White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/9019/

Article:
Muth, C., Gensichen, J., Beyer, M. et al. (2009) The systematic guideline review: method, rationale, and test on chronic heart failure. BMC Health Services Research, 9. Art. No.74. ISSN: 1472-6963

https://doi.org/10.1186/1472-6963-9-74

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

BMC Health Services Research

Research article

Open Access

The Systematic Guideline Review: Method, rationale, and test on chronic heart failure

Christiane Muth*[1], Jochen Gensichen[2], Martin Beyer[1], Allen Hutchinson[3] and Ferdinand M Gerlach[1]

Address: [1]Institute for General Practice, Johann Wolfgang Goethe University, Frankfurt/Main, Germany, [2]Institute for General Practice, University Hospital, Friedrich-Schiller-University, Jena, Germany and [3]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK

Email: Christiane Muth* - muth@allgemeinmedizin.uni-frankfurt.de; Jochen Gensichen - gensichen@allgemeinmedizin.uni-frankfurt.de; Martin Beyer - beyer@allgemeinmedizin.uni-frankfurt.de; Allen Hutchinson - Allen.Hutchinson@sheffield.ac.uk; Ferdinand M Gerlach - gerlach@allgemeinmedizin.uni-frankfurt.de

* Corresponding author

Published: 8 May 2009

BMC Health Services Research 2009, **9**:74 doi:10.1186/1472-6963-9-74

Received: 3 November 2008
Accepted: 8 May 2009

This article is available from: http://www.biomedcentral.com/1472-6963/9/74

© 2009 Muth et al; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Evidence-based guidelines have the potential to improve healthcare. However, their de-novo-development requires substantial resources – especially for complex conditions, and adaptation may be biased by contextually influenced recommendations in source guidelines. In this paper we describe a new approach to guideline development – the systematic guideline review method (SGR), and its application in the development of an evidence-based guideline for family physicians on chronic heart failure (CHF).

Methods: A systematic search for guidelines was carried out. Evidence-based guidelines on CHF management in adults in ambulatory care published in English or German between the years 2000 and 2004 were included. Guidelines on acute or right heart failure were excluded. Eligibility was assessed by two reviewers, methodological quality of selected guidelines was appraised using the AGREE instrument, and a framework of relevant clinical questions for diagnostics and treatment was derived. Data were extracted into evidence tables, systematically compared by means of a consistency analysis and synthesized in a preliminary draft. Most relevant primary sources were re-assessed to verify the cited evidence. Evidence and recommendations were summarized in a draft guideline.

Results: Of 16 included guidelines five were of good quality. A total of 35 recommendations were systematically compared: 25/35 were consistent, 9/35 inconsistent, and 1/35 un-rateable (derived from a single guideline). Of the 25 consistencies, 14 were based on consensus, seven on evidence and four differed in grading. Major inconsistencies were found in 3/9 of the inconsistent recommendations. We re-evaluated the evidence for 17 recommendations (evidence-based, differing evidence levels and minor inconsistencies) – the majority was congruent. Incongruity was found where the stated evidence could not be verified in the cited primary sources, or where the evaluation in the source guidelines focused on treatment benefits and underestimated the risks. The draft guideline was completed in 8.5 man-months. The main limitation to this study was the lack of a second reviewer.

Conclusion: The systematic guideline review including framework development, consistency analysis and validation is an effective, valid, and resource saving-approach to the development of evidence-based guidelines.

Background

Evidence-based clinical practice guidelines have considerable potential to improve health care, and their international production has increased substantially over the past two decades [1-3]. However, the de-novo development of an evidence-based guideline requires considerable time, expertise and resources [4,5], the latter of which are limited, even in developed countries [6,7]. It was therefore suggested that their development should be based on the adaptation of existing guidelines [8].

A number of projects have sought to adapt or adopt existing guidelines [9-14], on topics ranging from HIV/AIDS [10] to the management of acute low back pain [14]. Each, in its own way, has highlighted the challenges of deriving new guidelines from existing material. Deficient development methods [15-20], subjectivity [21] and conflicts of interest [22,23] in the source guidelines may all lead to biases.

Additionally, implicit normativity has been shown to be inherent in guidelines – from the search for and critical appraisal of evidence, to its presentation, and the formulation of recommendations [24]. Contextual features arise from subject and financial constraints [25,26], ethical considerations and social influences [27-29], as well as practical necessities [30]. But presuppositions and values change, vary between cultures and healthcare systems, and are sometimes inconsistent, even within healthcare systems [30-34]. They may distort an adapted guideline to suit a different target context. Moreover, explaining and addressing context-specific normative considerations have been seen as a necessary condition for the successful implementation of guidelines [27].

All of these issues give rise to methodological challenges for adaptation processes. In recognition of this, the ADAPTE group [35] has proposed a seven-step framework which they call 'trans-contextual' adaptation. Fervers et al. [35] pointed out that it is of crucial importance to analyze the coherence between evidence and recommendations and to take culture and systems into account, but they did not provide the practical means to deal with this problem. In this paper we describe a new method that we have named the systematic guideline review (SGR). The SGR method was designed to take both methodological shortcomings and context-specific normative issues in source guidelines into consideration, in order to develop a valid guideline for a different target context in a resource saving manner.

The SGR was conducted as the first step in the development of a new guideline on chronic heart failure for use by family physicians in Germany. After having carried out the SGR, the resulting draft guideline underwent further development in accordance with institutional standards [36], which are out of scope of this paper: To ensure the involvement of stakeholders and target users, this comprised a multistage internal and external peer review, a multi-professional, interdisciplinary formal consensus that included a patients' representative (nominal group process), and a pilot testing phase. The final guideline was authorized by the German Society of General Practice and Family Medicine (DEGAM) and published in 2006 [37]. In our article we focus on the systematic guideline review.

Methods

To develop an evidence-based guideline on chronic heart failure (CHF) for use in German primary care we defined, a priori, the key premises of the target guideline: the definition of the target condition, its epidemiology, the target setting, the need for change in healthcare, and the outcome parameters of interest.

We designed the SGR method comprising the nine steps grey tagged in FIGURE 1.

Systematic guideline search

In March 2004 one of us (CM) performed a systematic literature search for existing guidelines in MEDLINE, The Cochrane Library, DARE, and HSTAT; combining controlled terms and free text, complemented by a comprehensive hand search of web-based resources (see Additional file 1, Tables W1 and W2) and in reference lists of the retrieved guidelines.

Guideline Selection

Based on predefined criteria, two reviewers (CM, JG) independently appraised the retrieved guideline-documents for eligibility (yes or no). Discrepancies were resolved by consensus, and kappa-statistics were calculated for observer variability according to Cohen [38]. To be included a guideline had to be (1) dedicated to adult patients with chronic heart failure (CHF), (2) evidence-based, i.e. evidence levels (and/or a grading) were reported in the majority of recommendations with a clear link to supporting evidence, (3) released in the year 2000 or later, and (4) published in English or German.

A guideline was excluded, if it referred exclusively to pediatric patients, isolated right heart failure, special or inpatient care, or no information on development methods was given (at least, an institutional standard of the developing organization had to be provided). Guidelines with a focus only on particular aspects of heart failure management (e.g. guidelines on echocardiography) were excluded from the SGR.

