
34 Lead, Research and Development Support Unit, Dumfries and
35 Galloway Royal infirmary, Bankend Road, Dumfries, Scotland,
36 UK DG1 4AP.
37
38
39

40 MeSH headings

41 General practice / Family practice

42 HEALTH SERVICES ADMINISTRATION & MANAGEMENT

43 Protocols & guidelines

44 Quality in health care

45 Clinical governance
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Background

National guidelines have become an integral part of UK general practice and aim to effectively deliver quality and consistency in clinical practice. As with any intervention, there are negative as well as positive consequences. Guideline effectiveness depends on the quality of evidence used.

Aim

To quantify and analyse the quality of evidence that is presented in national guidelines.

Design and setting

Levels of evidence used in all the current valid recommendations in the Scottish Intercollegiate Guideline Network (SIGN) guidelines were reviewed and statistically analysed.

Method

The data was collected from published guidelines available online to the public. SIGN methodology entails a professional group selected by a national organisation to develop each of these guidelines. Statistical analysis of the relationship between the number of guideline recommendations and the quality of evidence used in its recommendations was performed.

Result

A significant correlation between the number of recommendations in a guideline and the use of level D evidence was discovered.

Conclusion

Practice guidelines should be brief and based on scientific evidence. Paradoxically the longest guidelines have the highest proportion of recommendations based on the lowest level of evidence. Guideline developers should be more aware of the need for brevity and a stricter application of evidence-based principles could achieve this. The findings support calls for a review of how evidence is used and presented in guidelines.

Article Summary

Article focus

Examines the hypothesis that larger guidelines rely disproportionately on poor evidence.

Key Messages

- One third of current national guidelines are supported only by case reports, case studies and expert opinion.
- Guidelines with large numbers of recommendations used a higher proportion of weak evidence.
- Guideline development groups appear to vary in their approach to offering recommendations.
- Guideline recommendations should be based on good evidence.
- Paucity of evidence should highlight topics for research.

Strengths and limitations of the study

Strengths.

- This is the first objective evidence of inconsistencies in approach by a national guideline developers
- This supports commentator suggestion that even without good evidence a group will prefer consensus.
- Adds to the current debate about how guidelines might be developed in the future

Limitations.

- The study is limited to only one set of national guidelines (SIGN).
- Reasons for the differences in quality of evidence preferred by the guideline development groups is unclear

Guidelines – is bigger better? A review of SIGN guidelines.

INTRODUCTION

The Scottish Intercollegiate Guidelines Network (SIGN) was founded in 1993. It is a national body, professionally led and publicly funded. SIGN's founding principles proposed direct links between evidence and recommendations, offering a brief and succinct quick-reference guide for clinicians¹. Guidelines anticipated presenting brief, evidence based clinical advice. They have developed into long and authoritative texts used by managers and politicians to inform policy. SIGN has responsibility to consider cost-effectiveness and directly inputs to the Quality Outcomes Framework (QOF). A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003.

The World Health Organisation (WHO) recognises that current grades of recommendation (Table 1) may be ambiguous² and encourages guideline developers to use a system which includes a category "Use only in the context of research" where doubt exists.

Guideline developers have conflict of interest policies reported as challenging to apply. Where doubt exists, groups of specialists may feel consensus more defensible than acknowledging uncertainty.³

Even with the best evidence, concerns are expressed about the relevance of guidelines in treating patients with multiple morbidities⁴, and the emergence of the phenomenon of reversal^{5,6}, where established practice, sometimes evidence based, is shown to be sub-optimal or harmful. This study looks at the quality of evidence used for SIGN guidelines, and describes a significant trend for some groups to emphasise poorly evidence-backed recommendations.

METHODS

SIGN guidelines were accessed online in September 2013. SIGN guidelines were chosen because they are internationally respected, the authors were familiar with

1
2
3 their format and they contribute to national government policy. Guidelines that
4 were "Current" and "Current 3-7 years. Some recommendations may be out of
5 date." were included. Those that had been "Withdrawn", "Recommendations being
6 updated", "Need for update being considered" and those with no recommendations
7 were excluded.
8
9
10

11
12
13 SIGN guideline 50 clearly describes an established process for developing
14 guidelines.⁷ It explains how the process is planned, how it is implemented and by
15 whom. This process is independent of this study, but is stated to be an objective
16 process. SIGN guidelines have four grades of recommendation outlined in table 1.
17 Table 2 describes the level of evidence SIGN uses to support the recommendation
18 grading. SIGN guideline development groups vary in size depending on the scope of
19 the topic under consideration, but generally comprise between 15 and 25 members.
20 SIGN states they are aware of the many psychosocial factors, including the problems
21 of overcoming professional hierarchies that can affect small group processes.
22
23
24
25
26
27
28
29
30

31 The level of evidence used by each examined guideline was independently
32 enumerated by 3 investigators and discrepancies resolved. They discounted any
33 duplication implicit in text-embedded key recommendations and also
34 implementation recommendations. A statistical analysis of the correlation between
35 the proportion of level D evidence and the total number of recommendations was
36 performed for the 42 guidelines.
37
38
39
40
41
42

43 RESULTS

44 The 42 guidelines consisted of 2559 pages (including references), ranging from 26 to
45 161 (median 59.5) pages. The longest guideline, number 116 was 61 pages longer
46 than the next largest. The number of recommendations per page ranged from 0.2 to
47 1.8 (median 0.7). The number of recommendations per guideline is presented in
48 table 3.
49
50
51
52
53
54

55
56 Of the 1999 recommendations, 480 (24.0%) were level A, 491(24.6 %) were level B,
57 318 (15.9%) level C, and 710 (35.5%) level D. Thus 51.4% were poorly evidenced
58
59
60

1
2
3 (C&D) and over a third (D) developed almost entirely on “expert opinion”. The
4 number of level A recommendations per guideline ranged from 0-57 (median 9),
5 level B from 2-62 (median 8.5) level C ranged from 0-26 (median 6) and D from 0-60
6 (median 14.5). 4 guidelines had no level A evidence.
7
8
9

10
11
12 The proportion of level D evidence increases with the number of recommendations
13 made. This correlation is significant with Kendall’s Tau=0.22 [approximate 95%
14 confidence interval 0.008-0.45] p value =0.04, and Spearman rho=0.22 [approximate
15 95% CI 0.02-0.57] p value= .04.
16
17
18

19 20 21 DISCUSSION

22 This study reveals that expert groups who produce long guidelines rely on poor
23 evidence more heavily than others. While this study only looks at SIGN, this study
24 highlights a problem that has escaped national guideline developers, a wide range of
25 professionals and the public to whom these guidelines are applied. National
26 guidelines are useful and important and there is a debate about how evidence is
27 best presented. Guidelines define standards of care, help busy clinicians and allow
28 managers and politicians to develop governance. An American study (using 3 not 4
29 levels of evidence) similarly found that 48% were “based on expert opinion, case
30 studies, or standards of care.”⁸; we show comparable results for current SIGN
31 guidelines. Where patients are involved in clinical decisions, honestly declaring
32 uncertainty has merit. In the absence of good scientific evidence, recommending a
33 course of action without understanding the circumstances of the individual to whom
34 it is applied seems both risky and, assuming a right to patient choice, unwarranted.
35 Other guidelines that use high levels of poor evidence should evaluate the
36 proportion of poorly evidenced recommendations and seek explanations for such
37 trends.
38
39
40
41
42
43
44
45
46
47
48
49
50

51
52 This study did not examine why longer guidelines use poorer evidence. Cloistered
53 groups of experts may view their own opinion as more authoritative than science
54 can support. It has been postulated that there is security in “just doing what
55 everyone else is doing – even if what everyone else is doing isn’t very good.”³
56
57
58
59
60

1
2
3 Reliance on expert opinion has a poor track record. Blinded by certainty, expert
4 groups defining established practice have, in the past, perpetuated radical
5 mastectomy instead of conservative surgery, Class 1C antiarrhythmics⁹, pulmonary
6 artery catheters in heart failure¹⁰, electronic foetal monitoring in low risk
7 pregnancies: even then practice can take a decade to reverse¹¹.
8
9

10
11
12
13
14 Even good evidence is subject to the phenomenon of reversal where new evidence
15 contradicts current practice. Reversal can affect around 13-16% of publications^{5,6}.
16 This may partly explain why the implementation of even the most soundly evidence
17 based national guidelines fails to improve outcome^{12,13,14}. There is potential
18 harm^{15,16} from guidelines in real clinical settings, for example increasing radiation
19 dose without benefit¹⁷ or increased risks of anticoagulation¹⁸.
20
21
22
23
24

25
26 SIGN 116 (diabetes), is a notable exception. It is more than 50% larger than the next
27 largest, 2.5 times longer than the average and yet uses the 4th lowest level D
28 recommendations. There are a number of hypotheses why this group reports
29 differently. SIGN guidelines inform Quality Outcomes Framework (QOF) policy.
30 Diabetes is the largest clinical UK QOF indicator and is associated with substantial
31 payment incentives. The need for objective evaluation of performance drives a use
32 of surrogate outcomes without appropriate clinical endpoints.¹⁹ Diabetes guidelines
33 have suffered several noteworthy reversals. Examples include the recommendation
34 of glycosylated haemoglobin reduction resulting in increased use of rosiglitazone
35 (still mentioned in the current document) both associated with harm²⁰ including
36 mortality.²¹ Aspirin recommendations have also been changed from previous
37 guidelines. Is it possible that the repeated use of surrogate outcomes arises from
38 group dynamics driven by a powerful external agenda?
39
40
41
42
43
44
45
46
47
48
49

50
51 Many doctors whose expertise cross several guidelines^{22,23}, express concerns about
52 guideline development groups. The inappropriate exclusion of disease groups from
53 general population data is common. Smoking cessation advice for testicular cancer
54 survivors is level C, although studies in the general population (without excepting
55 specific disease groups) advises everyone to stop smoking. Overall smoking cessation
56
57
58
59
60

1
2
3 was level D and C once each and B on three occasions. Using evidence in this way
4 may imply group dysfunction. Differently constituted groups, or greater oversight
5 might avoid problems.
6
7

8
9
10 In 1993, SIGN guidelines stated intention was to be evidence based, brief and
11 succinct. Brevity increases value as a quick reference guide. Removing or reducing
12 poorly evidenced recommendations would reduce size by more than a third overall
13 and in some up to two thirds. The two volumes Oxford Textbook of Primary Medical
14 Care (2005) is a relatively brief 1420 pages, more than a thousand less than the 2559
15 pages of guidelines. Evidence based medicine is described as “the use of
16 mathematical estimates of the risk of benefit and harm, derived from high-quality
17 research on population samples, to inform clinical decision-making in the diagnosis,
18 investigation or management of individual patients”²⁴. Guidelines relevance to daily
19 practice, the reliability of evidence and whether the application of evidence will
20 improve outcomes are important questions.
21
22
23
24
25
26
27
28
29

30
31 These results may reflect how professional groups deal with uncertainty. If so, this is
32 not good for individual patients faced with the same uncertainties (whether aware
33 of it or not), nor is it good for scientists who actively seek unanswered questions by
34 challenging established practice, an area in which medicine has a poor record from
35 Semmelweis to the present day.
36
37
38
39

40
41
42 The finding of a significant increase of level D recommendations in larger guidelines
43 has not happened by chance. A wider debate about how guideline groups can create
44 greater clarity about the reliability of evidence used is needed.²⁵ Reducing the use of
45 poorly evidenced recommendations has potential to create a shorter, more reliable
46 and usable clinical support. The GRADE working group was formed in 2000.²⁶ SIGN
47 proposed a move to a new grading system in 2001.²⁷ Whether the changes in
48 process at present being considered will resolve the challenges that underpin the
49 inconsistencies we have outlined remains to be seen.
50
51
52
53
54
55
56
57
58
59
60

Table 1

GRADES OF RECOMMENDATION

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table 2

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	-Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies
2+	High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

Table 3

Main Recommendations

Number	Name	Pages	A	B	C	D	Total	%age D
133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	Management of diabetes	161	57	62	23	16	158	10.1%
115	Management of Obesity	87	6	11	7	11	35	31.4%
114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	Management of stable angina	59	13	10	3	11	37	29.7%
95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%
87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
	TOTAL	2559	480	491	318	710	1999	

Acknowledgements.

Heather Barrington, Statistical Adviser; Bridget Bird, Administrative Assistant; Research and Development Support Unit, Dumfries and Galloway Royal infirmary, Bankend Road, Dumfries DG1 4AP.

Dr A Gordon Baird affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; any discrepancies from the study as planned have been explained.

Contributorship Statement

Both authors discussed the hypothesis that there was a disproportionate use of poor evidence in longer guidelines; both checked the raw data, and agreed on a statistical approach to discover whether the trend was significant or not. Both have been involved in writing and researching the evidence.

Funding.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing.

Technical appendix, statistical code, and dataset available from the corresponding author, who will provide a permanent, citable and open access home for the dataset.

Competing interests

No competing interests

MeSH Headings

- Practice Guideline N04.761.700.350.650
- General Practice H02.403.340
- Evidence-Based Medicine H02.249.750

How this fits in.

Guidelines should encourage an evidence-based approach to clinical practice. Longer guidelines used significantly higher levels of poor evidence. WHO has proposed a different system of grading evidence. The effect of group behavior altering guideline development has been hypothesised. New research often challenges established clinical practice. Improving the quality of evidence, acknowledging uncertainty and shortening guideline length would make guidelines more clinically relevant and effective.

REFERENCES

- 1 Harbour R, Lowe G, Twaddle S. Scottish Intercollegiate Guidelines Network; the
2 first 15 years (1993-2008). *J R Coll Physicians Edinb* 2011 Jun;41(2) 163-168
- 3
4
- 5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 1 de Joncheere K, Hill S, Klazinga N, Mäkelä M, Oxman AD. The Clinical Guideline
Programme of the National Institute for Health and Clinical Excellence (NICE) A
review by the World Health Organization May 2006
- 3 Lenzer J. Why we can't trust clinical guidelines: *BMJ* 2013;346:f3830
- 4 Aylett V. Do geriatricians need guidelines? *BMJ* 2010 Sep;341 :c5340
- 5 Prasad V, Cifu, A, Ioannidis JP. Reversals of Established Medical Practices: Evidence
to Abandon Ship. *JAMA*. 2012 Jan4;307(1):37-38,
- 6 Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Int Med*. 2011 Oct
10;. 171(18): 1675-1676
- 7 SIGN 50: A guideline developer's handbook <http://www.sign.ac.uk/pdf/sign50.pdf>
- 8 Tricoci P, Allen JM, Kramer JM, Califf RM, Smith Jr. SC. Scientific evidence
underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009 Feb25; 301 (8): 831-
841
- 9 Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, Arensberg
D, Baker A, Friedman L, Greene HL, Huther ML, Richardson DW. Mortality and
morbidity in patients receiving encainide, flecainide, or placebo - The Cardiac
Arrhythmia Suppression Trial. *New Eng J Med*. 1991 Mar21; 324 (12): 781-788
- 10 Binanay C, Califf RM, Hasselblad V, O'Connor CM, Shah MR, Sopko G, Stevenson
LW, Francis GS, Leier CV, Miller LW Evaluation study of congestive heart failure and
pulmonary artery catheterization effectiveness: the ESCAPE trial. *JAMA*. 2005 Oct
5;294(13):1625-33.
- 11 Tatsioni A, Bonitsis NG, Ioannidis JA. Persistence of Contradicted Claims in the
Literature. *JAMA*. 2007 Dec5 ;298(21):2517-2526.
- 12 Anderson JE, McKenzie C, Singh N, Gajree N, Giles L, Sharma P, McDonald A.
Compliance with the 62 day target does not improve long-term survival. Association
of Coloproctology of Great Britain and Ireland Annual Meeting 2012 Dublin Ireland.
- 13 Kerr J, Smith R, Gray S, Beard D, Robertson CE. An audit of clinical practice in the
management of head injured patients following the introduction of the Scottish
Intercollegiate Guidelines Network (SIGN) recommendations. *Emerg Med J*. 2005
Dec; 22(12):850-854

1
2
3
4
5 14 Caplan LR. How well does "evidence-based" medicine help neurologists care for
6 individual patients? *Revin Neurol Dis.*2007;4(2):75-84
7

8
9 15 Hutchison G. Guidelines can harm patients too. *BMJ (Clinical research ed.)*. 2012
10 Apr18;344: e2685.
11

12 16 Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines.
13 Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999
14 Feb20;318(7182):527-530
15

16
17 17 Miller L., Kidd A. Are sign guidelines key in the decision to CT? Post-head injury CT
18 scanning within a paediatric population. *Academic Emergency Medicine*.
19 Conference: 14th International Conference on Emergency Medicine, ICEM 2012
20 Dublin Ireland. Conference Publication: 19 (6) (pp 776), 2012. Date of Publication:
21 June 2012.
22

23
24 18 Thomson R., Eccles M., Wood R., Chinn D.J. A cautionary note on data sources for
25 evidence-based clinical decisions: Warfarin and stroke prevention. *MedDecis*
26 *Making*. 2007;27 (4): 438-447
27

28
29 19 Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ*. 2011
30 Dec28;343:d7995
31

32 20 Cohen D. Rosiglitazone: what went wrong? *BMJ*.2010 Sep; 341:c4848
33

34
35 21 Gerstein HC, Miller ME, Byington RP, Goff DC Jr., Bigger JT, Buse JB, Cushman WC,
36 Genuth S, Ismail-Beigi F, Grimm RH Jr. Probstfield JL, Simons-Morton DG, Friedewald
37 WT. Effects of intensive glucose lowering in type 2 diabetes. Action to Control
38 Cardiovascular Risk in Diabetes Study Group. *New Engl J Med*. 2008
39 Jun12;358(24):2545-59
40

41
42 22 Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of
43 implementation of clinical guidelines' prescribing recommendations in primary care.
44 *Health Policy*.2008 Feb; 85 (2) 148-161
45

46
47 23 Guy S, Wardlaw JM. Who writes guidelines, and who should? *ClinRadiol*. 2002 Oct
48 ;57 (10): 891-897.
49

50 24 Greenhalgh, Trisha. *How To Read a Paper: The Basics of Evidence-Based*
51 *Medicine*. Wiley-Blackwell, fourth edition, 2010, p. 1
52

53
54 25 Zuiderent-Jerak T, Forland F, Macbeth F. Guidelines should reflect all knowledge,
55 not just trials *BMJ* 2012;345:e6702
56

57 26 GRADE working group http://www.gradeworkinggroup.org/about_us.htm
58
59
60

1
2
3
4
5 27 Harbour R, Miller J. A new system for grading recommendations in evidence
6 based guidelines. BMJ. 2001 Aug 11;323(7308):334-6.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 **SIGN Guidelines –analysis of evidence levels for their**
7 **recommendations.**
8

9
10 *Alternatively*

11
12 **Guidelines – is bigger better? A review of SIGN national**
13 **guidelines.**
14
15
16
17
18
19

20
21
22 **National Guidelines – How evidence based are they?**
23
24
25
26
27
28
29
30
31
32
33
34

35 Dr A Gordon Baird, MB ChB MRCOG FRCGP R&D General
36 Practice Adviser, Research and Development Support Unit,
37 Dumfries and Galloway Royal infirmary, Bankend Road,
38 Dumfries Scotland, UK DG1 4AP.
39

40
41
42 Address for correspondence; Dr A Gordon Baird, The White
43 House, Sandhead, Wigtownshire UK. DG9 9JA. Email
44 gordon.baird@me.com. Phone/fax 01776830281
45
46
47

48 Dr James R Lawrence, BSc MD FRCP(Glas & Edin) R&D Clinical
49 Lead, Research and Development Support Unit, Dumfries and
50 Galloway Royal infirmary, Bankend Road, Dumfries Scotland,
51 UK DG1 4AP.
52
53
54

Formatted: Justified

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

MeSH headings
General practice / Family practice
HEALTH SERVICES ADMINISTRATION & MANAGEMENT
Protocols & guidelines
Quality in health care
Clinical governance

For peer review only

Formatted: Justified

ABSTRACT

Background

National guidelines have become an integral part of UK general practice and aim to deliver quality and consistency ~~within practices~~ in clinical practice. As with any intervention, there are negative as well as positive consequences. Guideline effectiveness depends on the quality of evidence used.

Aim

To quantify and analyse the quality of evidence that is presented in national guidelines.

Design and setting

Levels of evidence used in all the current valid recommendations in the Scottish Intercollegiate Guideline Network (SIGN) guidelines were reviewed and statistically analysed.

Method

The data was collected from published guidelines available online to the public. A SIGN methodology entails a professional group selected by a national organisation develops each of these guidelines. Statistical analysis of the relationship between the number of guideline recommendations and the quality of evidence used in its recommendations was performed.

Result

A significant correlation between the number of recommendations in a guideline and the use of level D evidence was discovered.

Conclusion

Practice guidelines should be brief and based on scientific evidence. Paradoxically the longest guidelines have the highest proportion of recommendations based on the lowest level of evidence. Guideline developers should be more aware of the need for brevity and a stricter application of evidence-based principles could achieve this. The findings support calls for a review of how evidence is used and presented in guidelines.

Dr A Gordon Baird affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; any discrepancies from the study as planned have been explained.

The data is in the public domain; there were no ethical conflicts; there was no funding.

Article Summary

Article focus

Examines the hypothesis that larger guidelines rely disproportionately on poor evidence.

Key Messages

- One third of current national guidelines ~~is~~are supported only by case reports, case studies and expert opinion.
- Guidelines with large numbers of recommendations are more likely to use weak evidence.
- Guideline development groups appear to vary in their approach to offering recommendations.
- Guideline recommendations should be based on good evidence.
- Without good science scientific evidence, avoiding a recommendation or highlighting the need for research should be considered.

Strengths and limitations of the study

Strengths.

- This is the first objective evidence of inconsistencies in approach by national guideline developers
- This supports commentator suggestion that even without good evidence a group will prefer consensus.
- Adds to the current debate about how guidelines might be developed in the future

Limitations.