Quality appraisal of guidelines

The quality of the guideline was assessed using the standardized Appraisal of Guidelines Research and Evaluation (AGREE) instrument [39,40]. AGREE consists of 23 items

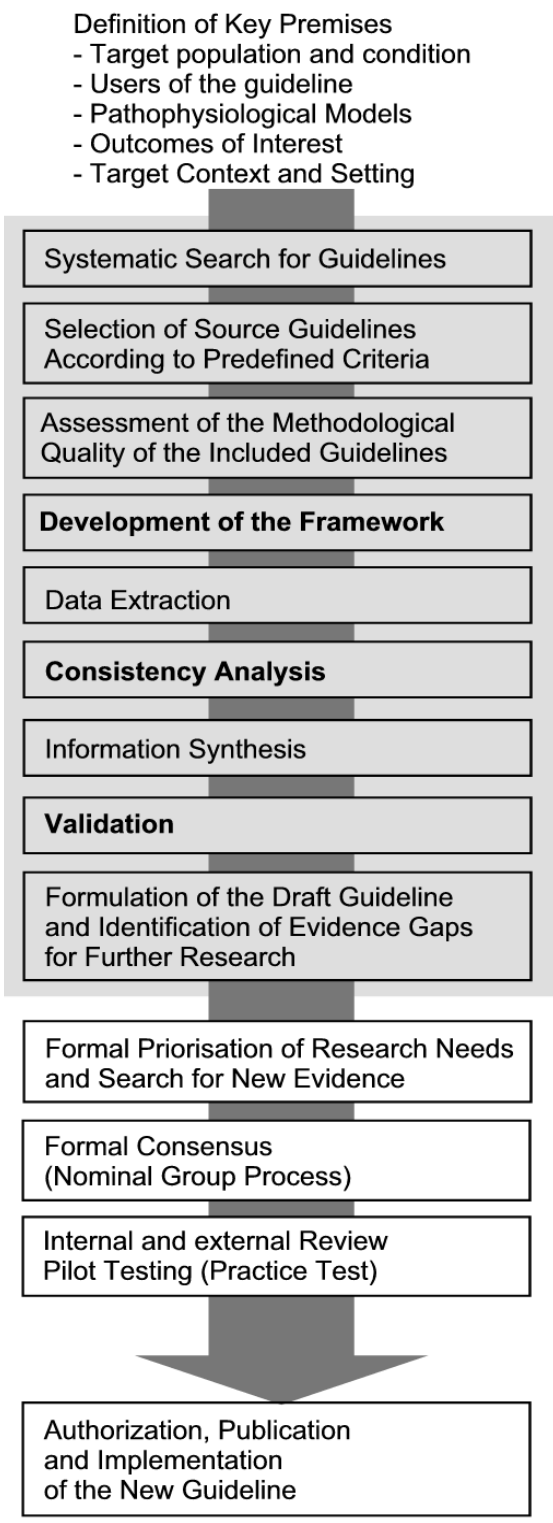

Figure 1
Development of the Evidence-based Guideline on Chronic Heart Failure in Primary Care. Gray Tagged: The Systematic Guideline Review.

which rate the various dimensions of guideline quality by means of four-point Likert scales. The items are organized into six independent domains. Domain scores are calculated by summing up the scores of the items per domain and are standardised as a percentage of the maximum possible score for that domain. [39]. There is no universal agreement on specific cut off scores to identify high quality guidelines, and differing approaches are applied [41,42]. To rate the overall guideline quality we defined a median target and result score per guideline (calculated as the median of the standardized domain scores over all domains) of 0.5 or above as indicating good quality.

AGREE is suggested for multiple raters. The appraisal is based on the consideration of the whole guideline and is therefore of wider scope than the evidence review and synthesis reported here. Nevertheless, it does include any recommendations on diagnosis, monitoring, or prevention and the proposed multiple rater scheme is thus rather time-consuming. Because we had only limited resources available to us, we performed a single rating (CM) for this step of the guideline development process.

Framework and data extraction

To build a framework of relevant issues, clinical questions from the included guidelines were derived by one reviewer (CM), appraised for their relevance and prioritized in terms of the target context by three others (JG, MB, FMG). For each clinical question the data were extractedinto evidence tables of a standardized format which included recommendation(s), evidence level(s), grading, critical appraisal of evidence, and cited sources.

Consistency analysis

One reviewer (CM) systematically compared guideline recommendations for every clinical question for their external consistency (e.g. between guidelines), and their reported evidence. We categorized our findings into six types (Table 1): four types of consistency and two types of inconsistency (major and minor). The categories were built on both dimensions (external consistency and evidence base) to support a further action plan for guideline development: low levels of evidence as well as inconsistencies between guidelines indicated the need for further research. The classification system was drafted by two of us (CM and MB) and consented with the others (JG, AH, FMG). One reviewer (CM) synthesized the information (recommendations, their consistency, as well as the reported underlying evidence from the included guidelines) in a preliminary draft.

Validation

For a critical re-appraisal of whether recommendations were supported by valid study results, the most relevant evidence cited in the included guidelines was selected by

Table 1: Categories of Consistency and Inconsistency

Type	Definition	Further Action Plan for Guideline Development
Consistency		
(1)	**Recommendations are consistent** in content, evidence level and grading, and **based on a *large* body of high level evidence** (e.g. multiple primary or secondary studies of high internal and external validity)	Verification of cited sources with highest evidence level, and update searches, if necessary
(2)	**Recommendations are consistent** in content, evidence level and grading, and **based on a *small* body of high level evidence** (e.g. a single or a few primary studies of high internal and external validity)	Verification of cited sources, further research on safety aspects in particular, and update searches
(3)	**Recommendations are consistent** in content, evidence level and grading, and **based on** evidence from studies of **low level evidence** (e.g. studies with design-related biases or where methodological flaws reduce internal or external validity) **or based on expert opinion** (where evidence is lacking)	Further research on evidence
(4)	**Recommendations are consistent in content**, but **evidence levels and grading conflict**	Verification of cited sources, and update searches
Inconsistency		
(A)	**Recommendations are completely inconsistent**, neither a mainstream trend nor even a common denominator can be identified	Further research on evidence
(B)	**Recommendations are consistent in the majority** of guidelines, but **differing** or even conflicting **recommendations are to be found in a minority**	Verification of cited sources to decide whether further research is necessary, and update-searches

applying the highest levels of evidence from the Oxford Centre for Evidence-Based Medicine [43]. Predominantly systematic reviews with or without meta-analyses were re-evaluated, along with clinical studies of an appropriate design when secondary publications did not provide the desired evidence. The selected studies were re-evaluated for their internal and external validity. For the assessment, checklists from the German Working Group on Health Technology Assessment (see Additional file 1, non-scoring standardized instruments, Table W3) were used, since one of us (CM) was trained and experienced in its use [44]. In particular, the studies were investigated for inherent bias, which might have influenced results, and for the applicability of their results (in accordance with the defined key premises of the target guideline: study population representative of target population, clinically relevant outcomes reported in the study, setting of the study appropriate with regard to the target setting).

Draft guideline

Finally, the evidence and recommendations were synthesized in a draft of the target guideline, and the results and possible methodological flaws of the included source-guidelines, and the re-assessed studies, were discussed. For specified clinical questions without identified high-quality evidence and for the necessary update-searches to ensure the actuality of the guideline, a structured list was formed to define the need for further research.

Results

Literature Search and Selection

Our literature search resulted in 699 citations, of which 52 were potentially relevant. A total of 16 evidence-based guidelines met our inclusion criteria (FIGURE 2). The inter-observer variability was excellent (one discrepancy; $\kappa = 0.95$). Of the included guidelines (Table 2), five were from Germany [45-49], four from the U.S. [50-54], two each from Canada [55-57] Australia & New Zealand [58,59], and one each from Finland [60], the U.K. [61], and Europe [62]. The context and setting of the guidelines varied: five guidelines were prepared for a specific health-care setting (e.g. ambulatory care in a health maintenance organization) [48,51-54,57], one guideline was directed towards CHF management throughout Europe [62], one guideline did not specify any target group [49], and the remaining guidelines referred to national healthcare. The scope of the guidelines varied considerably (all addressed pharmacotherapy in heart failure, 13/16 non-pharmaco-

Table 2: Characteristics of Included Guidelines

Source	Organisation, Country of Origin	Covered Scope†	No. of Authors	Literature Search: Period/Databases	No. of References
ACC/AHA 2001 [50]	American College of Cardiology/American Heart Association, U.S.A.	D, T	14	Period not specified MEDLINE, EMBASE	573
AKDAE 2001 [45]	Drug Commission of the German Medical Association, Germany	P	Not specified	No systematic search	216
CCS 2001 [56], 2002/3 [55]	Canadian Cardiovascular Society, Canada	D, T	Not specified	Period not specified MEDLINE	2001: 79; 2002/3: 42
DGK 2001 [46]	German Cardiac Society, Germany	T	2	1990–2000; Databases: not specified	213
DieM 2003/2004 [47]	Institute for Evidence-Based Medicine, Germany	T	3	Period not specified (Last 03/2003); MEDLINE, Cochrane Library, „Best Evidence"	42
Duodecim 2004 [60]	The Finnish Medical Society Duodecim, Finland	D, T	Not specified	Period not specified (Last 03/2004); DARE, „Best Evidence"	Ca. 50††
DVA & VHA 2002 [51]	Department of Veterans Affairs & Veterans Health Administration, U.S.A.	P	14*	Update search 01/2001 to 11/2002; MEDLINE	197
ESC 2002/2001 [62]	European Society of Cardiology, (Europe)	D, T	18	Not specified	196
ICSI 2003 [52]	Institute for Clinical System Improvement, U.S.A.	D, T	11	Not specified	Ca. 50
LLGH 2003 [48]	'Leitliniengruppe' (Group of Family Physicians in Hesse), Germany	(D), T	2	Period not specified MEDLINE	72
NHF/Austr & SANZ 2002 [58]	National Heart Foundation of Australia and Cardiac Society of Australia & New Zealand	D, T	33	Not specified	143
NHF/NZ 2001 [59]	The National Heart Foundation of New Zealand	D, T	19	Period not specified (Last 04/2000); MEDLINE	44
NICE 2003 [61]	The National Collaborating Centre for Chronic Conditions/National Institute for Clinical Excellence, United Kingdom	D, T	15**	Start of the Database until 09/2002; MEDLINE, EMBASE, CINAHL, PsycINFO, AMED, Cochrane Library, EconLit	347‡
OPOT 2000 [57]	Ontario Program for Optimal Therapeutics, Canada	P	Not specified	Not specified	24
UM 2001 [53,54]	University of Michigan, U.S.A.	D, T	6	1994 – 02/1998, + hand searches until 2001; MEDLINE	50
UWH 2001 [49]	Faculty of Medicine, University Witten/Herdecke, Germany	D, T	5 + 2‡‡	Not specified	157

†Limitation on recommendations with evidence levels and/or grading; ††Ascertainment possibly incomplete because of document structure (internet-based version with links to other documents); *Supported by 13 members of The Medical Advisory; **Support from 14 members of the Guideline Reference Group; ‡Additional citations in evidence tables not counted; ‡‡Additional advisors from specialized care; ***Abbreviations:*** D – diagnostics, (D) – diagnostics partly, T – therapy, P – pharmacotherapy only

logical therapy, 10/16 diagnostics), as did the size of the guideline development groups (2 to 33; median 13), the number of citations (24 to 573; median 132), the volume of the full text version (16 to 163 pages in the long versions), and the quantity and layout of additional tools. A complete list of excluded guidelines, together with the reasons for exclusion, is provided as additional web-material (see Additional file 1, Table W4).

Appraisal of the Methodological Quality

Five guidelines [48,50-52,61] were found to be of high quality according to our criteria. Methodological quality varied broadly between both guidelines (ranging from 0.2 to 0.7) and domains (Table 3) and the key results of the most important domains for the SGR were as follows: In domain 3 "rigour of development" 7/16 guidelines scored well: 13/16 used systematic methods in their search for evidence; 12/16 explained methods for formulating their recommendations – two of which included formal consensus techniques; 12/16 provided information on the update process – seven in detail. All guidelines discussed health benefits, side effects, and the risks of applying most (or at least the key) recommendations, 11/16 reported on a formal external peer review process. The main weakness appeared in the description of how the evidence was selected: only 4/16 guidelines reported explicit criteria. In domain 6 "editorial independence" four guidelines scored 0.0 (no information about funding source, no financial disclosures for their group members), one guideline scored 1.0 (full information on financing, and declarations on potential conflicts of interest), the remainder in between.

Framework Development and Data Extraction

In a stepwise approach we derived 27 clinical questions for the framework of our target-guideline. Some questions were partly complex, i.e. they comprised more than one research question (e.g. 'Should patients be encouraged to do exercise training, and if so, what intensity and duration should be aimed for?'). To extract the data we entered all relevant information into evidence tables in the above mentioned format.

Consistency Analysis

Within the framework we identified 35 recommendations. Of these, 25 recommendations were consistent, nine were inconsistent, and one was not rateable (derived from a single guideline). Of 25 consistencies, seven were based on strong evidence (types 1 and 2), 14/25 on weak evidence or expert consensus (type 3), and 4/25 were consistent in content, but had different evidence levels (type 4). Only one inconsistent recommendation (a minor inconsistency about salt and fluid restriction) was based on expert consensus, the remaining inconsistencies differed in their evaluation of the empirical findings (types A and B).

While we found type 1-, 2-, and 4-consistencies mainly in recommendations on pharmacotherapy (Table 4), type 3-consistencies addressed all kinds of clinical questions, from diagnostics to education, monitoring, and the definition of interfaces to specialized ambulatory and inpatient care.

Among nine inconsistencies we classified three as major inconsistencies (type A): (1) The use of brain natriuretic peptide testing in patients when CHF is suspected (Table 5); (2) Angiotensin II receptor blockers (ARB) in addition to ACE inhibitors and beta-blockers (triple-therapy); and (3) ARB in combination with ACE inhibitors in beta-blocker-intolerant patients (substitution of beta-blockers).

Validation

Seventeen partly complex recommendations were categorized as consistency types 1, 2, 4 (n = 11), and inconsistency type B (n = 6) (Table 4). They were further examined in the validation procedure, while type-3-consistencies and major inconsistencies indicated the need for further search for new evidence rather than validation (see Additional file 1, for a complete list of type-3-consistencies and type-A-inconsistencies see Table W5).

To address the clinical questions under investigation the guidelines cited 309 documents, of which we considered 21 studies for re-assessment: 14 systematic reviews (SR) with or without meta-analyses, six RCTs and one post-hoc subgroup analysis of an RCT, since this study was linked to a warning sign in one guideline. The reasons for preclusion of the remaining studies are given below (Table 6). The validation procedure (Table 4) showed most of the recommendations to be justified by the cited evidence sources. Nevertheless, we found incongruity – three concise examples are given below:

(1) Incongruity due to a lack of evidence

All included guidelines recommended the use of ACE inhibitors in all NYHA classes including asymptomatic patients. While 11 guidelines designed differentiated clinical outcomes as therapeutic goals in asymptomatic patients (e.g. improvement in prognosis and hospitalization), three guidelines ranked evidence-level highest for a mortality reduction in this subpopulation, and two guidelines used ambiguous formulations. We found no evidence from the cited studies for a mortality reduction in asymptomatic patients, since this subpopulation was underrepresented in RCTs. Evidence was found only when we identified an health technology assessment report in

Figure 2
Results of the Systematic Guideline Search (Flowchart).

Table 3: Summarized Methodological Quality of Included Guidelines (AGREE Instrument, Standardized Domain Scores in Brackets)

	Domain 1 'Scope and Purpose' 3–12 Points	**Domain 2** 'Stakeholder Involvement' 4–16 Points	**Domain 3** 'Rigour of Development' 7–28 Points	**Domain 4** 'Clarity and Presentation' 4–16 Points	**Domain 5** 'Application' 3–12 Points	**Domain 6** 'Editorial Independence' 2–8 Points
Median	6.5 (0.39)	8.5 (0.38)	17 (0.48)	12.5 (0.71)	4.5 (0.17)	4 (0.33)
Range	4–11 (0.11 – 0.89)	4–12 (0 – 0.67)	11–23 (0.19 – 0.76)	10–16 (0.5 – 1.0)	3–9 (0 – 0.67)	2–8 (0 – 1.0)

further searches [74]. Comparable incongruity was found in recommendations for beta-blocker therapy in asymptomatic patients without prior myocardial infarction.