- The study is limited to only one set of national guidelines.
- Reasons for the differences in quality of evidence preferred by the guideline development groups is unclear

Formatted: Left

Guidelines – is bigger better? A review of SIGN guidelines.
SIGN Guidelines – How evidence based are they?

INTRODUCTION

The Scottish Intercollegiate Guidelines Network (SIGN) was founded in 1993. It is a national body, professionally led and publicly funded. SIGN’s founding principles proposed direct links between evidence and recommendations, offering a brief and succinct quick-reference guide for clinicians¹. Guidelines anticipated presenting brief, evidence based clinical advice. They have developed into long and authoritative texts ~~that are~~ used by managers and politicians to inform policy. SIGN has responsibility to consider cost-effectiveness and directly inputs to the Quality Outcomes Framework (QOF). A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003. ~~SIGN has no responsibility to consider cost-effectiveness and no direct input into the Quality Outcomes Framework (QOF). A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003.~~

The World Health Organisation (WHO) recognises that current grades of recommendation categories (Table 1) may be ambiguous² and has encouraged guideline developers to use a different form of grading, including a category “Use only in the context of research” where doubt exists.

Guideline developers have conflict of interest policies reported as challenging to apply. Where doubt exists, groups of specialists may feel consensus more defensible than acknowledging uncertainty.³

Even with the best evidence, concerns are expressed about the relevance of guidelines in treating patients with multiple morbidities⁴, and the emergence of the phenomenon of reversal^{5,6}, where established practice, sometimes evidence based,

1
2
3
4
5
6 is shown to be sub-optimal or harmful. This study looks at the quality of evidence
7 used for SIGN guidelines, and describes a significant trend for some groups to
8 emphasise poorly evidence-backed recommendations.
9
10

11 12 13 14 15 16 METHODS

17 SIGN guidelines were accessed online in September 2013. SIGN guidelines were
18 chosen because they are internationally respected, the authors were familiar with
19 their format and they contribute to national government policy. Guidelines that
20 were "Current" and "Current 3-7 years. Some recommendations may be out of
21 date." ~~in September 2013~~ were included. Those that had been "Withdrawn",
22 "Recommendations being updated", "Need for update being considered" and those
23 with no recommendations were excluded. Key recommendations and
24 implementation recommendations were excluded.
25
26
27
28
29
30

31 SIGN guideline 50 clearly describes an established process for developing
32 guidelines.⁷ It explains how the process is planned, how it is implemented and by
33 whom. This process is independent of this study, but is stated to be an objective
34 process. SIGN guidelines have four grades of recommendation outlined in table 1.
35 Table 2 describes the level of evidence SIGN uses to support the recommendation
36 grading. SIGN guideline development groups vary in size depending on the scope of
37 the topic under consideration, but generally comprise between 15 and 25 members.
38 SIGN states they are aware of the many psychosocial factors, including the problems
39 of overcoming professional hierarchies that can affect small group processes.
40 ~~SIGN guideline 50 describes an established process for developing guidelines.⁸ This~~
41 ~~process is independent of this study, but is stated to be an objective process. SIGN~~
42 ~~guidelines have four grades of recommendation outlined in table 1. Table 2~~
43 ~~describes the level of evidence supporting the recommendation grading.~~
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 The level of evidence used by each examined guideline was independently
7 enumerated by 3 investigators and discrepancies resolved~~used by each guideline~~
8 ~~was independently recorded by 3 investigators and errors resolved.~~ They discounted
9 any duplication implicit in text-embedded key recommendations and also
10 implementation Arecommendations. A statistical analysis of the correlation between
11 the proportion of level D evidence and the total number of recommendations was
12 performed for the 42 guidelines.
13
14
15
16
17

18 RESULTS

19
20 The 42 guidelines consisted of 2559 pages (including references), ranging from 26 to
21 161 (median 59.5) pages. The longest guideline, number 116 was 61 pages longer
22 than the next largest. The number of recommendations per page ranged from 0.2 to
23 1.8 (median 0.7). The number of recommendations per guideline is presented in
24 table 3.
25
26
27

28
29 Of the 1999 recommendations, 480 (24.0%) were level A, 491(24.6 %) were level B,
30 318 (15.9%) level C, and 710 (35.5%) level D. Thus 51.4% ~~40.2%~~ were poorly
31 evidenced (C&D) and over a third (D) developed almost entirely on “expert opinion”.
32 The number of level A recommendations per guideline ranged from 0-57 (median 9),
33 level B from 2-62 (median 8.5) level C ranged from 0-26 (median 6) and D from 0-60
34 (median 14.5). 4 guidelines had no level A evidence.
35
36
37
38
39

40 The proportion of level D evidence increases with the number of recommendations
41 made. This correlation is significant with Kendall’s Tau=0.22 [approximate 95%
42 confidence interval 0.008-0.45] p value =0.04, and Spearman rho=0.22 [approximate
43 95% CI 0.02-0.57] p value= .04.
44
45
46
47

48 DISCUSSION

49 This study reveals that expert groups who produce long guidelines rely on poor
50 evidence more heavily than others. While this study only looks at SIGN, this study
51 highlights a problem that has escaped national guideline developers, a wide range of
52 professionals and the public to whom these guidelines are applied. National
53
54
55
56
57
58
59
60

1
2
3
4
5
6 guidelines are useful and important and there is a debate about how evidence is
7 best presented. ~~Guidelines~~Guidelines define standards of care, help busy clinicians
8 and allow managers and politicians to develop governance. An American study
9 (using 3 not 4 levels of evidence) similarly found that 48% were “based on expert
10 opinion, case studies, or standards of care.”⁹ we show comparable results for
11 current SIGN guidelines. Where patients are involved in clinical decisions, honestly
12 declaring uncertainty has merit. In the absence of good scientific evidence,
13 recommending a course of action without understanding the circumstances of the
14 individual to whom it is applied seems both risky and, assuming a right to patient
15 choice, unwarranted~~imbalanced~~.
16
17 Other guidelines that use high levels of poor evidence should evaluate the
18 proportion of poorly evidenced recommendations and seek explanations for such
19 trends.
20
21
22
23
24
25
26
27

28 This study did not examine why longer guidelines use poorer evidence. Cloistered
29 groups of experts may view their own opinion as more authoritative than science
30 can support. It has been postulated that there is security in “just doing what
31 everyone else is doing – even if what everyone else is doing isn’t very good.”³
32
33 ~~Cloistered groups of experts may view their own opinion as more authoritative than~~
34 ~~science can support.~~ Reliance on expert opinion has a poor track record. Blinded by
35 certainty, expert groups defining established practice have perpetuated radical
36 mastectomy instead of conservative surgery, Class 1C antiarrhythmics¹⁰, pulmonary
37 artery catheters in heart failure¹¹, electronic foetal monitoring in low risk
38 pregnancies: even then practice can take a decade to reverse¹².
39
40
41
42
43
44

45 Even good evidence is subject to the phenomenon of reversal where new evidence
46 contradicts current practice. Reversal can affect around 13-16% of publications^{5,6}.
47 This may partly explain why the implementation of even the most soundly evidence
48 based national guidelines fails to improve outcome^{13,14,15}. There is potential
49 harm^{16,17} from guidelines in real clinical settings, for example increasing radiation
50 dose without benefit¹⁸ or increased risks of anticoagulation¹⁹.
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 SIGN 116 (diabetes), is a notable exception. It is more than 50% larger than the next
7 largest, 2.5 times longer than the average and yet uses the 4th lowest level D
8 recommendations. There are a number of hypotheses why this group reports
9 differently. ~~SIGN guidelines inform Quality Outcomes Framework (QOF) policy~~
10 ~~The advisory committee to SIGN's English equivalent (NICE) informs Quality Outcomes~~
11 ~~Framework (QOF) policy.~~ Diabetes is the largest clinical UK QOF indicator and is
12 associated with substantial payment incentives. The need for objective evaluation of
13 performance drives a use of surrogate outcomes without appropriate clinical
14 endpoints.²⁰ Diabetes guidelines have suffered several noteworthy reversals.
15 Examples include the recommendation of glycosylated haemoglobin reduction
16 resulting in increased use of rosiglitazone (still mentioned in the current document)
17 both associated with harm²¹ including mortality.²² Aspirin recommendations have
18 also been changed from previous guidelines. Is it possible that the repeated use of
19 surrogate outcomes arises from group dynamics driven by a powerful external
20 agenda?
21
22
23
24
25
26
27
28
29
30

31 Many doctors whose expertise cross several guidelines^{23,24}, express concerns about
32 guideline development groups. The inappropriate exclusion of disease groups from
33 general population data is common. Smoking cessation advice for testicular cancer
34 survivors is level C, although studies in the general population (without excepting
35 specific disease groups) advises everyone to stop smoking. Overall smoking cessation
36 was level D and C once each and B on three occasions. Using evidence in this way
37 may imply group dysfunction. Differently constituted groups, or greater oversight
38 might avoid problems.
39
40
41
42
43
44

45 In 1993, SIGN guidelines stated intention was to be evidence based, brief and
46 succinct. Brevity increases value as a quick reference guide. Removing or reducing
47 poorly evidenced recommendations would reduce size by more than a third overall
48 and in some up to two thirds. ~~The two volumes Oxford Textbook of Primary Medical~~
49 ~~Care (2005) is a relatively brief 1420 pages, more than a thousand less than the 2559~~
50 ~~pages of guidelines. 2559 pages is longer than the Oxford Textbook of Primary~~
51 ~~Medical Care.~~ Evidence based medicine is described as “the use of mathematical
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 estimates of the risk of benefit and harm, derived from high-quality research on
7 population samples, to inform clinical decision-making in the diagnosis, investigation
8 or management of individual patients²⁵. Guidelines relevance to daily practice, the
9 reliability of evidence and whether the application of evidence will improve
10 outcomes are important questions.
11
12
13

14
15
16 These results may reflect how professional groups deal with uncertainty. If so, this is
17 not good for individual patients faced with the same uncertainties (whether aware
18 of it or not), nor is it good for scientists who actively seek unanswered questions by
19 challenging established practice, an area in which medicine has a poor record from
20 Semmelweis to the present day.
21
22
23

24
25 The finding of a significant increase of level D recommendations in larger guidelines
26 has not happened by chance. A wider debate about how guideline groups can create
27 greater clarity about the reliability of evidence used is needed.²⁶ Reducing the use of
28 poorly evidenced recommendations has potential to create a shorter, more reliable
29 and usable clinical support. The GRADE working group was formed in 2000.²⁷ SIGN
30 proposed a move to a new grading system in 2001.²⁸ Whether the changes in
31 process at present being considered will resolve the challenges that underpin the
32 inconsistencies we have outlined remains to be seen.
33
34
35
36
37
38
39

40 Acknowledgements.

41
42 Heather Barrington, Statistical Adviser; Bridget Bird, Administrative Assistant;
43 Research and Development Support Unit, Dumfries and Galloway Royal infirmary,
44 Bankend Road, Dumfries DG1 4AP.
45
46
47

48
49 Conflict of interest; none
50
51
52
53
54
55
56
57
58
59
60

Table 1

GRADES OF RECOMMENDATION

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table 2

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	-Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies
2+	High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

Table 3

Main Recommendations

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Number	Name	Pages	A	B	C	D	Total	%age D
133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	Management of diabetes	161	57	62	23	16	158	10.1%
115	Management of Obesity	87	6	11	7	11	35	31.4%
114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	Management of stable angina	59	13	10	3	11	37	29.7%
95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%
87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
		TOTAL	2559	480	491	318	710	1999

Contributorship Statement

Both authors discussed the hypothesis that there was a disproportionate use of poor evidence in longer guidelines; both checked the raw data, and agreed on a statistical approach to discover whether the trend was significant or not. Both have been involved in writing and researching the evidence.

Funding.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing.

Technical appendix, statistical code, and dataset available from the corresponding author, who will provide a permanent, citable and open access home for the dataset.

Competing interests

No competing interests

MeSH Headings

- Practice Guideline N04.761.700.350.650
- General Practice H02.403.340
- Evidence-Based Medicine H02.249.750

How this fits in.

Guidelines should encourage an evidence-based approach to clinical practice. Longer guidelines used significantly higher levels of poor evidence. WHO has proposed a different system of grading evidence. The effect of group behavior altering guideline development has been hypothesised. New research often challenges established clinical practice. Improving the quality of evidence, acknowledging uncertainty and shortening guideline length would make guidelines more clinically relevant and effective.

REFERENCES

1 Harbour R, Lowe G, Twaddle S. Scottish Intercollegiate Guidelines Network; the first 15 years (1993-2008). J R Coll Physicians Edinb 2011 Jun;41(2) 163-168

2 de Joncheere K, Hill S, Klazinga N, Mäkelä M, Oxman AD. The Clinical Guideline Programme of the National Institute for Health and Clinical Excellence (NICE) A review by the World Health Organization May 2006

3 Lenzer J. Why we can't trust clinical guidelines: BMJ 2013;346:f3830

4 Aylett V. Do geriatricians need guidelines? BMJ 2010 Sep;341 :c5340

5 Prasad V, Cifu, A, Ioannidis JP. Reversals of Established Medical Practices: Evidence to Abandon Ship. JAMA. 2012 Jan4;307(1):37-38,

6 Prasad V, Gall V, Cifu A. The frequency of medical reversal. Arch Int Med. 2011 Oct 10;. 171(18): 1675-1676

7 [SIGN 50: A guideline developer's handbook http://www.sign.ac.uk/pdf/sign50.pdf](http://www.sign.ac.uk/pdf/sign50.pdf)

8 <http://www.sign.ac.uk/pdf/sign50.pdf>

9 Tricoci P, Allen JM, Kramer JM, Califf RM, Smith Jr. SC. Scientific evidence underlying the ACC/AHA clinical practice guidelines. JAMA 2009 Feb25; 301 (8): 831-841

10 Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, Arensberg D, Baker A, Friedman L, Greene HL, Huther ML, Richardson DW. Mortality and morbidity in patients receiving encainide, flecainide, or placebo - The Cardiac Arrhythmia Suppression Trial. New Eng J Med. 1991 Mar21; 324 (12): 781-788

11 Binanay C, Califf RM, Hasselblad V, O'Connor CM, Shah MR, Sopko G, Stevenson LW, Francis GS, Leier CV, Miller LW Evaluation study of congestive heart failure and

1
2
3
4
5
6
7 pulmonary artery catheterization effectiveness: the ESCAPE trial. JAMA. 2005 Oct
8 5;294(13):1625-33.
9

10 12 Tatsioni A, Bonitsis NG, Ioannidis JA. Persistence of Contradicted Claims in the
11 Literature. JAMA. 2007 Dec 5 ;298(21):2517-2526.
12

13 13 Anderson JE, McKenzie C, Singh N, Gajree N, Giles L, Sharma P, McDonald A.
14 Compliance with the 62 day target does not improve long-term survival. Association
15 of Coloproctology of Great Britain and Ireland Annual Meeting 2012 Dublin Ireland.
16

17 14 Kerr J, Smith R, Gray S, Beard D, Robertson CE. An audit of clinical practice in the
18 management of head injured patients following the introduction of the Scottish
19 Intercollegiate Guidelines Network (SIGN) recommendations. Emerg Med J. 2005
20 Dec; 22(12):850-854
21

22 15 Caplan LR. How well does "evidence-based" medicine help neurologists care for
23 individual patients? Rev Neurol Dis. 2007;4(2):75-84
24

25 16 Hutchison G. Guidelines can harm patients too. BMJ (Clinical research ed.). 2012
26 Apr 18;344: e2685.
27

28 17 Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines.
29 Potential benefits, limitations, and harms of clinical guidelines. BMJ 1999
30 Feb 20;318(7182):527-530
31

32 18 Miller L, Kidd A. Are sign guidelines key in the decision to CT? Post-head injury CT
33 scanning within a paediatric population. Academic Emergency Medicine.
34 Conference: 14th International Conference on Emergency Medicine, ICEM 2012
35 Dublin Ireland. Conference Publication: 19 (6) (pp 776), 2012. Date of Publication:
36 June 2012.
37

38 19 Thomson R., Eccles M., Wood R., Chinn D.J. A cautionary note on data sources for
39 evidence-based clinical decisions: Warfarin and stroke prevention. MedDecis
40 Making. 2007;27 (4): 438-447
41

42 20 Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. BMJ. 2011
43 Dec 28;343:d7995
44

45 21 Cohen D. Rosiglitazone: what went wrong? BMJ. 2010 Sep; 341:c4848
46

47 22 Gerstein HC, Miller ME, Byington RP, Goff DC Jr., Bigger JT, Buse JB, Cushman WC,
48 Genuth S, Ismail-Beigi F, Grimm RH Jr. Probstfield JL, Simons-Morton DG, Friedewald
49 WT. Effects of intensive glucose lowering in type 2 diabetes. Action to Control
50 Cardiovascular Risk in Diabetes Study Group. New Engl J Med. 2008
51 Jun 12;358(24):2545-59
52
53
54
55
56
57
58
59
60

23 Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of implementation of clinical guidelines' prescribing recommendations in primary care. *Health Policy*. 2008 Feb; 85 (2) 148-161

24 Guy S, Wardlaw JM. Who writes guidelines, and who should? *ClinRadiol*. 2002 Oct ;57 (10): 891-897.

25 Greenhalgh, Trisha. *How To Read a Paper: The Basics of Evidence-Based Medicine*. Wiley-Blackwell, fourth edition, 2010, p. 1

26 Zuiderent-Jerak T, Forland F, Macbeth F. Guidelines should reflect all knowledge, not just trials
BMJ 2012;345:e6702

TABLE 2

Number	Name	Pages	Main Recommendations				Total
			A	B	C	D	
133	Management of hepatitis C	57	20	24	7	52	103
132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30
131	Management of schizophrenia	64	10	19	3	15	47
130	Brain injury rehabilitation in adults	68	0	14	7	8	29
129	Antithrombotics: indication and management	68	25	11	6	19	61
127	Management of perinatal mood disorders	47	0	5	6	15	26
126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74
125	Management of atopic eczema in primary care	34	3	5	3	2	13
124	Management of adult testicular germ cell tumours	63	6	6	9	21	42
123	Management of early rheumatoid arthritis	27	3	7	2	0	12
122	Prevention and management of venous thromboembolism	88	26	15	14	55	110
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59
120	Management of chronic venous leg ulcers	46	5	3	4	7	19
119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78
117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20
116	Management of diabetes	161	57	62	23	16	158
115	Management of Obesity	87	6	11	7	11	35
114	Nonpharmaceutical management of depression	37	5	4	0	0	9
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17
111	Management of hip fracture in old people	49	10	9	8	14	41
110	Early management of patients with a head injury	76	1	7	6	17	31
109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101
107	Diagnosis and management of headache in adults	81	17	16	9	34	76
106	Control of pain in adults with cancer	71	5	7	3	19	34
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36
103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22
102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37
99	Management of cervical cancer	73	1	13	19	29	62
97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34

96	Management of stable angina	59	13	10	3	11	37
95	Management of chronic heart failure	55	9	12	1	1	23
94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69
93	Acute coronary syndromes	60	11	14	9	8	42
91	Bronchiolitis in children	42	4	3	6	14	27
90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136
89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30
87	Management of oesophageal and gastric cancer	70	3	26	23	28	80
61	Investigation of postmenopausal bleeding	26	2	7	4	4	17
TOTAL		2559	480	491	318	710	1999

[27 GRADE working group http://www.gradeworkinggroup.org/about_us.htm](http://www.gradeworkinggroup.org/about_us.htm)

[28 Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. BMJ. 2001 Aug 11;323\(7308\):334-6.](#)

Table 1

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table 2

1	
2	
3	
4	
5	
6	
7	
8	
9	1++ High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
10	
11	
12	1+ Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
13	
14	
15	
16	1 - -Meta-analyses, systematic reviews, or RCTs with a high risk of bias
17	
18	2++ High quality systematic reviews of case control or cohort studies
19	
20	2+ High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
21	
22	
23	2 - Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
24	
25	
26	
27	3 Non-analytic studies, eg case reports, case series
28	
29	4 Expert opinion
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	

Table 3 – for editing

Number	Number	Name	Pages	Main Recommendations				Total	%age D
				A	B	C	D		
133	133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	116	Management of diabetes	161	57	62	23	16	158	10.1%
115	115	Management of Obesity	87	6	11	7	11	35	31.4%
114	114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%

Table 3 – for editing

113	113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	96	Management of stable angina	59	13	10	3	11	37	29.7%
95	95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%

Table 3 – for editing

87	87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
<u>TOTAL</u>			<u>2559</u>	<u>480</u>	<u>491</u>	<u>318</u>	<u>710</u>	<u>1999</u>	

For peer review only



Scottish Intercollegiate Guideline Network—analysis of evidence levels for their recommendations.

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2013-004278.R2
Article Type:	Research
Date Submitted by the Author:	13-Jan-2014
Complete List of Authors:	Baird, Alastair; The White House, Lawrence, James; Dumfries & Galloway Royal Infirmary, Research & Development Support Unit
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	General practice / Family practice
Keywords:	HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical governance < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

1
2
3
4
5 **Scottish Intercollegiate Guideline Network—analysis of**
6 **evidence levels for their recommendations.**
7
8

9 *Alternatively*

10
11
12 **Guideline recommendations are more better? A review of**
13 **Scottish Intercollegiate Guideline Network national**
14 **guidelines.**
15
16
17
18
19

20 Dr A Gordon Baird, MB ChB MRCOG FRCGP R&D General
21 Practice Adviser, Research and Development Support Unit,
22 Dumfries and Galloway Royal infirmary, Bankend Road,
23 Dumfries, Scotland, UK DG1 4AP.
24
25
26
27

28 Address for correspondence; Dr A Gordon Baird, The White
29 House, Sandhead, Wigtownshire UK. DG9 9JA. Email
30 gordon.baird@me.com. Phone/fax 01776830281
31
32
33

34 Dr James R Lawrence, BSc MD FRCP(Glas & Edin) R&D Clinical
35 Lead, Research and Development Support Unit, Dumfries and
36 Galloway Royal infirmary, Bankend Road, Dumfries, Scotland,
37 UK DG1 4AP.
38
39
40
41

42 MeSH headings

43 General practice / Family practice

44 HEALTH SERVICES ADMINISTRATION & MANAGEMENT

45 Protocols & guidelines

46 Quality in health care

47 Clinical governance
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Background

National guidelines have become an integral part of UK general practice and aim to effectively deliver quality and consistency in clinical practice. As with any intervention, there are negative as well as positive consequences. Guideline effectiveness depends on the quality of evidence used.