(2) Incongruity due to the ambiguous use of evidence levels

Fourteen guidelines recommended the use of diuretics, reported evidence levels and graded their recommendations: in 2/14 diuretics were recommended to reduce mortality and morbidity in CHF. The supporting evidence was ranked highest and linked to a systematic review including meta-analysis (published in 2002) [75], which was not accepted by another guideline on the basis of its equivocal methodological quality. In 12 guidelines the main therapeutic goal of diuretics was to control fluid retention: 7/12 ranked the evidence highest, 2/12 second highest and 2/12 lowest (expert opinion). We reappraised the meta-analysis of 17 small sample sized RCTs and confirmed concerns about its methodological quality (no discussion of the methodological quality of included studies, homogeneity assumption after the χ^2-test detected no heterogeneity despite high clinical heterogeneity between the studies, no sensitivity analyses to test the robustness of the results). In conclusion, the evidence level remained unclear: formally, multiple RCTs have shown positive effects on (different) clinical outcomes of diuretics, but the methodological quality of the meta-analyses (and the RCTs themselves) give a high probability of biased results.

(3) Other kinds of incongruence

The evaluation of pharmacotherapy in the source guidelines often focused on treatment benefits and underestimated the risks, such as of adverse events, following the combination of ARBs and ACE inhibitors, or the risk of hyperkalemia following the use of aldosterone antagonists, as recently shown [76,77].

Draft guideline and formulation of the needs for further research

We summarized the SGR results in a draft guideline for the following steps in development and formed a list of specific clinical questions for further research comprising all type-3-consistencies, type-A-inconsistencies and incongruity from the validation process. To adjust the research needs to the given resources these research questions were formally prioritized for the search for new evidence by representatives from the targeted user groups after conducting the SGR (outside the scope of this paper).

Resources

Our methodological concept of the systematic guideline review had to be planned with exceedingly limited resources for the whole project (the budget was € 75.000), but was successfully completed after 8.5 man-months. Starting in January 2004 with the development of the methods concept and systematic searches for guidelines, the first part of the SGR was finished by one researcher (part time: 75%) by the end of June 2004, accounting for 4.5 man-months. From July 2004 to February 2005 the validation procedure was conducted and the draft guideline was written by one researcher (50%), accounting for the remaining four man-months.

Discussion

Our study addressed the need for a high-quality evidence-based guideline to manage patients with CHF in the German primary care context. Internationally, several high-ranking guidelines were/are already in existence, but adaptation methods have not been established yet [78,79]. The guidelines included in this study were heterogeneous in many respects, such as origin, coverage, and the extent of appraised evidence. Nevertheless, the analysis of these guidelines by means of the systematic guideline review enabled us to develop a new evidence-based guideline on a complex condition with comparatively limited resources. In our opinion three steps in the SGR were responsible for resource saving in terms of net effort, as well as for improving the validity and transparency of the target guideline: (1) construction of a framework, (2) consistency analysis, and (3) validation.

Framework Building

CHF is a syndrome with a number of distinct clinical presentations [80]. 'Asking the right questions and asking them right' was named as the core task of guideline development [81], and it is essential for both appraisal of the identified evidence [82-84] and the practical needs to be met by a guideline. Often it is a time-consuming and untransparent process, and there is inevitably a trade-off between depth and breadth of scope in accordance with

Table 4: Results of Consistency Analysis and Validation

Type	Clinical Question	Consistency Analysis†	Validation	Comment
Consistencies				
(1)	Use of ACE inhibitors in systolic CHF, all NYHA classes (incl. asymptomatic patients NYHA class I, with or without history of myocardial infarction)	16/16 'recommended'	Partly justified	Benefit was shown for symptomatic patients (all outcomes incl. mortality), in asymptomatic patients NYHA class I: improvement of prognosis and morbidity, but no evidence for a mortality reduction (see text)
	Use of beta-blockers in systolic CHF, NYHA I post myocardial infarction	11/11 'recommended'	Completely justified	Cited sources provided the reported evidence in form and content
	Use of beta-blockers in systolic CHF, NYHA II-III	16/16 'recommended'	Completely justified	Cited sources provided the reported evidence in form and content
(2)	Use of aldosterone antagonists in systolic CHF, NYHA III/IV	16/16 'recommended'	Justified	Cited sources provided evidence on effectiveness; further research is needed on safety (see text)
	Use of digoxin in systolic CHF with tachyarrhythmia	15/15 'recommended'	Partly justified	Evidence level were revised (see text)
	Control of hypertension in diastolic CHF	2/2 'recommended'	Not justified	Insufficient evidence, further research is needed (see text)
	Use of anticoagulants in patients with the combination of CHF and atrial fibrillation and/or a history of thromboembolism	12/12 'recommended'	-	No re-assessment: recommendations referred to atrial fibrillation (out of scope in the target guideline)
(4)	Exercise Training	13/13 'recommended'	-	No re-assessment: evidence was to be found in a newly identified meta-analysis [63]
	Diuretics in systolic CHF, NYHA II-IV	14/14 'recommended'	Partly justified	Evidence level was revised (see text)
	Use of hydralazine plus ISDN in ACE inhibitor-/ARB-intolerant patients	10/10 'recommended'	-	No re-assessment: no market availability for the fixed combination in the target context
	Harmlessness of long-acting dihydropyridines	7/7 'recommended'	Partly justified	Evidence levels not justified; evidence insufficient, further research is needed
Inconsistencies				
(B)	Salt and fluid restriction (varying quantification)	9/10 'recommended', 1/10 'not recommended'	-	No validation: recommendations based on expert consensus
	Beta-blockers in clinical stable systolic CHF, NYHA IV	13/15 'recommended', 2/15 'not recommended'	Majority was justified, minority was rejected	Positive recommendations completely justified, negative recommendations based on insufficient evidence

Table 4: Results of Consistency Analysis and Validation *(Continued)*

	Beta-blockers in all systolic CHF, NYHA I – no matter whether post myocardial infarction or non-ischemic genesis	7/8 'recommended', 1/8 'consideration recommended'	Majority was not justified, minority was accepted	No evidence for strong recommendation (see text)
	ARB in ACE intolerant patients	15/16 'recommended', 1/16 potentially harmful therapy	Majority justified, minority rejected	Positive recommendations justified, negative recommendations based on insufficient evidence

†Numerical proportion of the mandating to guidelines which covered the scope and reported evidence levels and graded their recommendation. Type-3-consistencies – based on weak evidence – and type-A-inconsistencies are not listed in this table, as they were not included in the validation procedure but needed further research for evidence (a list is provided as additional web-based material, TABLE W5).

development resources [85]. In our approach we extracted questions that were seen as relevant by other guideline development groups, which we then prioritized and refined to assure relevance to our setting. Thus it was a time-saving, systematic and (by publishing it in the methods report) a transparent process.

Consistency Analysis

We carefully compared recommendations and the judgments of underlying evidence reported in the guidelines. Like Kulig et al., we found good overall (external) consistency [86]. Nevertheless, many recommendations were based only on expert consensus or weak evidence. We categorized our findings to bring to light these 'grey zones of clinical practice' [73], but also fields of controversy in CHF management such as BNP testing (Table 5). Our classification supported the specification of questions for further research, and their transparent prioritization. When supported by update-searches to ensure the timeliness of our target guideline, the SGR prevents unnecessary repetition of extensive evidence searches and appraisals: evidence-based consistencies (type-1-consistencies) allow a focused re-appraisal of the most important evidence sources. In type 2-consistencies further research may concentrate on safety aspects, as controlled studies are not generally sufficient to identify risks or their frequency [72].

Validation

We critically re-appraised the most important evidence sources cited in the guidelines to assess whether design, study-population and results were able to support the recommendations in our setting. Certainly, the vast majority of recommendations was justified by empirical findings, but some were not. Examples such as strong recommendations for drug therapy in asymptomatic patients, which were not sufficiently based on the highest level of evidence (not being obsolete, but solely based on the consensus based extrapolation of studies with patients of higher NYHA-classes) shed a light on methodological

Table 5: Case Study about Brain Natriuretic Peptides (BNP) in the Diagnosis of Heart Failure*

Results from the SGR: The recommendations in the source guidelines on the use of BNP tests in patients suspected of heart failure showed a major inconsistency (type A): The test was treated in 7/16 guidelines; recommendations differed completely in content (2/7 'not recommended', 3/7 'recommended under certain circumstances', 1/7 'recommended in every case', 1/7 'recommended ruling out CHF before an echocardiogram'), and in grading.
Further research: We conducted a systematic review of the diagnostic accuracy of this test in primary care. We did not find strong evidence in favour of its use in this setting (most studies were undertaken after referral which implies a potential spectrum bias [64,65], a clear cut-off was not defined, study results were inconsistent, in particular for concomitant diseases and medication) [66-69], supported by two subsequently published systematic reviews [70,71]. The consensus panel agreed not to recommend the test in our target guideline.
Discussion: Inconsistent recommendations in source guidelines may be due to (i) *methodological shortcomings* (recommendations were not setting-specific in 6/7 guidelines that addressed both primary and secondary care; literature searches were stated to be comprehensive in only 2/7 guidelines), or to (ii) *potential conflicts of interests* (4/7 guidelines did not provide any financial disclosures for the authors). Moreover, (iii) *contextual influences* may have guided the recommendations, such as availability (1/7 guidelines recommended the test, as echocardiograms are not widely available), access (BNP tests have market approval throughout Europe but costs are reimbursed by public funding – e.g. in the U.K. – or privately, e.g. in Germany), or intended resource allocation (1/7 guidelines restricted access to the more expensive echocardiogram, as CHF had to be ruled out by BNP and/or electrocardiogram before the referral). Last but not least the BNP test is an emerging technology where typically only limited information of its benefit is available, and initial studies show predominantly optimistic results [72]. Variations in the adoption of a new (healthcare) technology from one country to another, and also from one physician to another are shown to be large, and 'inextricably interwoven' with culture [72]. We know that more highly trained and committed physicians in the community tend to be 'early adopters' [72], and that specialists are less conservative than generalists [72,73]. It might be that the selection of guideline developing groups and their attitudes influenced the decision to include the BNP test.

*This test was developed to distinguish between heart failure and other conditions that show typical symptoms and signs in a patient. In this paper we use BNP as a synonym for itself and others such as NT-proBNP.

Table 6: Selection of Studies for Re-assessment

Publication Type	No. of Cited	No. of Excluded	Reason for Preclusion
Systematic Reviews/Meta-analyses (SR)	36	13	Results outdated by more recent SR
		3	SR reported surrogate outcomes where clinical outcomes were available
		2	SR did not contain target population
		4	SR was out of the scope of the target guideline
Randomized Controlled Trials (RCT)	170	132	RCT included in re-evaluated SR
		8	Results on surrogate outcomes where clinical outcomes were available
		12	Set aside for further comprehensive research
		12	RCT was out of the scope of the target guideline
Traditional Reviews, Editorials	33	33	Provided no systematic evidence
Miscellaneous Clinical Studies	70	69	Study design and/or sample size N<50 were not expected to provide strong evidence; further search for high-level evidence was seen to be more effective than re-appraisal
Total	309	288	

shortcomings even in those guidelines that were of high quality according to the AGREE appraisal. Our reported findings fit in well with recently published studies: McAlister et al. [87] have shown that less than one-third of treatment recommendations (and less than half of those citing RCTs in support of the advocated treatment) were based on high-quality evidence in national evidence-based guidelines for common conditions, in particular when external validity was adequately taken into account. Also Watine et al. [42,88] demonstrated that guideline quality was not necessarily associated with valid recommendations.

Some of our judgments might seem overcritical to the reader. But the questions on how to compare and combine different sorts of evidence (e.g. benefits and risks) or conflicting study results reflect not only the under-use of current approaches to grade the evidence (e.g. the GRADE- or CHEP-system) [89,90] but also epistemological problems in evidence-based medicine [91,92]. Moreover, they point out that the concept of effectiveness is inherently interwoven with normative values [93], which vary depending on cultural context.

In comparison to adaptation procedures – e.g. the most ambitious approach by Fervers et al [35] – our SGR method differs mainly in the systematic consistency analysis and the validation procedure. Other steps are comparable (the systematic search for, selection and critical appraisal of the guideline) or necessary requirements for the SGR (data extraction and information synthesis). Fervers et al [35] emphasized the importance of proving coherence between evidence and recommendations and their applicability and acceptance to the target context [79]. We have shown that consistency analyses and the validation procedure of the SGR demonstrate coherence and further steps in guideline development proved applicability and acceptance.

Furthermore, our SGR method contributed to resource saving: The total costs for the guideline development (incl. the consensus process) were about € 75,000. Though the following examples do not allow a head-to-head comparison, they describe the context: for the 1997 Scottish guideline on 'The management of mild, non-proteinuric hypertension in pregnancy' the total costs were given as £ 66,809 (in those days about € 95,000) [94]; the average expenditure on evidence-based guidelines of the Agency for Health Care Policy and Research (AHCPR) in the nineties was approximately USD1 Million per guideline [95], and at the beginning of 2000 the Agency for Health Quality and Research (AHRQ) spent about USD 250,000 to conduct a systematic review [96].

Our study had some limitations such as the lack of a second review in data extraction and further steps of appraisal. Therefore, a certain observer bias cannot be ruled out. Our budget allowed only a single AGREE appraisal. Fortunately, we were able to compare our results with another working group of the German Agency for Quality in Medicine (AQuMed/ÄZQ) (two appraisers), which carried out a quality appraisal of the same guidelines subsequently and independently of us. Their results were quite similar with a slightly better rating than ours (see Additional file 1, Table W6). Furthermore, we disclosed all SGR material (e.g. evidence tables) to the participants in the consensus process and to the public (methods report in German is available online: http://www.allgemeinmedizin.uni-frankfurt.de/forschung2/herzinsuffizienz_internet.html) to diminish the potential effects of an observer bias on the target guideline.

Also, due to limited resources, we confined the publication languages to English and German, which may lead to a system-related bias (to focus on a specific healthcare system which may vary from the target context). However, this bias is less probable in our study, since we included guidelines from several countries, differing in funding and reimbursement [97,98], availability and accessibility [99,100] of healthcare.

Further, our findings on the resource saving effect of the SGR have to be interpreted within the context of this study: our paper reports on an uncontrolled N = 1 group study and the resource data were collected retrospectively.

Nevertheless, our findings in the SGR were relevant to the recommendations in our guideline, assured their validity, contributed to its transparency, highlighted fields of implicit normativity to prepare an open discourse during the later stages of guideline development (the peer reviews, the consensus process, and the pilot testing), and saved resources by avoiding unnecessary redundancies in the literature search and appraisal. Further studies should be carried out with the involvement of a second reviewer to avoid observer bias.

Conclusion

The systematic guideline review method is a valid means of making use of existing guidelines, and allows a systematic approach to the development of a new guideline. A systematic guideline search aims at the inclusion of guidelines from different healthcare systems, and a systematic comparison of recommendations brings to light both mainstream recommendations and controversies, but also the grey zones of clinical practice. A careful re-evaluation of the most relevant evidence sources assures validity. In the trade-off between breadth and depth, the SGR allows reasonable and transparent prioritization in order to concentrate the development resources on the most important questions for further research. The SGR helps to highlight fields of implicit normativity in guidelines and prepares them for an open discourse within the target context. It abbreviates the initial full evidence review stages, and helps to avoid unnecessary repetition of high-quality research by other guideline authors. In our example it allowed the development of a guideline on the complex clinical issue of chronic heart failure with comparatively small resources. Further studies will have to confirm our results to develop a valid guideline by means of the SGR.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CM had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. She designed the study, acquired, analyzed and interpreted the data and drafted the manuscript. JG, MB and FMG were involved in the study concept and design, supported data acquisition, analysis and interpretation and the critical revision of the manuscript for important intellectual content. AH substantially contributed to the data analysis and interpretation and critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript.