Aim

To quantify and analyse the quality of evidence that is presented in national guidelines.

Design and setting

Levels of evidence used in all the current valid recommendations in the Scottish Intercollegiate Guideline Network (SIGN) guidelines were reviewed and statistically analysed.

Method

The data was collected from published guidelines available online to the public. SIGN methodology entails a professional group selected by a national organisation to develop each of these guidelines. Statistical analysis of the relationship between the number of guideline recommendations and the quality of evidence used in its recommendations was performed.

Result

A significant correlation between the number of recommendations in a guideline and the use of level D evidence was discovered.

Conclusion

Practice guidelines should be brief and based on scientific evidence. Paradoxically the longest guidelines have the highest proportion of recommendations based on the lowest level of evidence. Guideline developers should be more aware of the need for brevity and a stricter application of evidence-based principles could achieve this. The findings support calls for a review of how evidence is used and presented in guidelines.

Article Summary

Article focus

Examines the hypothesis that larger guidelines rely disproportionately on poor evidence.

Key Messages

- One third of current national guidelines are supported only by case reports, case studies and expert opinion.
- Guidelines with large numbers of recommendations used a higher proportion of weak evidence.
- Guideline development groups appear to vary in their approach to offering recommendations.
- Guideline recommendations should be based on good evidence.
- Paucity of evidence should highlight topics for research.

Strengths and limitations of the study

Strengths.

- This is the first objective evidence of inconsistencies in approach by a national guideline developer
- This supports commentator suggestion that even without good evidence a group will prefer consensus.
- Adds to the current debate about how guidelines might be developed in the future

Limitations.

- The study is limited to only one set of national guidelines (SIGN).
- Reasons for the differences in quality of evidence preferred by the guideline development groups is unclear

Guidelines – is bigger better? A review of SIGN guidelines.

INTRODUCTION

The Scottish Intercollegiate Guidelines Network (SIGN) was founded in 1993. It is a national body, professionally led and publicly funded. SIGN's founding principles proposed direct links between evidence and recommendations, offering a brief and succinct quick-reference guide for clinicians [1]. Guidelines anticipated presenting brief, evidence based clinical advice. They have developed into long and authoritative texts used by managers and politicians to inform policy. A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003. Both have responsibility to consider cost-effectiveness and input to the Quality Outcomes Framework (QOF).

The World Health Organisation (WHO) recognises that current grades of recommendation (Table 1) may be ambiguous[2] and encourages guideline developers to use a system which includes a category "Use only in the context of research" where doubt exists.

Guideline developers have conflict of interest policies reported as challenging to apply. Where doubt exists, groups of specialists may feel consensus more defensible than acknowledging uncertainty. [3]

Even with the best evidence, concerns are expressed about the relevance of guidelines in treating patients with multiple morbidities[4], and the emergence of the phenomenon of reversal[5,6], where established practice, sometimes evidence based, is shown to be sub-optimal or harmful. This study looks at the quality of evidence used for SIGN guidelines, and describes a significant trend for some groups to emphasise poorly evidence-backed recommendations.

METHODS

1
2
3 SIGN guidelines were accessed online in September 2013. SIGN guidelines were
4 chosen because they are internationally respected, the authors were familiar with
5 their format and they contribute to national government policy. Guidelines that
6 were "Current" and "Current 3-7 years". Some recommendations may be out of
7 date." were included. Those that had been "Withdrawn", "Recommendations being
8 updated", "Need for update being considered" and those with no recommendations
9 were excluded.

10
11
12
13
14
15
16
17 SIGN guideline 50 clearly describes an established process for developing
18 guidelines.[7] It explains how the process is planned, how it is implemented and by
19 whom. This process is independent of this study, but is stated to be an objective
20 process. SIGN guidelines have four grades of recommendation outlined in table 1.
21 Table 2 describes the level of evidence SIGN uses to support the recommendation
22 grading. SIGN guideline development groups vary in size depending on the scope of
23 the topic under consideration, but generally comprise between 15 and 25 members.
24 SIGN states they are aware of the many psychosocial factors, including the problems
25 of overcoming professional hierarchies that can affect small group processes.

26
27
28
29
30
31
32
33
34
35 3 investigators (JRL, AGB, ABB) independently enumerated the level of evidence
36 used by each guideline. They discounted any duplication implicit in text-embedded
37 key recommendations and also implementation recommendations. There were no
38 discrepancies. A statistical analysis of the correlation between the proportion of
39 level D evidence and the total number of recommendations was performed for the
40 42 guidelines.

41 42 43 44 45 46 47 RESULTS

48
49 The 42 guidelines consisted of 2559 pages (including references), ranging from 26 to
50 161 (median 59.5) pages. The longest guideline, number 116 was 61 pages longer
51 than the next largest. The number of recommendations per page ranged from 0.2 to
52 1.8 (median 0.7). The number of recommendations per guideline is presented in
53 table 3.
54
55
56
57
58
59
60

1
2
3 Of the 1999 recommendations, 480 (24.0%) were level A, 491(24.6 %) were level B,
4 318 (15.9%) level C, and 710 (35.5%) level D. Thus 51.4% were poorly evidenced
5 (C&D) and over a third (D) developed almost entirely on “expert opinion”. The
6 number of level A recommendations per guideline ranged from 0-57 (median 9),
7 level B from 2-62 (median 8.5) level C ranged from 0-26 (median 6) and D from 0-60
8 (median 14.5). 4 guidelines had no level A evidence.
9

10
11
12
13
14
15
16 The proportion of level D evidence increases with the number of recommendations
17 made. This correlation is significant with Kendall’s Tau=0.22 [approximate 95%
18 confidence interval 0.008-0.45] p value =0.04, and Spearman rho=0.22 [approximate
19 95% CI 0.02-0.57] p value= .04.
20
21
22

23 24 DISCUSSION

25
26 This study reveals that expert groups who produce long guidelines rely on poor
27 evidence more heavily than others. While this study only looks at SIGN, this study
28 highlights a problem that has escaped national guideline developers, a wide range of
29 professionals and the public to whom these guidelines are applied. National
30 guidelines are useful and important and there is a debate about how evidence is
31 best presented. Guidelines define standards of care, help busy clinicians and allow
32 managers and politicians to develop governance. An American study (using 3 not 4
33 levels of evidence) similarly found that 48% were “based on expert opinion, case
34 studies, or standards of care.” [8]; we show comparable results for current SIGN
35 guidelines. Where patients are involved in clinical decisions, honestly declaring
36 uncertainty has merit. In the absence of good scientific evidence, recommending a
37 course of action without understanding the circumstances of the individual to whom
38 it is applied seems both risky and, assuming a right to patient choice, unwarranted.
39 Other guidelines that use high levels of poor evidence should evaluate the
40 proportion of poorly evidenced recommendations and seek explanations for such
41 trends.
42
43
44
45
46
47
48
49
50
51
52
53

54
55
56 This study did not examine why longer guidelines use poorer evidence. Groups of
57 experts, indulging in “group think” may view their own opinion as more authoritative
58
59
60

1
2
3 than science can support[9]. It has been postulated that there is security in “just
4 doing what everyone else is doing – even if what everyone else is doing isn’t very
5 good.”[3] Reliance on expert opinion has a poor track record. Blinded by certainty,
6 expert groups defining established practice have, in the past, perpetuated radical
7 mastectomy instead of conservative surgery, Class 1C antiarrhythmics[10],
8 pulmonary artery catheters in heart failure[11], electronic foetal monitoring in low
9 risk pregnancies: even then practice can take a decade to reverse[12].

16
17 Even good evidence is subject to the phenomenon of reversal where new evidence
18 contradicts current practice. Reversal can affect around 13-16% of publications [5,6].
19 This may partly explain why the implementation of even the most soundly evidence
20 based national guidelines fails to improve outcome [13-15] . There is potential
21 harm[16,17] from guidelines in real clinical settings, for example increasing radiation
22 dose without benefit[18] or increased risks of anticoagulation[19].

28
29 SIGN 116 (diabetes), is a notable exception. It is more than 50% larger than the next
30 largest, 2.5 times longer than the average and yet uses the 4th lowest level D
31 recommendations. There are a number of hypotheses why this group reports
32 differently. SIGN guidelines inform Quality Outcomes Framework (QOF) policy.
33 Diabetes is the largest clinical UK QOF indicator and is associated with substantial
34 payment incentives. The need for objective evaluation of performance drives a use
35 of surrogate outcomes without appropriate clinical endpoints. [20] Diabetes
36 guidelines have suffered several noteworthy reversals. Examples include the
37 recommendation of glycosylated haemoglobin reduction resulting in increased use
38 of rosiglitazone (still mentioned in the current document) both associated with harm
39 including mortality. [21,22] Aspirin recommendations have also been changed from
40 previous guidelines. Is it possible that the repeated use of surrogate outcomes arises
41 from group dynamics driven by a powerful external agenda?

52
53
54 Many doctors whose expertise cross several guidelines[23,24] express concerns
55 about guideline development groups. The inappropriate exclusion of disease groups
56 from general population data is common. Smoking cessation advice for testicular
57
58
59
60

1
2
3 cancer survivors is level C, although studies in the general population (without
4 excepting specific disease groups) advises everyone to stop smoking. Overall
5 smoking cessation was level D and C once each and B on three occasions. Using
6 evidence in this way may imply group dysfunction. Differently constituted groups, or
7 greater oversight might avoid problems.
8
9
10

11
12
13 In 1993, SIGN guidelines stated intention was to be evidence based, brief and
14 succinct. Brevity increases value as a quick reference guide. Removing or reducing
15 poorly evidenced recommendations would reduce size by more than a third overall
16 and in some up to two thirds. The two volumes Oxford Textbook of Primary Medical
17 Care (2005) is a relatively brief 1420 pages, more than a thousand less than the 2559
18 pages of guidelines. Evidence based medicine is described as “the use of
19 mathematical estimates of the risk of benefit and harm, derived from high-quality
20 research on population samples, to inform clinical decision-making in the diagnosis,
21 investigation or management of individual patients”[25]. Guidelines relevance to
22 daily practice, the reliability of evidence and whether the application of evidence will
23 improve outcomes are important questions.
24
25
26
27
28
29
30
31
32

33
34
35 These results may reflect how professional groups deal with uncertainty. If so, this is
36 not good for individual patients faced with the same uncertainties (whether aware
37 of it or not), nor is it good for scientists who actively seek unanswered questions by
38 challenging established practice, an area in which medicine has a poor record from
39 Semmelweis to the present day.
40
41
42
43
44

45
46 The finding of a significant increase of level D recommendations in larger guidelines
47 has not happened by chance. A wider debate about how guideline groups can create
48 greater clarity about the reliability of evidence used is needed.[26] Reducing the use
49 of poorly evidenced recommendations has potential to create a shorter, more
50 reliable and usable clinical support. The GRADE working group was formed in
51 2000.[27] SIGN moved to a new grading system in 2001[28] and from 2013 a new
52 system based on GRADE principles. Whether these changes will resolve the
53 challenges that underpin the inconsistencies we have outlined remains to be seen.
54
55
56
57
58
59
60

The authors would like to acknowledge the help of Heather Barrington, Dumfries & Galloway Royal Infirmary, Research & Development Support Unit Dumfries, Dumfries & Galloway, United Kingdom who gave statistical advice and support, and Anne B Baird, The white House, Sandhead, Wigtownshire DG9 9JA who checked the data for inconsistencies.