Additional material

Additional file 1

***Supporting Web-based Material**. The file contains the following tables, as mentioned in the text: – List of websites consulted in hand searches: International Organizations (Table W1). – List of websites consulted in hand searches: National Organizations (Table W2). – Checklist by the German Working Group on Health Technology Assessment: an example for systematic reviews with/without meta-analyses (Table W3). – List of excluded guidelines with the reason for preclusion (Table W4). – Results of Consistency Analyses: List of Type-3-Consistencies and Type-A-Inconsistencies (Table W5). – Comparison of different quality appraisals for the included guidelines (Table W6).*

Click here for file

[http://www.biomedcentral.com/content/supplementary/1472-6963-9-74-S1.doc]

Acknowledgements

This project was conducted in cooperation with the „Kompetenznetz Herzinsuffizienz" and the German Society for General Practice and Family Medicine (DEGAM). It was supported by grant 01GI0205 from the German Ministry for Education and Research (BMBF). We thank **Prof. Paul Glasziou**, director of the Centre for Evidence-Based Medicine at the University of Oxford, U.K. for important comments on an earlier draft.

References

1. Thomas L, Cullum N, McColl E, Rousseau N, Soutter J, Steen N: **Guidelines in professions allied to medicine.** *The Cochrane Database of Systematic Reviews* 1999:Art. No.: CD000349.. DOI: 10.1002/14651858.CD000349.
2. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J: **Clinical guidelines. Potential benefits, limitations, and harms of clinical guidelines.** *BMJ* 1999, **318:**527-530.
3. Komajda M, Lapuerta P, Hermans N, Gonzalez-Juanatey JR, van Veldhuisen DJ, *et al.*: **Adherence to guidelines is a predictor of outcome in chronic heart failure: the MAHLER survey.** *Eur Heart J* 2005, **26:**1653-1659.
4. Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J: **Using clinical guidelines.** *BMJ* 1999, **318:**728-730.
5. Shekelle P, Eccles MP, Grimshaw JM, Woolf SH: **When should clinical guidelines be updated?** *BMJ* 2001, **323:**155-157.
6. Woolf SH: **The meaning of translational research and why it matters.** *JAMA* 2008, **299:**211-213.
7. Emanuel EJ, Fuchs VR, Garber AM: **Essential elements of a technology and outcomes assessment initiative.** *JAMA* 2007, **298:**1323-1325.
8. Council of Europe: **Developing a methodology for drawing up guidelines on best medical practices.** *(Recommendation (2001)13 and explanatory memorandum)* 2002 [https://wcd.coe.int/]. ISBN 92-871-4788-4.
9. Glasier A, Brechin S, Raine R, Penney G: **A consensus process to adapt the World Health Organization selected practice recommendations for UK use.** *Contraception* 2003, **68:**327-333.
10. Wabitsch R, Margolis CZ, Malkin JE, Abu-Saad K, Waksman J: **Evaluating the process of tailoring the global HIV/AIDS practice guidelines for use in developing countries.** *Int J Qual Health Care* 1998, **10:**147-154.
11. MacLeod FE, Harrison MB, Graham ID: **The process of developing best practice guidelines for nurses in Ontario: risk assessment and prevention of pressure ulcers.** *Ostomy Wound Manage* 2002, **48:**30-32. 34–38.
12. Voellinger R, Berney A, Baumann P, Annoni JM, Bryois C, *et al.*: **Major depressive disorder in the general hospital: adaptation of clinical practice guidelines.** *Gen Hosp Psychiatry* 2003, **25:**185-193.
13. Graham ID, Harrison MB, Lorimer K, Piercianowski T, Friedberg E, *et al.*: **Adapting national and international leg ulcer practice guidelines for local use: the Ontario Leg Ulcer Community Care Protocol.** *Adv Skin Wound Care* 2005, **18:**307-318.
14. Van Tulder M, Becker A, Bekkering T, Breen A, Gil del Real MT, Hutchinson A, on behalf of the COST B13 Working Group on Guidelines for the Management of Acute Low Back Pain in Primary Care, *et al.*: **European guidelines for the management of acute nonspecific low back pain in primary care. 2006.** [http://www.backpaineurope.org/web/files/WG1_Guidelines.pdf]. Accessed January 28, 2008.
15. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J: **Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature.** *JAMA* 1999, **281:**1900-1905.
16. Burgers JS, Bailey JV, Klazinga NS, Bij AK Van Der, Grol R, Feder G, AGREE Collaboration: **Inside guidelines: comparative analysis of recommendations and evidence in diabetes guidelines from 13 countries.** *Diabetes Care* 2002, **25:**1933-1939.
17. Thomson R, McElroy H, Sudlow M: **Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment.** *BMJ* 1998, **316:**509-513.
18. Beck C, Cody M, Souder E, Zhang M, Small GW: **Dementia diagnostic guidelines: methodologies, results and implementation costs.** *J Am Geriatr Soc* 2000, **48:**1195-1203.
19. Silagy CA, Stead LF, Lancaster T: **Use of systematic review in clinical practice guidelines: case study of smoking cessation.** *BMJ* 2001, **323:**833-836.
20. Ravago TS, Mosniam J, Alem F: **Evaluation of community acquired pneumonia guidelines.** *J Med Sys* 2000, **24:**289-296.
21. McIntosh A: **Guidance: what role should consensus play in guideline development?** In *Making use of guidelines in clinical practice* Edited by: Hutchinson A, Baker R. Oxon: Radcliff Medical Press Ltd; 1999:63-81.
22. Choudhry NK, Stelfox HT, Detsky AS: **Relationships between authors of clinical practice guidelines and the pharmaceutical industry.** *JAMA* 2002, **287:**612-617.
23. Taylor R, Giles J: **Cash interests taint drug advise.** *Nature* 2005, **437:**1070-1071.
24. Molewijk AC, Stiggelbout AM, Otten W, Dupuis HM, Kievit J: **Implicit normativity in evidence-based medicine: a plea for integrated empirical ethics research.** *Health Care Anal* 2003, **11:**69-92.
25. DeMaeseneer J, Derese A: **European general practice guidelines: a step too far?** *Eur J Gen Pract* 1999, **5:**86-87.
26. Aldrich R, Kemp L, Williams JS, Harris E, Simpson S, *et al.*: **Using socioeconomic evidence in clinical practice guidelines.** *BMJ* 2003, **327:**1283-1285.
27. Berg M, Ter Meulen R, Burg M Van Den: **Guidelines for appropriate care: the importance of empirical normative analysis.** *Health Care Anal* 2001, **9:**77-99.
28. Eisinger F, Geller G, Burke W, Holtzman NA: **Cultural basis for differences between US and French clinical recommendations for women at increased risk of breast and ovarian cancer.** *Lancet* 1999, **353:**919-920.
29. Christiaens T, De Backer D, Burgers J, Baerheim A: **Guidelines, evidence, and cultural factors. Comparison of four European guidelines on uncomplicated cystitis.** *Scand J Prim Health Care* 2004, **22:**141-145.
30. Jonsen AR, Siegler M, Winslade WJ: **Clinical ethics. A practical approach to ethical decisions in clinical medicine.** 6th edition. New York, NY: McGraw-Hill Medical Publishing Division; 2006.
31. Relman AS: **Medical professionalism in a commercialized health care market.** *JAMA* 2007, **298:**2668-2670.
32. Welle DA, Ross RS, Detsky AS: **What is the different about the market for health care?** *JAMA* 2007, **298:**2785-2787.
33. Budetti PP: **Market justice and US health care.** *JAMA* 2008, **299:**92-94.
34. Dixon J, Chantler C, Billings J: **Competition on outcomes and physician leadership are not enough to reform health care.** *JAMA* 2007, **298:**1445-1447.
35. Fervers B, Burgers JS, Haugh MC, Latreille J, Mlika-Cabanne N, *et al.*: **Adaptation of clinical guidelines: literature review and proposition for a framework and procedure.** *Int J Qual Health Care* 2006, **18:**167-176.
36. Gerlach FM, Beyer M, Berndt M, Szecsenyi J, Abholz HH, *et al.*: **The DEGAM-concept – development, dissemination, implementation and evaluation of guidelines for general practice.** *[German] Z Arztl Fortbild Qualitatssich* 1999, **93:**111-20.
37. Muth C, Gensichen J, Butzlaff M: **German Society for General Practice and Family Medicine. Guideline No. 9 'Heart Failure'.** Duesseldorf: Omikron Publishing; 2006. [in German] ISBN: 3-936572-07-0 (Part 1) and 3-936572-10-0 (Part 2). [Abridged Internet-Versions are available at: http://www.degam.de/leitlinien/9_herzinsuffizienzs.html]
38. Cohen J: **A coefficient of agreement for nominal scales.** *Educ Psychol Meas* 1960, **20:**37-46.
39. The AGREE Collaboration: **Appraisal of Guidelines for Research & Evaluation (AGREE) instrument training manual.** 2003 [http://www.agreecollaboration.org/pdf/aitraining.pdf]. Accessed May 10, 2004.
40. The AGREE Collaboration: **Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project.** *Quality and Safety in Health Care* 2003, **12:**18-23.
41. Graham ID, Harrison MB: **Evaluation and adaptation of clinical practice guidelines.** *Evid Based Nurs* 2005, **8:**68-72.
42. Watine J, Friedberg B, Nagy E, Onody R, Oosterhuis W, *et al.*: **Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners.** *Clin Chem* 2006, **52:**65-72.
43. Centre for Evidence Based Medicine (CEBM): **Levels of evidence.** 2001 [http://www.cebm.net/index.aspx?o=1025]. last accessed April, 14th, 2008