Table 1

GRADES OF RECOMMENDATION	
A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table 2

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	-Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies
2+	High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series

4	Expert opinion
---	----------------

Table 3

Main Recommendations

For peer review only

Number	Name	Pages	A	B	C	D	Total	%age D
133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	Management of diabetes	161	57	62	23	16	158	10.1%
115	Management of Obesity	87	6	11	7	11	35	31.4%
114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	Management of stable angina	59	13	10	3	11	37	29.7%
95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%
87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
		TOTAL	2559	480	491	318	710	1999

Acknowledgements.

Heather Barrington, Statistical Adviser; Bridget Bird, Administrative Assistant; Research and Development Support Unit, Dumfries and Galloway Royal infirmary, Bankend Road, Dumfries DG1 4AP.

Dr A Gordon Baird affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; any discrepancies from the study as planned have been explained.

Contributorship Statement

Both authors discussed the hypothesis that there was a disproportionate use of poor evidence in longer guidelines; both checked the raw data, and agreed on a statistical approach to discover whether the trend was significant or not. Both have been involved in writing and researching the evidence.

Funding.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing.

Technical appendix, statistical code, and dataset available from the corresponding author, who will provide a permanent, citable and open access home for the dataset.

Competing interests

No competing interests

MeSH Headings

- Practice Guideline N04.761.700.350.650
- General Practice H02.403.340
- Evidence-Based Medicine H02.249.750

How this fits in.

Guidelines should encourage an evidence-based approach to clinical practice. Longer guidelines used significantly higher levels of poor evidence. WHO has proposed a different system of grading evidence. The effect of group behavior altering guideline development has been hypothesised. New research often challenges established clinical practice. Improving the quality of evidence, acknowledging uncertainty and shortening guideline length would make guidelines more clinically relevant and effective.

REFERENCES

1. Harbour R, Lowe G, Twaddle S. Scottish Intercollegiate Guidelines Network; the first 15 years (1993-2008). *J R Coll Physicians Edinb* 2011 Jun;41(2) 163-168
2. de Joncheere K, Hill S, Klazinga N, et al. The Clinical Guideline Programme of the National Institute for Health and Clinical Excellence (NICE) A review by the World Health Organization May 2006
3. Lenzer J. Why we can't trust clinical guidelines: *BMJ* 2013;346:f3830
4. Aylett V. Do geriatricians need guidelines? *BMJ* 2010 Sep;341:c5340
5. Prasad V, Cifu A, Ioannidis JP. Reversals of Established Medical Practices: Evidence to Abandon Ship. *JAMA*. 2012 Jan4;307(1):37-38,
6. Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Int Med*. 2011 Oct 10;. 171(18): 1675-1676
7. SIGN 50: A guideline developer's handbook <http://www.sign.ac.uk/pdf/sign50.pdf>
8. Tricoci P, Allen JM, Kramer JM, et al. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009 Feb25; 301 (8): 831-841
9. Raine, R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. *BMJ*. 331(7517):631-633, September 17, 2005
10. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo - The Cardiac Arrhythmia Suppression Trial. *New Eng J Med*. 1991 Mar21; 324 (12): 781-788
11. Binanay C, Califf RM, Hasselblad V, et al. Evaluation study of congestive heart failure and pulmonary artery catheterization effectiveness: the ESCAPE trial. *JAMA*. 2005 Oct 5;294(13):1625-33.
12. Tatsioni A, Bonitsis NG, Ioannidis JA. Persistence of Contradicted Claims in the Literature. *JAMA*. 2007 Dec5 ;298(21):2517-2526.
13. Anderson JE, McKenzie C, Singh N, et al. Compliance with the 62 day target does not improve long-term survival. Association of Coloproctology of Great Britain and Ireland Annual Meeting 2012 Dublin Ireland.
14. Kerr J, Smith R, Gray S, et al. An audit of clinical practice in the management of head injured patients following the introduction of the Scottish Intercollegiate Guidelines Network (SIGN) recommendations. *Emerg Med J*. 2005 Dec; 22(12):850-854

- 1
2
3 15. Caplan LR. How well does "evidence-based" medicine help neurologists care for
4 individual patients? *Revin Neurol Dis.*2007;4(2):75-84
5
6
7 16.Hutchison G. Guidelines can harm patients too. *BMJ (Clinical research ed.)*. 2012
8 Apr18;344: e2685.
9
10 17. Woolf SH, Grol R, Hutchinson A, et al. Clinical guidelines. Potential benefits,
11 limitations, and harms of clinical guidelines. *BMJ* 1999 Feb20;318(7182):527-530
12
13 18. Miller L., Kidd A. Are sign guidelines key in the decision to CT? Post-head injury
14 CT scanning within a paediatric population. *Academic Emergency Medicine*.
15 Conference: 14th International Conference on Emergency Medicine, ICEM 2012
16 Dublin Ireland. Conference Publication: 19 (6) (pp 776), 2012. Date of Publication:
17 June 2012.
18
19
20 19. Thomson R., Eccles M., Wood R., et al. A cautionary note on data sources for
21 evidence-based clinical decisions: Warfarin and stroke prevention. *MedDecis*
22 *Making*. 2007;27 (4): 438-447
23
24 20. Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ*. 2011
25 Dec28;343:d7995
26
27
28 21. Cohen D. Rosiglitazone: what went wrong? *BMJ*.2010 Sep; 341:c4848
29
30
31 22. Gerstein HC, Miller ME, Byington RP, et al. Effects of intensive glucose lowering
32 in type 2 diabetes. Action to Control Cardiovascular Risk in Diabetes Study Group.
33 *New Engl J Med*. 2008 Jun12;358(24):2545-59
34
35 23. Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of
36 implementation of clinical guidelines' prescribing recommendations in primary care.
37 *Health Policy*.2008 Feb; 85 (2) 148-161
38
39
40 24.Guy S, Wardlaw JM. Who writes guidelines, and who should? *ClinRadiol*. 2002 Oct
41 ;57 (10): 891-897.
42
43 25. Greenhalgh, Trisha. *How To Read a Paper: The Basics of Evidence-Based*
44 *Medicine*. Wiley-Blackwell, fourth edition, 2010, p. 1
45
46 26. Zuiderent-Jerak T, Forland F, Macbeth F. Guidelines should reflect all knowledge,
47 not just trials *BMJ* 2012;345:e6702
48
49
50 27.GRADE working group. Grading quality of evidence and strength of
51 recommendations. *BMJ* 2004;328:1490
52
53
54 28. Harbour R, Miller J. A new system for grading recommendations in evidence
55 based guidelines. *BMJ*. 2001 Aug 11;323(7308):334-6
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8 **~~SIGN-Scottish Intercollegiate Guideline Network Guidelines—~~**
9 **analysis of evidence levels for their recommendations.**
10

11 *Alternatively*

12
13
14 **Guideline ~~s— is bigger recommendations are more~~ better? A**
15 **review of Scottish Intercollegiate Guideline Network SIGN**
16 **national guidelines.**
17
18
19
20

21 Dr A Gordon Baird, MB ChB MRCOG FRCGP R&D General
22 Practice Adviser, Research and Development Support Unit,
23 Dumfries and Galloway Royal infirmary, Bankend Road,
24 Dumfries, Scotland, UK DG1 4AP.
25
26
27

28 Address for correspondence; Dr A Gordon Baird, The White
29 House, Sandhead, Wigtownshire UK. DG9 9JA. Email
30 gordon.baird@me.com. Phone/fax 01776830281
31
32
33

34 Dr James R Lawrence, BSc MD FRCP(Glas & Edin) R&D Clinical
35 Lead, Research and Development Support Unit, Dumfries and
36 Galloway Royal infirmary, Bankend Road, Dumfries, Scotland,
37 UK DG1 4AP.
38
39
40

41 MeSH headings

42 General practice / Family practice

43 HEALTH SERVICES ADMINISTRATION & MANAGEMENT

44 Protocols & guidelines

45 Quality in health care

46 Clinical governance
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Background

National guidelines have become an integral part of UK general practice and aim to effectively deliver quality and consistency in clinical practice. As with any intervention, there are negative as well as positive consequences. Guideline effectiveness depends on the quality of evidence used.

Aim

To quantify and analyse the quality of evidence that is presented in national guidelines.

Design and setting

Levels of evidence used in all the current valid recommendations in the Scottish Intercollegiate Guideline Network (SIGN) guidelines were reviewed and statistically analysed.

Method

The data was collected from published guidelines available online to the public. SIGN methodology entails a professional group selected by a national organisation to develop each of these guidelines. Statistical analysis of the relationship between the number of guideline recommendations and the quality of evidence used in its recommendations was performed.

Result

A significant correlation between the number of recommendations in a guideline and the use of level D evidence was discovered.

Conclusion

Practice guidelines should be brief and based on scientific evidence. Paradoxically the longest guidelines have the highest proportion of recommendations based on the lowest level of evidence. Guideline developers should be more aware of the need for brevity and a stricter application of evidence-based principles could achieve this. The findings support calls for a review of how evidence is used and presented in guidelines.

Article Summary

Article focus

Examines the hypothesis that larger guidelines rely disproportionately on poor evidence.

Key Messages

- One third of current national guidelines are supported only by case reports, case studies and expert opinion.
- Guidelines with large numbers of recommendations used a higher proportion of weak evidence.
- Guideline development groups appear to vary in their approach to offering recommendations.
- Guideline recommendations should be based on good evidence.
- Paucity of evidence should highlight topics for research.

Strengths and limitations of the study

Strengths.

- This is the first objective evidence of inconsistencies in approach by a national guideline developers
- This supports commentator suggestion that even without good evidence a group will prefer consensus.
- Adds to the current debate about how guidelines might be developed in the future

Limitations.

- The study is limited to only one set of national guidelines (SIGN).
- Reasons for the differences in quality of evidence preferred by the guideline development groups is unclear

Guidelines – is bigger better? A review of SIGN guidelines.

INTRODUCTION

The Scottish Intercollegiate Guidelines Network (SIGN) was founded in 1993. It is a national body, professionally led and publicly funded. SIGN's founding principles proposed direct links between evidence and recommendations, offering a brief and succinct quick-reference guide for clinicians [1]. Guidelines anticipated presenting brief, evidence based clinical advice. They have developed into long and authoritative texts used by managers and politicians to inform policy. A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003. ~~SIGN has~~Both have responsibility to consider cost-effectiveness and ~~directly~~ inputs to the Quality Outcomes Framework (QOF). ~~A formal arrangement between SIGN and the National Institute of Clinical Excellence (NICE) has existed from 2003.~~

The World Health Organisation (WHO) recognises that current grades of recommendation (Table 1) may be ambiguous [2] and encourages guideline developers to use a system which includes a category "Use only in the context of research" where doubt exists.