44. Siebert U, Muth C, Sroczynski G, Velasco-Garrido M, Gerhardus A, Gibis B: **[Liquid-based slide preparations and automated devices in examining cervical smears – systematic review of medical efficacy and health economics, and decision analysis].** In *Health Technology Assessment 35* Sankt Augustin: Asgard-Verlag; 2003. ISBN 3-537-27035-6.
45. Arzneimittelkommission der deutschen Aerzteschaft: **Chronische Herzinsuffizienz. Empfehlungen zur Therapie der chronischen Herzinsuffizienz. Gemeinsame Empfehlung mit der Deutschen Gesellschaft fuer Kardiologie – Herz- und Kreislaufforschung.** *AVP-Sonderheft Therapieempfehlungen* 2001.
46. Hoppe UC, Erdmann E, Kommission Klinische Kardiologie: **[Guidelines for the treatment of chronic heart failure. Issued by the Executive Committee of the German Society of Cardiology – Heart and Circulation Research, compiled on behalf of the Commission of Clinical Cardiology in cooperation with Pharmaceutic Commission of the German Physicians' Association].** *Z Kardiol* 2001, **90:**218-237.
47. Kaiser T, Jennen E, Sawicki PT: **Entscheidungsgrundlage zur evidenzbasierten Diagnostik und Therapie bei Disease Management Programmen für Herzinsuffizienz bei systolischer linksventrikulaerer Dysfunktion.** 2003 [http://www.di-em.de/publikationen.php?idtopic=17&la=de]. Accessed March 15, 2004.
48. Schubert I, General Practitioners' Guideline Group of Hessen: **[Guideline report of the Guideline Group of Hessen – general practice therapy circle: management of chronic heart failure].** *Z Arztl Fortbild Qualitatssich* 2003, **97:**145-150.
49. University Witten/Herdecke: **Leitlinie Diagnose und Therapie der chronischen Herzinsuffizienz.** 2001 [http://www.evidence.de/Leitlinien/leitlinien-intern/Herzinsuffizienz_Start/herzinsuffizienz_start.html]. Accessed March 15, 2004.
50. Hunt SA, Baker DW, Chin MH, Cinquegrani MP, Feldman AM, Francis GS, *et al.*: **ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice guidelines (Committee to Revise the 1995 Guidelines for the Evaluation and Management of Heart Failure).** *Circulation* 2001, **104:**2996.
51. Department of Veterans Affairs & Veterans Health Administration: **Pharmacy Benefits Management Strategic Healthcare Group and the Medical Advisory Panel. The pharmacologic management of chronic heart failure.** 2001 [http://www.oqp.med.va.gov/cpg/CHF/CHF_Base.htm]. Washington, DC: Veterans Health Administration, Department of Veterans Affairs. Updated December 2002. PBM-MAP Publication No. 00-0015 Accessed March 15, 2004.
52. Institute for Clinical System Improvement: **Health care guideline. Congestive heart failure in Adults.** 2003 [http://www.icsi.org/knowledge/detail.asp?catID=29&itemID=161]. Accessed May 11, 2004.
53. Chavey WE 2nd, Blaum CS, Bleske BE, Harrison RV, Kesterson S, Nicklas JM, American Heart Association: **Guideline for the management of heart failure caused by systolic dysfunction: Part I. Guideline development, etiology and diagnosis.** *Am Fam Physician* 2001, **64:**769-774.
54. Chavey WE 2nd, Blaum CS, Bleske BE, Harrison RV, Kesterson S, Nicklas JM, American Heart Association: **Guideline for the management of heart failure caused by systolic dysfunction: Part II. Treatment.** *Am Fam Physician* 2001, **64:**1045-1054.
55. Canadian Cardiovascular Society: **The 2002/3 Canadian Cardiovascular Society consensus Guideline. Update for the diagnosis and management of heart failure.** [http://www.ccs.ca/download/CCSHFConsUpdateDraft0103EMail.pdf]. Accessed March 15, 2004.
56. Canadian Cardiovascular Society: **The 2001 Canadian Cardiovascular Society consensus guideline update for the management and prevention of heart failure.** *Printed version: Can J Cardiol* 2001, **17(Suppl E):**5E-25E.
57. Ontario Program for Optimal Therapeutics: **Ontario drug therapy guidelines for chronic heart failure in primary care.** 2000 [http://www.opot.org/guidelines/chf.pdf]. ISBN: 1-894706-03-X. Accessed February 17, 2004.
58. National Heart Foundation of Australia and Cardiac Society of Australia & New Zealand: **Guidelines on the contemporary management of the patient with chronic heart failure in Australia.** 2002 [http://www.csanz.edu.au/guidelines/practice/Chronic_Heart_Failure.pdf]. Accessed March 9, 2004.
59. The National Heart Foundation of New Zealand: **A guideline for the management of heart failure.** 2001 [http://www.nzgg.org.nguidelines/dsp_guideline_popup.cfm?guidelineCatID=32&guideli neID=26]. Accessed March 17, 2004.
60. The Finnish Medical Society Duodecim: **Chronic heart failure.** 2004 [http://www.ebm-guidelines.com]. Accessed March 22, 2004.
61. The National Collaborating Centre for Chronic Conditions: **Chronic heart failure. National clinical guideline for diagnosis and management in primary and secondary care. NICE Guideline No. 5.** *National Institute for Clinical Excellence; ISBN: 1-84257-324-1* 2003 [http://www.nice.org.uk]. Accessed March 17, 2004.
62. Remme WJ, Swedberg K: **Guidelines for the diagnosis and treatment of chronic heart failure.** *Eur Heart J* 2001, **22:**1527-1560. Erratum in Eur Heart J 2001, **22:** 2217–2218.
63. Piepoli MF, Davos C, Francis DP, Coats AJ, ExTraMATCH Collaborative: **Exercise training meta-analysis of trials in patients with chronic heart failure (ExTraMATCH).** *BMJ* 2004, **328(7433):**189.
64. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, *et al.*: **Empirical evidence of design-related bias in studies of diagnostic tests.** *JAMA* 1999, **282:**1061-1066.
65. Knottnerus JA, van Weel C, Muris JW: **Evaluation of diagnostic procedures.** *BMJ* 2002, **324:**477-480.
66. Doust JA, Glasziou PP, Pietrzak E, Dobson AJ: **A systematic review of the diagnostic accuracy of natriuretic peptides for heart failure.** *Arch Int Med* 2004, **164:**1978-1984.
67. Alberta Heritage Foundation for Medical Research (AHFMR): **B-type natriuretic peptide for diagnosing congestive heart failure.** *TechNote 46* 2004. (Updated January 2005). [http://www.ihe.ca/documents/tn46.pdf. Last assessed January 3, 2008.]
68. Craig J, Bradbury I, Cummings E, Downie S, Foster L, Stout A: **The use of B-type natriuretic peptides (BNP and NT-proBNP) in the investigation of patients with suspected heart failure.** *NHS Quality Improvement Scotland (NHS QIS). Health Technology Assessment Report No. 6, 2005. ISDN 1-903961-49-1* . [http://www.nhshealthquality.org/nhsqis/files/bnp_final_report.pdf. Assessed June 21, 2005.]
69. Alehagen U, Lindstedt G, Eriksson H, Dahlstrom U: **Utility of the aminoterminal fragment of pro-brain natriuretic peptide in plasma for the evaluation of cardiac dysfunction in elderly patients in primary health care.** *Clin Chem* 2003, **49(8):**1337-1346.
70. Davenport C, Cheng EY, Kwok YT, Lai AH, Wakabayashi T, *et al.*: **Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: a systematic review and meta-analysis.** *Br J Gen Pract* 2006, **56:**48-56.
71. Battaglia M, Pewsner D, Jüni P, Egger M, Bucher HC, Bachmann LM: **Accuracy of B-type natriuretic peptide tests to exclude congestive heart failure: systematic review of test accuracy studies.** *Arch Intern Med* 2006, **166:**1073-1080.
72. Banta D, Luce B: **Health technology and its assessment.** Oxford: Oxford University Press; 1993:35 ff-65 ff.
73. Naylor CD: **Grey zones of clinical practice: some limits to evidence-based medicine.** *Lancet* 1995, **345:**840-842.
74. Shekelle P, Rich M, Morton S, *et al.*: **Pharmacologic management of heart failure and left ventricular systolic dysfunction: Effect in female, black, and diabetic patients, and cost effectiveness.** *Evidence Report/Technology Assessment No. 82 (Prepared by the Southern California-RAND Evidence-based Center under Contract No 290-97-0001)* 2003 [http://www.ahrq.gov]. AHRQ Publication No. 03-E045. Rockville, MD: Agency for Healthcare Research and Quality Accessed March 19, 2004.
75. Faris R, Flather M, Purcell H, Henein M, Poole-Wilson P, Coats A: **Current evidence supporting the role of diuretics in heart failure: a meta-analysis of randomized controlled trials.** *Int J Cardiol* 2002, **82:**149-58.
76. Phillips CO, Kashani A, Ko DK, Francis G, Krumholz HM: **Adverse effects of combination of angiotensin II receptor blockers plus angiotensin-converting enzyme inhibitors for left ventricular dysfunction.** *Arch Intern Med* 2007, **167:**1930-1936.