Guideline developers have conflict of interest policies reported as challenging to apply. Where doubt exists, groups of specialists may feel consensus more defensible than acknowledging uncertainty. [3]

Even with the best evidence, concerns are expressed about the relevance of guidelines in treating patients with multiple morbidities [4], and the emergence of the phenomenon of reversal [5,6], where established practice, sometimes evidence based, is shown to be sub-optimal or harmful. This study looks at the quality of evidence used for SIGN guidelines, and describes a significant trend for some groups to emphasise poorly evidence-backed recommendations.

METHODS

SIGN guidelines were accessed online in September 2013. SIGN guidelines were chosen because they are internationally respected, the authors were familiar with their format and they contribute to national government policy. Guidelines that were "Current" and "Current 3-7 years". Some recommendations may be out of date." were included. Those that had been "Withdrawn", "Recommendations being updated", "Need for update being considered" and those with no recommendations were excluded.

SIGN guideline 50 clearly describes an established process for developing guidelines.^[7] It explains how the process is planned, how it is implemented and by whom. This process is independent of this study, but is stated to be an objective process. SIGN guidelines have four grades of recommendation outlined in table 1. Table 2 describes the level of evidence SIGN uses to support the recommendation grading. SIGN guideline development groups vary in size depending on the scope of the topic under consideration, but generally comprise between 15 and 25 members. SIGN states they are aware of the many psychosocial factors, including the problems of overcoming professional hierarchies that can affect small group processes.

3 investigators (JRL, AGB, ABB) independently enumerated the level of evidence used by each examined guideline ~~was independently enumerated by 3 investigators, and discrepancies resolved.~~ They discounted any duplication implicit in text-embedded key recommendations and also implementation recommendations. There were no discrepancies. A statistical analysis of the correlation between the proportion of level D evidence and the total number of recommendations was performed for the 42 guidelines.

RESULTS

The 42 guidelines consisted of 2559 pages (including references), ranging from 26 to 161 (median 59.5) pages. The longest guideline, number 116 was 61 pages longer than the next largest. The number of recommendations per page ranged from 0.2 to

1
2
3
4
5
6 1.8 (median 0.7). The number of recommendations per guideline is presented in
7 table 3.
8
9

10
11 Of the 1999 recommendations, 480 (24.0%) were level A, 491(24.6 %) were level B,
12 318 (15.9%) level C, and 710 (35.5%) level D. Thus 51.4% were poorly evidenced
13 (C&D) and over a third (D) developed almost entirely on “expert opinion”. The
14 number of level A recommendations per guideline ranged from 0-57 (median 9),
15 level B from 2-62 (median 8.5) level C ranged from 0-26 (median 6) and D from 0-60
16 (median 14.5). 4 guidelines had no level A evidence.
17
18
19

20
21 The proportion of level D evidence increases with the number of recommendations
22 made. This correlation is significant with Kendall’s Tau=0.22 [approximate 95%
23 confidence interval 0.008-0.45] p value =0.04, and Spearman rho=0.22 [approximate
24 95% CI 0.02-0.57] p value= .04.
25
26
27

28 29 30 DISCUSSION

31 This study reveals that expert groups who produce long guidelines rely on poor
32 evidence more heavily than others. While this study only looks at SIGN, this study
33 highlights a problem that has escaped national guideline developers, a wide range of
34 professionals and the public to whom these guidelines are applied. National
35 guidelines are useful and important and there is a debate about how evidence is
36 best presented. Guidelines define standards of care, help busy clinicians and allow
37 managers and politicians to develop governance. An American study (using 3 not 4
38 levels of evidence) similarly found that 48% were “based on expert opinion, case
39 studies, or standards of care.” [8]; we show comparable results for current SIGN
40 guidelines. Where patients are involved in clinical decisions, honestly declaring
41 uncertainty has merit. In the absence of good scientific evidence, recommending a
42 course of action without understanding the circumstances of the individual to whom
43 it is applied seems both risky and, assuming a right to patient choice, unwarranted.
44 Other guidelines that use high levels of poor evidence should evaluate the
45 proportion of poorly evidenced recommendations and seek explanations for such
46 trends.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8 This study did not examine why longer guidelines use poorer evidence. ~~Cloistered~~
9 ~~G~~groups of experts, indulging in “group think” may view their own opinion as more
10 authoritative than science can support^[9]. It has been postulated that there is
11 security in “just doing what everyone else is doing – even if what everyone else is
12 doing isn’t very good.”^[3] Reliance on expert opinion has a poor track record.
13 Blinded by certainty, expert groups defining established practice have, in the past,
14 perpetuated radical mastectomy instead of conservative surgery, Class 1C
15 antiarrhythmics^[10], pulmonary artery catheters in heart failure^[11], electronic foetal
16 monitoring in low risk pregnancies: even then practice can take a decade to
17 reverse^[12].

Comment [GB1]: New reference

24
25 Even good evidence is subject to the phenomenon of reversal where new evidence
26 contradicts current practice. Reversal can affect around 13-16% of publications ^[5,6].
27 This may partly explain why the implementation of even the most soundly evidence
28 based national guidelines fails to improve outcome ^[13-15]. There is potential
29 harm^[16,17] from guidelines in real clinical settings, for example increasing radiation
30 dose without benefit^[18] or increased risks of anticoagulation^[19].

31
32
33
34
35
36 SIGN 116 (diabetes), is a notable exception. It is more than 50% larger than the next
37 largest, 2.5 times longer than the average and yet uses the 4th lowest level D
38 recommendations. There are a number of hypotheses why this group reports
39 differently. SIGN guidelines inform Quality Outcomes Framework (QOF) policy.
40 Diabetes is the largest clinical UK QOF indicator and is associated with substantial
41 payment incentives. The need for objective evaluation of performance drives a use
42 of surrogate outcomes without appropriate clinical endpoints. ^[20] Diabetes
43 guidelines have suffered several noteworthy reversals. Examples include the
44 recommendation of glycosylated haemoglobin reduction resulting in increased use
45 of rosiglitazone (still mentioned in the current document) both associated with harm
46 including mortality. ^[21,22] Aspirin recommendations have also been changed from
47 previous guidelines. Is it possible that the repeated use of surrogate outcomes arises
48 from group dynamics driven by a powerful external agenda?
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8 | Many doctors whose expertise cross several guidelines^[23,24] express concerns
9 about guideline development groups. The inappropriate exclusion of disease groups
10 from general population data is common. Smoking cessation advice for testicular
11 cancer survivors is level C, although studies in the general population (without
12 excepting specific disease groups) advises everyone to stop smoking. Overall
13 smoking cessation was level D and C once each and B on three occasions. Using
14 evidence in this way may imply group dysfunction. Differently constituted groups, or
15 greater oversight might avoid problems.
16
17
18
19
20

21
22 In 1993, SIGN guidelines stated intention was to be evidence based, brief and
23 succinct. Brevity increases value as a quick reference guide. Removing or reducing
24 poorly evidenced recommendations would reduce size by more than a third overall
25 and in some up to two thirds. The two volumes Oxford Textbook of Primary Medical
26 Care (2005) is a relatively brief 1420 pages, more than a thousand less than the 2559
27 pages of guidelines. Evidence based medicine is described as “the use of
28 mathematical estimates of the risk of benefit and harm, derived from high-quality
29 research on population samples, to inform clinical decision-making in the diagnosis,
30 investigation or management of individual patients”^[25]. Guidelines relevance to
31 daily practice, the reliability of evidence and whether the application of evidence will
32 improve outcomes are important questions.
33
34
35
36
37
38
39

40
41 These results may reflect how professional groups deal with uncertainty. If so, this is
42 not good for individual patients faced with the same uncertainties (whether aware
43 of it or not), nor is it good for scientists who actively seek unanswered questions by
44 challenging established practice, an area in which medicine has a poor record from
45 Semmelweis to the present day.
46
47
48
49

50
51 The finding of a significant increase of level D recommendations in larger guidelines
52 has not happened by chance. A wider debate about how guideline groups can create
53 greater clarity about the reliability of evidence used is needed.^{[26[†]]}—Reducing the
54 use of poorly evidenced recommendations has potential to create a shorter, more
55
56
57
58
59
60

reliable and usable clinical support. The GRADE working group was formed in 2000.^{[27]²} SIGN ~~proposed a move~~ moved to a new grading system in 2001^{[28]³} and ~~from 2013 a new system based on GRADE principles.~~ Whether these changes ~~in process at present being considered~~ will resolve the challenges that underpin the inconsistencies we have outlined remains to be seen.

The authors would like to acknowledge the help of Heather Barrington, Dumfries & Galloway Royal Infirmary, Research & Development Support Unit Dumfries, Dumfries & Galloway, United Kingdom who gave statistical advice and support, and Anne B Baird, The white House, Sandhead, Wigtownshire DG9 9JA who checked the data for inconsistencies.

Table 1

GRADES OF RECOMMENDATION

A	At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

Table 2

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	-Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

2+	High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

Table 3
Main Recommendations

Number	Name	Pages	A	B	C	D	Total	%age D
133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	Management of diabetes	161	57	62	23	16	158	10.1%
115	Management of Obesity	87	6	11	7	11	35	31.4%
114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%
113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	Management of stable angina	59	13	10	3	11	37	29.7%
95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%
87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
		TOTAL	2559	480	491	318	710	1999

Acknowledgements.

1
2
3
4
5
6
7 Heather Barrington, Statistical Adviser; Bridget Bird, Administrative Assistant;
8 Research and Development Support Unit, Dumfries and Galloway Royal infirmary,
9 Bankend Road, Dumfries DG1 4AP.

10 Dr A Gordon Baird affirms that this manuscript is an honest, accurate, and
11 transparent account of the study being reported; that no important aspects of the
12 study have been omitted; any discrepancies from the study as planned have been
13 explained.
14

15 16 17 **Contributorship Statement**

18 **Both authors discussed the hypothesis that there was a disproportionate use of**
19 **poor evidence in longer guidelines; both checked the raw data, and agreed on a**
20 **statistical approach to discover whether the trend was significant or not. Both**
21 **have been involved in writing and researching the evidence.**
22

23 24 **Funding.**

25 This research received no specific grant from any funding agency in the public,
26 commercial or not-for-profit sectors.
27

28 29 **Data Sharing.**

30 Technical appendix, statistical code, and dataset available from the corresponding
31 author, who will provide a permanent, citable and open access home for the dataset.
32

33 34 **Competing interests**

35 No competing interests
36

37 **MeSH Headings**

- 38 • Practice Guideline N04.761.700.350.650
- 39 • General Practice H02.403.340
- 40 • Evidence-Based Medicine H02.249.750

41 42 **How this fits in.**

43 Guidelines should encourage an evidence-based approach to clinical practice. Longer
44 guidelines used significantly higher levels of poor evidence. WHO has proposed a
45 different system of grading evidence. The effect of group behavior altering guideline
46 development has been hypothesised. New research often challenges established
47 clinical practice. Improving the quality of evidence, acknowledging uncertainty and
48 shortening guideline length would make guidelines more clinically relevant and
49 effective.
50

51 **REFERENCES**

52
53 **[1. Harbour R, Lowe G, Twaddle S. Scottish Intercollegiate Guidelines Network; the](#)**
54 **[first 15 years \(1993-2008\). J R Coll Physicians Edinb 2011 Jun;41\(2\) 163-168](#)**
55
56
57
58
59
60

- 1
2
3
4
5
6 [2. de Joncheere K, Hill S, Klazinga N, Mäkelä M, Oxman AD. The Clinical Guideline Programme of the National Institute for Health and Clinical Excellence \(NICE\) A review by the World Health Organization May 2006](#)
- 7
8
9
10 [3. Lenzer J. Why we can't trust clinical guidelines: BMJ 2013;346:f3830](#)
- 11
12 [4. Aylett V. Do geriatricians need guidelines? BMJ 2010 Sep;341:c5340](#)
- 13
14
15 [5. Prasad V, Cifu, A, Ioannidis JP. Reversals of Established Medical Practices: Evidence to Abandon Ship. JAMA. 2012 Jan4;307\(1\):37-38,](#)
- 16
17
18 [6. Prasad V, Gall V, Cifu A. The frequency of medical reversal. Arch Int Med. 2011 Oct 10; 171\(18\): 1675-1676](#)
- 19
20
21 [7. SIGN 50: A guideline developer's handbook <http://www.sign.ac.uk/pdf/sign50.pdf>](#)
- 22
23 [8. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith Jr. SC. Scientific evidence underlying the ACC/AHA clinical practice guidelines. JAMA 2009 Feb25; 301 \(8\): 831-841](#)
- 24
25
26
27 [9. Raine, R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. BMJ. 331\(7517\):631-633, September 17, 2005](#)
- 28
29
30
31 [10. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, Arensberg D, Baker A, Friedman L, Greene HL, Huther ML, Richardson DW. Mortality and morbidity in patients receiving encainide, flecainide, or placebo - The Cardiac Arrhythmia Suppression Trial. New Eng J Med. 1991 Mar21; 324 \(12\): 781-788](#)
- 32
33
34
35 [11. Binanay C, Califf RM, Hasselblad V, O'Connor CM, Shah MR, Sopko G, Stevenson LW, Francis GS, Leier CV, Miller LW Evaluation study of congestive heart failure and pulmonary artery catheterization effectiveness: the ESCAPE trial. JAMA. 2005 Oct 5;294\(13\):1625-33.](#)
- 36
37
38
39
40 [12. Tatsioni A, Bonitsis NG, Ioannidis JA. Persistence of Contradicted Claims in the Literature. JAMA. 2007 Dec5 ;298\(21\):2517-2526.](#)
- 41
42
43
44 [13. Anderson JE, McKenzie C, Singh N, Gajree N, Giles L, Sharma P, McDonald A. Compliance with the 62 day target does not improve long-term survival. Association of Coloproctology of Great Britain and Ireland Annual Meeting 2012 Dublin Ireland.](#)
- 45
46
47
48 [14. Kerr J, Smith R, Gray S, Beard D, Robertson CE. An audit of clinical practice in the management of head injured patients following the introduction of the Scottish Intercollegiate Guidelines Network \(SIGN\) recommendations. Emerg Med J. 2005 Dec; 22\(12\):850-854](#)
- 49
50
51
52
53 [15. Caplan LR. How well does "evidence-based" medicine help neurologists care for individual patients? Rev Neurol Dis. 2007;4\(2\):75-84](#)
- 54
55
56
57
58
59
60

16. [Hutchison G. Guidelines can harm patients too. BMJ \(Clinical research ed.\). 2012 Apr18;344: e2685.](#)
17. [Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines. Potential benefits, limitations, and harms of clinical guidelines. BMJ 1999 Feb20;318\(7182\):527-530](#)
18. [Miller L., Kidd A. Are sign guidelines key in the decision to CT? Post-head injury CT scanning within a paediatric population. Academic Emergency Medicine. Conference: 14th International Conference on Emergency Medicine, ICEM 2012 Dublin Ireland. Conference Publication: 19 \(6\) \(pp 776\), 2012. Date of Publication: June 2012.](#)
19. [Thomson R., Eccles M., Wood R., Chinn D.J. A cautionary note on data sources for evidence-based clinical decisions: Warfarin and stroke prevention. MedDecis Making. 2007;27 \(4\): 438-447](#)
20. [Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. BMJ. 2011 Dec28;343:d7995](#)
21. [Cohen D. Rosiglitazone: what went wrong? BMJ.2010 Sep; 341:c4848](#)
22. [Gerstein HC, Miller ME, Byington RP, Goff DC Jr., Bigger JT, Buse JB, Cushman WC, Genuth S, Ismail-Beigi F, Grimm RH Jr. Probstfield JL, Simons-Morton DG, Friedewald WT. Effects of intensive glucose lowering in type 2 diabetes. Action to Control Cardiovascular Risk in Diabetes Study Group. New Engl J Med. 2008 Jun12;358\(24\):2545-59](#)
23. [Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of implementation of clinical guidelines' prescribing recommendations in primary care. Health Policy.2008 Feb; 85 \(2\) 148-161](#)
24. [Guy S, Wardlaw JM. Who writes guidelines, and who should? ClinRadiol. 2002 Oct ;57 \(10\): 891-897.](#)
25. [Greenhalgh, Trisha. How To Read a Paper: The Basics of Evidence-Based Medicine. Wiley-Blackwell, fourth edition, 2010, p. 1](#)
26. [Zuiderent-Jerak T, Forland F, Macbeth F. Guidelines should reflect all knowledge, not just trials BMJ 2012;345:e6702](#)
27. [GRADE working group. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490](#)
28. [Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. BMJ. 2001 Aug 11;323\(7308\):334-6](#)

1
2
3
4
5
6
7
8
9
10
11 [1-Zuiderent Jerak T, Forland F, Macbeth F. Guidelines should reflect all knowledge,
12 not just trials—BMJ 2012;345:e6702](#)

13 [2-GRADE working group http://www.gradeworkinggroup.org/about-us.htm](http://www.gradeworkinggroup.org/about-us.htm)

14
15
16 [3-Harbour R, Miller J. A new system for grading recommendations in evidence-based
17 guidelines. BMJ. 2001 Aug 11;323\(7308\):334-6.](#)

TABLE 1

Grades of recommendation

A	At least one meta-analysis ,systematic review, or RCT rated as 1++and directly applicable to the target population; or a body of evidence consisting principally of studies rated as 1+,directly applicable to the target population, and demonstrating overall consistency of results.
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+.
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2+.
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2=.

Table 2

Levels of evidence

1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias
1 -	Meta-analyses, systematic reviews, or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies. High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2+	Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
2 -	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, eg case reports, case series
4	Expert opinion

Table 3 – for editing

Number	Number	Name	Pages	Main Recommendations				Total	%age D
				A	B	C	D		
133	133	Management of hepatitis C	57	20	24	7	52	103	50.5%
132	132	Long-term follow-up of survivors of childhood cancer	62	0	7	9	14	30	46.7%
131	131	Management of schizophrenia	64	10	19	3	15	47	31.9%
130	130	Brain injury rehabilitation in adults	68	0	14	7	8	29	27.6%
129	129	Antithrombotics: indication and management	68	25	11	6	19	61	31.1%
127	127	Management of perinatal mood disorders	47	0	5	6	15	26	57.7%
126	126	Diagnosis and management of colorectal cancer	56	11	19	15	29	74	39.2%
125	125	Management of atopic eczema in primary care	34	3	5	3	2	13	15.4%
124	124	Management of adult testicular germ cell tumours	63	6	6	9	21	42	50.0%
123	123	Management of early rheumatoid arthritis	27	3	7	2	0	12	0.0%
122	122	Prevention and management of venous thromboembolism	88	26	15	14	55	110	50.0%
121	121	Diagnosis and management of psoriasis and psoriatic arthritis in adults	65	11	16	6	26	59	44.1%
120	120	Management of chronic venous leg ulcers	46	5	3	4	7	19	36.8%
119	119	Management of patients with stroke: identification and management of dysphagia	42	0	6	4	20	30	66.7%
118	118	Management of patients with stroke: rehabilitation, prevention and management of complications, and discharge planning	101	21	29	7	21	78	26.9%
117	117	Management of sore throat and indications for tonsillectomy	37	9	3	4	4	20	20.0%
116	116	Management of diabetes	161	57	62	23	16	158	10.1%
115	115	Management of Obesity	87	6	11	7	11	35	31.4%
114	114	Nonpharmaceutical management of depression	37	5	4	0	0	9	0.0%

Table 3 – for editing

113	113	Diagnosis and pharmacological management of Parkinson's disease	61	12	6	6	4	28	14.3%
112	112	Management of attention deficit and hyperkinetic disorders in children and young people	45	6	4	3	4	17	23.5%
111	111	Management of hip fracture in old people	49	10	9	8	14	41	34.1%
110	110	Early management of patients with a head injury	76	1	7	6	17	31	54.8%
109	109	Management of genital Chlamydia trachomatis infection	40	3	6	9	29	47	61.7%
108	108	Management of patients with stroke or TIA: assessment, investigation, immediate management and secondary prevention	100	42	27	18	14	101	13.9%
107	107	Diagnosis and management of headache in adults	81	17	16	9	34	76	44.7%
106	106	Control of pain in adults with cancer	71	5	7	3	19	34	55.9%
105	105	Management of acute upper and lower gastrointestinal bleeding	57	14	5	2	15	36	41.7%
103	103	Diagnosis and management of chronic kidney disease	50	9	6	4	3	22	13.6%
102	102	Management of invasive meningococcal disease in children and young people	46	1	4	6	26	37	70.3%
99	99	Management of cervical cancer	73	1	13	19	29	62	46.8%
97	97	Risk estimation and the prevention of cardiovascular disease	72	16	12	2	4	34	11.8%
96	96	Management of stable angina	59	13	10	3	11	37	29.7%
95	95	Management of chronic heart failure	55	9	12	1	1	23	4.3%
94	94	Cardiac arrhythmias and coronary heart disease	42	22	11	13	23	69	33.3%
93	93	Acute coronary syndromes	60	11	14	9	8	42	19.0%
91	91	Bronchiolitis in children	42	4	3	6	14	27	51.9%
90	90	Diagnosis and management of head and neck cancer	92	42	8	26	60	136	44.1%
89	89	Diagnosis and management of peripheral arterial disease	37	11	2	0	4	17	23.5%
88	88	Management of suspected bacterial urinary tract infection in adults	45	8	10	2	10	30	33.3%

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Table 3 – for editing

87	87	Management of oesophageal and gastric cancer	70	3	26	23	28	80	35.0%
61	61	Investigation of postmenopausal bleeding	26	2	7	4	4	17	23.5%
<u>TOTAL</u>			<u>2559</u>	<u>480</u>	<u>491</u>	<u>318</u>	<u>710</u>	<u>1999</u>	

For peer review only