77. Juurlink DN, Mamdani MM, Lee DS, Kopp A, Austin PC, *et al.*: **Rates of hyperkalemia after publication of the randomized aldactone evaluation study.** *N Engl J Med* 2004, **351:**543-551.
78. Qaseem A, Vijan S, Snow V, Cross T, Weiss KB, Owens DK: **Clinical Efficacy Assessment Subcommittee of the American College of Physicians. Glycemic control and type 2 diabetes mellitus: the optimal hemoglobin A1c targets. A guidance statement from the American College of Physicians.** *Ann Intern Med* 2007, **147:**417-22.
79. Fervers B, Remy-Stockinger M, Graham ID, Burnand B, Harrison M, *et al.*: **Guideline adaptation: an appealing alternative to de novo guideline development.** *Ann Intern Med* 2008, **148:**563-4.
80. Adams KF Jr: **Developing clinical practice guidelines for heart failure: creative process and practice implications.** *Pharmacother* 2000, **20(11 Pt 2):**379S-384S.
81. Oosterhuis WP, Bruns DE, Watine J, Sandberg S, Horvath AR: **Evidence-based guidelines in laboratory medicine: Principles and methods.** *Clin Chem* 2004, **50:**806-818.
82. Rosenberg W, Donald A: **Evidence based medicine: an approach to clinical problem-solving.** *BMJ* 1995, **310:**1122-1126.
83. Richardson WS, Wilson MC, Nishikawa J, Hayward RSA: **The well-built clinical question: a key to evidence-based decisions.** *ACP J Club* 1995, **123:**A-12.
84. Helou A, Perleth M, Schwartz FW: **[Determining priorities in the development of medical guidelines. 1: Criteria, procedures and actors: a methodological review of international experiences].** *Z Arztl Fortbild Qualitatssich* 2000, **94:**53-60.
85. Hadorn DC, Baker D: **Development of the AHCPR-sponsored heart failure guideline: methodological and procedural issues.** *Jt Comm J Qual Patient Saf* 1994, **20:**539-47.
86. Kulig M, Schulte E, Willich SN: **Comparing methodological quality and consistency of international guidelines for the management of patients with chronic heart failure.** *Eur J Heart Failure* 2003, **5:**327-335.
87. McAlister FA, van Diepen S, Padwal RS, Johnson JA, Majumdar SR: **How evidence-based are the recommendations in evidence-based guidelines?** *PLoS Med* 2007, **4:**e250.
88. Watine JC, Bunting PS: **Mass colorectal cancer screening: Methodological quality of practice guidelines is not related to their content validity.** *Clin Biochem* 2008, **41:**459-66.
89. GRADE Working Group: **Grading quality of evidence and strength of recommendations.** *BMJ* 2004, **328:**1490.
90. McAlister FA: **The Canadian Hypertension Education Program (CHEP) – A unique Canadian initiative.** *Can J Cardiol* 2006, **22:**559-564.
91. Ashcroft RE: **Current epistemological problems in evidence based medicine.** *J Med Ethics* 2004, **30:**131-135.
92. Vineis P: **Evidence based medicine and ethics: a practical approach.** *J Med Ethics* 2004, **30:**126-130.
93. Ashcroft R: **What is clinical effectiveness?** *Stud Hist Phil Biol & Biomed Sci* 2002, **33:**219-233.
94. Foy R, Ramsay CR, Grimshaw JM, Penney GC, Vale L, Thomson A, Greer IA: **The impact of guidelines on mild hypertension in pregnancy: time series analysis.** *BJOG* 2004, **111:**765-70.
95. American College of Cardiology's Foundation: [http://www.acc.org/clinical/bethesda/beth28/9991/convrted/]. accessed 20.08.2005
96. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, Woolf SH: **Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated?** *JAMA* 2001, **286:**1461-7.
97. Busse R: **Health care systems: Germany and Britain compared. Health care delivery – towards an agenda for policy learning between Britain and Germany.** 2002 [http://www.mig.tu-berlin.de/fileadmin/a38331600/2002.publications/2002.busse_AGF.BritainGermany.pdf]. London: Anglo-German Foundation for the Study of Industrial Society Accessed April, 2009.
98. Busse R, Schlette S, eds.: **Health policy developments – international trends and analyses.** 2003:19-36 [http://www.bertelsmann-stiftung.de/bst/de/media/xcms_bst_dms_15030_15031_2.pdf]. Guetersloh: Bertelsmann Foundation Publishers 57–66. Accessed February 14, 2007.
99. Busse R, Schlette S, Weinbrenner S, eds.: **Health policy developments. Issue 4: Focus on access, primary care, health care organization.** 2005 [http://www.wm.tu-berlin.de/fachgebiet/mig/papers/index.html]. Guetersloh: Bertelsmann Foundation Publishers Accessed March 9, 2006.
100. Van Doorslaer E, Masseria C: **OECD Health Equity Research Group. Income-related inequality in the use of medical care in 21 OECD countries.** *Towards high-performing health systems: policy studies* 2004:109-66 [http://www.oecd.org]. Paris: OECD Accessed February 22, 2006.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1472-6963/9/74/prepub

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral