

Quality Measurement Recommendations Relevant to Clinical Guidelines in Germany and the United Kingdom: (What) Can We Learn From Each Other?

INQUIRY: The Journal of Health Care
Organization, Provision, and Financing
Volume 55: 1–10
© The Author(s) 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0046958018761495
journals.sagepub.com/home/inq

Thomas Petzold, PhD, M.Sc.¹, Stefanie Deckert, MPH¹, Paula R. Williamson, Prof.²,
and Jochen Schmitt, MPH, MD¹

Abstract

We conducted a systematic review of clinical guidelines (CGs) to examine the methodological approaches of quality indicator derivation in CGs, the frequency of quality indicators to check CG recommendations in routine care, and clinimetric properties of quality indicators. We analyzed the publicly available CG databases of the Association of the Scientific Medical Societies in Germany (AWMF) and National Institute for Health and Care Excellence (NICE). Data on the methodology of subsequent quality indicator derivation, the content and definition of recommended quality indicators, and clinimetric properties of measurement instruments were extracted. In Germany, no explicit methodological guidance exists, but 3 different approaches are used. For NICE, a general approach is used for the derivation of quality indicators out of quality standards. Quality indicators were defined in 34 out of 87 CGs (39%) in Germany and for 58 out of 133 (43%) NICE CGs. Statements regarding measurement properties of instruments for quality indicator assessment were missing in German and NICE documents. Thirteen pairs of CGs (32%) have associated quality indicators. Thirty-four quality indicators refer to the same aspect of the quality of care, which corresponds to 27% of the German and 7% of NICE quality indicators. The development of a standardized and internationally accepted methodology for the derivation of quality indicators relevant to CGs is needed to measure and compare quality of care in health care systems.

Keywords

clinical guideline, quality indicator, quality measurement, evaluation, core outcome set

Introduction

Since *To Err Is Human*¹ in 2000, efforts on the measurement and gradual improvement of the quality of medical treatment have been important parts of health care systems. The assurance and steady development of quality of care is a nationwide aim of national funded and voluntary initiatives in many countries and health care systems around the world.^{2–5} For several years, these efforts have been complemented by various research activities in the fields of health services and implementation research.⁶ With regard to high-quality care, great importance was ascribed to clinical guidelines (CGs).⁷ CGs provide a widespread and well-established method to disseminate systematic and structured recommendations for patient care.⁸ The recipients of the recommendations comprise outpatient, inpatient, and rehabilitative care and therefore a large proportion of medical care providers. The recommendations are based on the best available evidence at the time of creation. The responsibility for and applied methods of CG development differ internationally. Germany and

the United Kingdom constitute 2 examples for the development of CGs. The classification of German CGs is divided according to the approach to their development and contains 3 grades: S1, S2 (separated in S2e [evidence] or S2k [consensus]), and S3 (systematic evidence synthesis and consecutive structured consensus).⁹ The development of CGs is carried out through voluntary initiatives of medical societies or in funded structured CG programs (Disease Management

¹Centre for Evidence-based Healthcare, Technical University Dresden, Germany

²MRC North West Hub for Trials Methodology Research, Institute of Translational Medicine, University of Liverpool, UK

Received 23 December 2016; revised 25 April 2017; accepted 4 December 2017

Corresponding Author:

Jochen Schmitt, Center for Evidence-Based Healthcare, University Hospital Carl Gustav Carus Dresden, Technical University Dresden, Fetscherstraße 74, 01307 Dresden, Germany.
Email: Jochen.Schmitt@mdk-sachsen.de



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons

Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Guidelines and German Guideline Program in Oncology). Both are hosted by the Association of the Scientific Medical Societies in Germany (AWMF). In the United Kingdom, CGs are developed and disseminated in 3 different frameworks through the National Institute for Health and Care Excellence (NICE) for England and Wales, the Scottish Intercollegiate Guidelines Network (SIGN) for Scotland, and the Guidelines and Audit Implementation Network (GAIN) for Northern Ireland. We only include CGs of NICE because of their widespread dissemination. NICE was established in 1999 by the government to “reduce variation in the availability and quality of National Health Service treatments and care.”¹⁰

Different approaches and tools are used in routine practice to measure the quality of medical care. Quality indicators are most commonly used to determine the quality of care.^{11,12} Defined reference intervals distinguish between different levels of health care quality. Data to define reference intervals may be derived from the literature, expert opinions, or statistical specifications.¹³ Quality indicators are instruments to measure structural, processual, and outcome quality in routine care.¹⁴ They are thus measurement instruments that can be expressed as definitions, single-item instruments, or multi-item instruments. They must meet specific requirements, such as clear definitions, rationale for use, and a transparent definition process to achieve conclusive results. For process and outcome quality indicators that are measured by measurement instruments, those requirements include criteria of methodology and feasibility.¹⁵ Methodological criteria are clinimetric properties, that is, content validity, construct validity, internal consistency, interobserver reliability, test-retest reliability, and sensitivity to change/responsiveness.^{16,17} Feasibility/acceptability criteria include the level of implementation of procedures or the sustainability of treatment recommendations for the quality of care.¹⁵

The rigorous and transparent development of meaningful quality indicators with high validity, reliability, and feasibility is a prerequisite for an evidence-driven development of health care systems toward medical care that is safer, more effective, and more efficient.^{18,19} Ideally, processual and outcome quality indicators for the same medical problem should be comparable across health care systems to allow international comparisons as an additional source of evidence for health care system research and development. Quality indicators for structural quality can differ according to the economic or health care system environment. The aims of the present study were (1) the systematic analysis and comparison of the methodological approaches in the development of quality indicator recommendations relevant to CGs in Germany and NICE and (2) the critical appraisal of the content, validity, and reliability of CG-recommended quality indicators between the 2 countries with a specific focus on quality indicators from pairs of guidelines on the same medical problem.

Methods

We undertook a systematic review on all German S3 guidelines (effective November 30, 2013)²⁰ and all NICE CGs

(effective April 1, 2014). The relevant literature was independently searched and assessed by 2 reviewers, and relevant information concerning quality indicators was extracted in a standardized manner. We analyzed the publicly available CG databases of the AWMF and NICE. Following an a priori defined protocol, we extracted the following information from each available, valid and published CG identified: corresponding, leading, or first author; year published; number of defined quality indicators; and quality dimensions according to Donabedian,¹⁴ that is, indicator for structure, process, or patient outcome quality. The extracted data were summarized on CG level, for German CG on the level of the methodological approaches and for NICE CG on the level of the title of the related quality standard and the year of whose publication. If the independent reviewers (T.P., S.D.) extracted different information, a third reviewer (J.S.) was consulted and the information was discussed.

Further analyses were conducted for pairs of CG. Analyses of quality indicators of pairs of CG focusing on identical medical problems were compared according to the content of the quality indicators, the methodology used to define quality indicators, data sources of quality indicators, the rationale and CG recommendation of the quality indicators, and their clinimetric properties. For NICE quality indicators, also the information of the quality standard was extracted and analyzed. Quality standards are used to develop outcome measures which represent the quality of care for a specific medical problem.⁵ The relevance of the different clinimetric properties differs between quality indicators which are based either on definitions or on (latent) constructs. Thus, the relevant clinimetric properties for quality indicators were categorized as shown in Figure 1.

In addition, the methodology manuals of the German and NICE CG developers were analyzed to identify instruments for measuring the feasibility of guidelines concerning the assessments of the quality indicator recommendations.

The execution of a risk of bias assessment was not possible due to the research question and the information sources.

Results

Methodology of Quality Indicator Derivation in CGs

The definition and dissemination of quality indicators relevant to CGs occurs both in Germany and NICE.²⁰⁻²⁴ The recommended methodology of the derivation of quality indicators differed between Germany and NICE.

In Germany, 3 groups developed CGs with different methodological approaches:

- Voluntary initiatives of medical societies (VIMS)
- The Disease Management Guidelines (DMG)
- The German Guideline Program in Oncology (GGPO).

In DMG und GGPO, quality indicators are derived from CG recommendations graded “A” or with the syntax/modality “shall.” In voluntary CG initiatives of medical societies, the

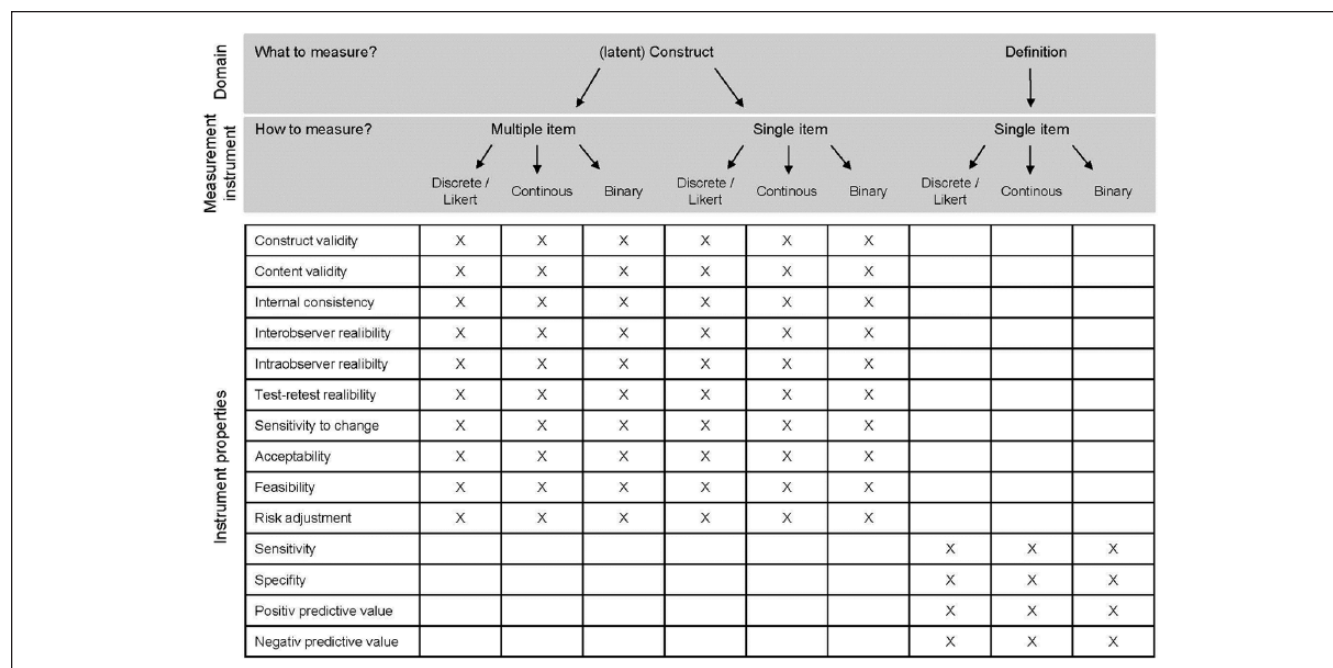


Figure 1. Categorization of quality indicators by their measurement basis.

process of quality indicator derivation is not described. In the funded structured German CG programs DMG und GGPO, quality indicators were derived from CG recommendations (Figure 2). Existing quality indicators pertaining to specific medical issues were identified by a systematic literature search and evaluated by experts with the help of the QUALIFY instrument.²⁵ The QUALIFY instrument provides guidance for the structured assessment of quality indicator development and validation. Selected experts are asked for each quality indicator to assess the following:

- Relevance: importance of the quality characteristic captured by the quality indicator for patients and the health care system, benefit, consideration of potential risks/side effects
- Scientific soundness/indicator evidence: clarity of definition, reliability, ability of statistical differentiation, risk adjustment, sensitivity, specificity, validity;
- Feasibility: interpretability for patients and the interested public; interpretability for physicians and nurses; indicator expression can be influenced by providers; data availability; data collection effort; barriers for implementation considered; correctness of data can be verified; completeness of data can be verified; complete count of data sets can be verified.²⁵

According to the QUALIFY instrument,²⁵ positively rated quality indicators shall be linked to guideline recommendations.²¹

There is no standard specifying at what stage in the process of the guideline work quality indicators are to be selected for recommendation. The definition of quality indicators in

voluntary initiatives of medical societies is optional. In DMG and GGPO, a systematic approach is pursued,^{21,22} which is based on a systematic review of existing and used quality indicators. The detected quality indicators are derived only for strong CG recommendations or with the syntax/modality “shall” and evaluated by an expert committee applying the QUALIFY instrument (Figure 2).

In England and Wales, the NICE will be mandated to develop CGs together with medical societies and organizations. Out of these CGs, quality standards were derived in which quality statements are described. The quality statements are based on CG recommendations from NICE and other organisations.²⁴ Quality standards are associated with the CG and a concise set of measurable quality statements to improve the quality of care in a particular area of health care. Quality statements are composed of quality indicators to measure specific structures, processes, or outcomes in health care.²⁶ Quality indicators are clearly defined²⁷ and critically appraised and evaluated by experts with regard to their utility for the sustainable CG implementation. For new quality standards in new settings or services, a field testing in the form of consultation is recommended. The consultation includes examining the relevance, utility, clarity, potential impact, and the acceptability of the quality standards with providers, payers, and patients.²⁸ The comments of providers, payers, and patients as results of the field testing are discussed by the Quality Standard Advisory Committee (QSAC). The quality standard is refined with the input from providers, payers, and patients and the QSAC members. This is followed by the process of internal quality assurance, consistency checking, and approval by an associate director in the NICE quality

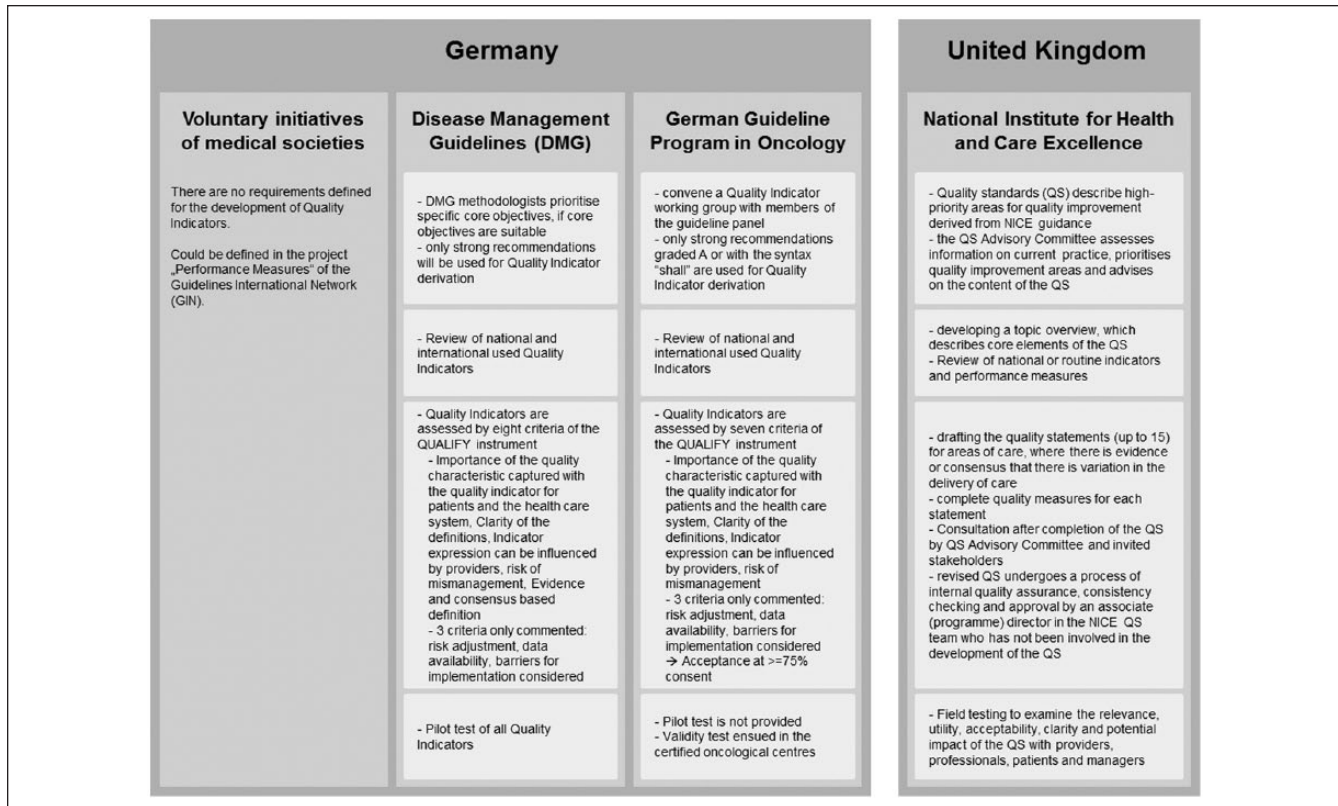


Figure 2. Methodology of the development of quality indicators in clinical guidelines in Germany and the United Kingdom.

standard team. This is to ensure the validity of the content prior to the publication of the quality standards (Figure 2).

The Department of Health determines new relevant public health topics and requests NICE to develop quality standards. New topics for CGs and quality standard development are derived through new government or commissioner priorities.²⁸

Quality Indicator Recommendations

In Germany, 183 S3 CGs (effective November 30, 2013) were identified. Eighty-seven of 183 S3 CGs are available and published. The remaining 96 S3 CGs were preliminarily registered or under revision and therefore not included in further analyses.

In 34 (39%) of the included 87 CGs, quality indicators have been defined (Figure 3), with a total of 394 quality indicators being recommended (2-52 per CG). Three hundred forty-one quality indicators in 29 CGs (85%) reported information on the precise definition (numerator and denominator). The rationale and CG recommendation were documented in 5 CGs (15%). Fifty-eight quality indicators (15%) are based on an underlying construct (single- or multi-item instrument) as measurement basis. Three hundred thirty-six quality indicators (85%) are based on definitions by quality indicator developers. Statements regarding data source and clinimetric properties of quality indicators were reported for none of the 394 quality indicators. Twenty-one (5%)

recommended quality indicators measure structural quality, 340 (86%) processual quality, and 33 (8%) outcome quality (Online Supplemental Table 3).

From NICE CGs, a total of 133 CGs and 58 quality standards were available (effective April 1, 2014), all of which recommended quality indicators relevant to CG (Figure 3). A total of 1795 quality indicators (9-78 per CG) were included in the 58 identified quality standards. All quality indicators out of quality standards that are relevant to CG were derived from CG recommendations.^{24,29} For all quality indicators, information on numerator, denominator, data source, rationale, and CG recommendation of the quality indicators was provided. Four hundred eleven quality indicators (23%) are based on single- or multi-item instruments and 1384 (77%) on definitions. As in German CG, statements regarding clinimetric properties of the defined quality indicators are missing for all quality indicators recommended. All quality indicators out of quality standards that are relevant to CG were tested, if they are applicable for routine care. The results of the pilot testing were collected at NICE and used to improve the quality indicator, for example, suitability of the data source. The results of the field testing were not listed in the quality standard report. Six hundred ninety-nine (39%) recommended quality indicators measure structural quality, 717 (40%) process quality, and 379 (21%) outcome quality (Online Supplemental Table 3).

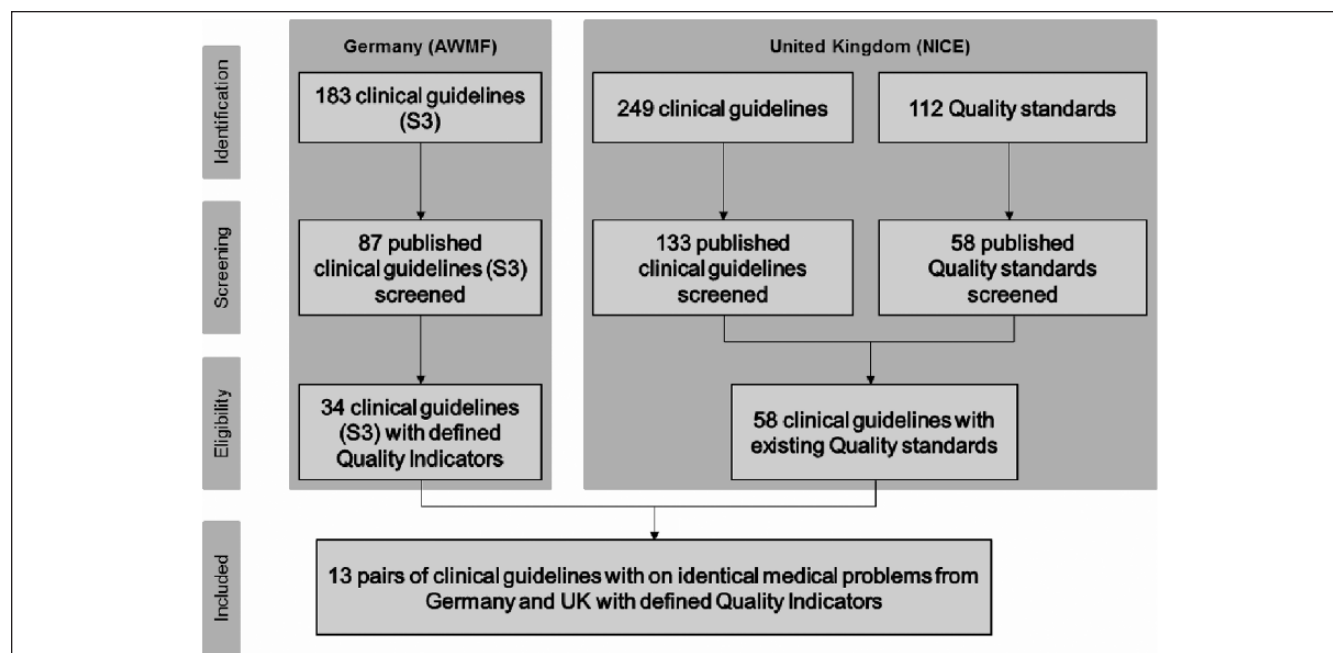


Figure 3. Number of clinical guidelines with quality indicators in Germany and the United Kingdom.

Table 1. Derived Quality Indicators of CGs in Germany and the United Kingdom.

		CGs in Germany		
		With quality indicators	Without quality indicators	No CG available
CGs of NICE	With quality indicators	13	12	38
	Without quality indicators	7	9	54
	No CG available	14	32	

Note. NICE = National Institute for Health and Care Excellence; CG = clinical guideline.

Analysis of Quality Indicators Related to Pairs of NICE and German CGs

Forty-one of the German and British CGs focus on the same medical problem. Thirteen of these 41 pairs of CGs include quality indicators (Table 1). In these 13 pairs, a total of 128 quality indicators were defined in German CGs. A total 468 quality indicators were defined in NICE CGs. Four (3%) of the 128 quality indicators of German CG embrace structural quality, 112 (88%) process quality, and 12 (9%) outcome quality. One hundred seventy-two (37%) indicators relevant to NICE CGs capture structural quality, 232 (50%) process quality, and 64 (14%) outcome quality (Table 2).

In the 13 pairs of CGs related to identical medical problems, 34 quality indicators refer to the same aspect of the quality of care. This corresponds to 27% (34 of 128) of the German quality indicators and 7% (34 of 468) of NICE quality indicators. Ten of 13 German CGs are exclusively process indicators compared with 0 of 13 for NICE. The rationales

for quality indicator recommendation also differed substantially. Twenty-three of 34 rationales refer to the same aspect of quality of care. German and NICE quality indicators differ in 11 rationales. Rationale is a statement in which the specific aspect of health care is described, assessed by the quality indicator.³⁰ In the following CGs, there was a similar alignment of quality indicators: depression in adults. In this pair of CGs,^{31,32} the severity of symptoms was considered as a measurable outcome in Germany and by NICE. In other CGs such as those for psoriasis,³³ no common objectives of quality indicators in NICE and German CGs were identified. In the German CGs, the average Psoriasis Area Severity Index (PASI),³⁴ Dermatology Life Quality Index (DLQI),³⁵ and the inability to work are measured as the defined quality indicators.³³ In the corresponding NICE CG “Psoriasis,”³⁶ 15 quality indicators were recommended to measure disease severity at diagnosis; the impact of the disease on physical, psychological, and social well-being at diagnosis; and the receipt of diagnostic measures.³⁶

Table 2. Comparison of CGs (n = 13) in Germany and From NICE With Defined Quality Indicators Related to the Same Aspect of Quality of Care.

AWMF CG	Number of quality measures for				Methodological approach for quality indicator definition	NICE CG (number of CGs and QS)	QS	Number of quality measures for				Number of quality statements
	Structure	Process	Outcome	Total				Structure	Process	Outcome	Total	
Smoking cessation in COPD	0	3	0	3 ²	Not specified	COPD: management of COPD in adults in primary and secondary care (CG 101, QS 10)	COPD	25	17	8	50 ²	13
Prophylaxis of VTE	0	21	0	21 ¹	Consensus	VTE: reducing the risk (CG 92, QS 3)	VTE prevention	8	8	0	16 ¹	7
Epidemiology, diagnosis, treatment, and management of uncomplicated bacterial community-acquired urinary tract infections in adult patients	2	3	5	10	Not specified	Lower urinary tract symptoms in men (CG 97, QS 45)	Lower urinary tract symptoms	8	10	2	20	8
Chronic heart insufficiency	0	9	0	9 ⁴	QUALIFY	Chronic heart failure: management of chronic heart failure in adults in primary and secondary care (CG 108, QS 9)	Chronic heart failure	16	25	1	42 ⁴	13
Asthma	0	9	0	9 ⁴	QUALIFY	Asthma (CG 12, QS 25)	Asthma	12	12	1	25 ⁴	11
Kidney disease in diabetes in adulthood	0	6	0	6 ¹	QUALIFY	Chronic kidney disease (CG 182, QS 5)	Chronic kidney disease	16	14	7	37 ¹	15
Prevention, diagnosis, treatment, and aftercare of lung cancer	0	9	0	9 ⁴	Consensus	Lung cancer: the diagnosis and treatment of lung cancer (CG 121, QS 17)	Lung cancer for adults	15	34	18	67 ⁴	15
Unipolar depression	0	17	0	17 ⁷	Not specified	Depression in adults (CG 90, QS 8)	Depression in adults	15	16	0	31 ⁷	13
Diagnosis, treatment, and follow-up of breast cancer	0	12	0	12 ⁵	Consensus, QUALIFY	Early and locally advanced breast cancer advanced breast cancer (CG 80, QS 12)	Breast cancer	18	30	9	57 ⁵	13
Colorectal cancer	0	10	0	10 ²	Consensus, Study based	Colorectal cancer: the diagnosis and management of colorectal cancer (CG 131, QS 20)	Colorectal cancer	8	12	4	24 ²	8
Type 2 diabetes: education	2	1	0	3 ³	Consensus	Diabetes in adults (CG 17, QS 6)	Diabetes in adults	17	36	13	66 ³	14
Diagnostics, therapy, and follow-up of malignant ovarian tumors	0	12	0	12 ¹	Consensus, Study based	Ovarian cancer: the recognition and initial management of ovarian cancer (CG 122, QS 18)	Ovarian cancer	8	10	0	18 ¹	8
Treatment of psoriasis vulgaris	0	1	7	8	Consensus	Psoriasis: the assessment and management of psoriasis (CG 153, QS 40)	Psoriasis	6	8	1	15	6

Note. High-ranked number of the “total number of quality indicators” represent quality indicators related to identical medical problems in pairs of clinical guidelines. NICE = National Institute for Health and Care Excellence; AWMF = Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften; CG = clinical guideline; QS = quality standard; VTE = venous thromboembolism; COPD = chronic obstructive pulmonary disease.

Discussion

This systematic review highlights substantial differences in the process and methodology of quality indicator development and recommendation between German and NICE CGs. While in Germany 3 different methodological approaches are applied, one common approach is applied in NICE. Quality indicators from NICE CGs have the advantage that unambiguous definitions, data sources, and information on pilot tests are available. Such information would be helpful for quality indicators in German CGs as well to sustainably improve quality measurement according to CGs. Therefore, CG developers in Germany would benefit from a standardized methodology for the derivation of quality indicators relevant to CGs. Quality indicators are measurable indicators of CG implementation to ascertain the improvement of structures, processes, and outcomes in health care. This is why the methodology of quality indicator derivation has to meet standards as high as those for the development of CG recommendations. Other studies^{37,38} that examined quality indicators in CG focused also on sophisticated, evidence-driven methodology of measurable endpoints.

The developmental process and the evaluation of quality indicators in routine care are important steps of quality measurement and improvement. The methodological approaches of the German DMG, GGPO, and NICE represent a structured and transparent approach to the development of quality indicators in CGs. Voluntary CG initiatives of medical societies, which develop the vast majority of S3 CGs in Germany (82%), have currently no clear methodological guidance on how to derive quality indicators based on CG recommendations. As 2 similar methodological approaches (DMG and GGPO) already exist in Germany, a corresponding methodology should be developed and published to support medical societies. There is no defined procedure in all 4 methodological approaches (3 from Germany, 1 from NICE) for the evaluation of quality indicators. In the NICE CG manual, a collection of feedback is reported, which influences the updating process of the quality indicators. A clinimetric testing of the quality indicators in routine care is not provided in any methodology. As a consequence, no statements on the quality of the quality indicators are available. This is an important limitation as we know from evidence-based medicine that measurement instruments of unclear quality overestimate, underestimate, or misjudge quality effects of interventions.¹²

Three hundred ninety-four quality indicators were found in German CGs and 1795 in CGs of NICE. The majority of these quality indicators (86% in Germany and 40% of NICE) are used to measure process quality. Only 8% in German CGs and 21% of NICE CGs of the defined quality indicators measure outcome quality. The measurement of health care outcomes is of great interest, especially for patients and their relatives. Also health insurance funds need data on outcome quality to support their assured when choosing a service provider or single treatment options. Quality indicators should be based on health care data to derive realistic treatment options. The majority of quality indicators (85% in German CGs and 77%

in NICE CGs) are based on definitions of experts or developers. The use of single- or multi-item instruments would be helpful to prove the goodness of quality indicators.

This systematic analysis of CGs identified pairs of German and British CGs focusing on the same medical issue. However, the quality indicators recommended in these pairs of CGs differ in their quantity, definition, and rationale. The rationale and CG recommendations are relevant for the derivation of quality indicators. In Germany, it is unclear which guideline recommendations are used to derive quality indicators.

The missing objectives for health care performance and outcomes in Germany are a possible reason for the development and implementation different quality measures. Only with distinct objectives for health care can measurable, relevant and sustainable endpoints of health care be developed and implemented. In the United Kingdom, health care objectives exist to improve the patient safety and quality of health and health care, whereas no common health care objectives exist in Germany.³⁹

Both in CGs from NICE and from Germany, information on clinimetric properties of the quality indicators recommended, that is, information on their validity and reliability, is missing. However, such information on quality indicators is crucial for their implementation in routine care and a necessary prerequisite to obtain resilient and convincing results.^{20,40} Therefore, a standardized method appears necessary in both countries for the implementation of these requirements.

Core Outcome Sets (COS) could be a possibility to derive quality indicators related to CG. COS represent the minimum set of outcomes that should be measured and reported in investigations or analyses of a specific medical issue. This is a necessary prerequisite to compare results of studies and to optimize the use of research evidence for clinical decision making. COS are being increasingly developed for clinical trials and effectiveness studies, on the basis of a systematic appraisal of research evidence and consensus by an (international) multidisciplinary expert panel.⁴¹ The COMET (Core Outcome Measures in Effectiveness Trials; www.comet-initiative.org) Initiative provides an overview of existing COS.^{42,43} Core outcome domains should be assessed with adequate measurement instruments, that is, instruments with high validity, reliability, and responsiveness.^{44,45} A roadmap for the systematic development and implementation of COS^{46,47} has been suggested by the international Harmonizing Outcomes for Eczema (HOME) initiative.⁴⁸ Currently, the development of quality indicators from German CGs does not include the review of an existing COS for the clinical problem. This might be an effective strategy to translate clinical outcomes into patient-relevant quality indicators to be used in routine care, as done in NICE CG program.

We furthermore suggest to apply the COS principle in the field of quality assurance. National Core Indicator Sets applied in different settings and fields might help set the agenda for process development in health care. An international COS of quality indicators would allow comparisons between different

health care systems. A first comparison of the overlap between COS and the quality indicators defined in pairs of CG from Germany and NICE concerning the same medical condition identified 3 areas in which a COS as well as quality indicators from German and NICE guidelines exist (effective December 20, 2013). These 3 medical fields are prophylaxis of venous thromboembolism,⁴⁹ asthma,⁵⁰ and low back pain.^{51,52} These 3 medical fields might be a good starting point for a comprehensive international process to bridge the assessment of effectiveness in trials with quality assurance in routine care and thus to help overcome the evidence-practice gap.

The Appraisal of Guidelines for Research and Evaluation (AGREE) instrument was developed for the methodical evaluation of CG and is internationally applied.⁵³ A translation for German guidelines exists Deutsches Leitlinien-Bewertungs-Instrument (DELBI).⁵⁴ These 2 checklists determine the quality of CG. Both instruments contain criteria for the evaluation of CG recommendations. These criteria do not refer to quality indicators related to CG. The sustainability of the treatment recommendations in CG is not guaranteed and does not appear in the form of quality indicators in AGREE or DELBI. Evaluated and sustainable CG recommendations are necessary for the methodological quality of CG and their implementation in routine care. Quality indicators could be a suitable method to attest the sustainability in routine care.

Strengths and Limitations of This Study

We applied systematic review methodology to guarantee completeness and validity of the information extracted from CG. The study was conducted based on a priori defined inclusion and exclusion criteria. A systematic literature search in relevant CGs was independently assessed by 2 reviewers, and relevant information concerning quality indicators was extracted in a standardized manner. After comparing the data extraction and consensus discussion, the evidence synthesis took place. A limitation of the work is possibly the restriction in German guidelines on the group of the S3 CGs. These are evidence- and consensus-based and currently represent the gold standard for CG in Germany. Therefore, a transfer of the results of our work to other German guidelines may be limited, and the proportion of guidelines with formulated quality indicators in S1 and S2 CGs may be different. The systematic review was carried out between 30.11.2013 and 30.04.2014. Since that period, the NICE quality standard program has been advanced with 115 published NICE quality standards. We did not extend our data set because in-depth analyses of CG-recommended quality indicators are highly time-consuming. Furthermore, adding further quality standards is not likely to contribute more substantial insights.

Implications for Future Research

The comparison of quality indicators in German and NICE CGs emphasized that different quality indicators are defined

and used in Germany and NICE for similar medical conditions. One possible reason could be the different structures of the underlying health care systems. The different methodologies used to derive quality indicators related to CGs may represent another issue. To enable a learning system of quality measurement with the help of CGs, an international forum should be established. The Guidelines International Network (GIN; www.g-i-n.net) could serve as a suitable organization for promoting the issue of quality measurement in international CGs. A uniform and internationally accepted methodology for the derivation of quality indicators in CGs should be a major aim. This approach would allow a first international comparison of the quality of care in different health care systems.

Conclusion

In Germany and the NICE, almost the same diseases are represented in the guideline framework. In these CGs, similar CG recommendations are derived. However, the efficacy of those recommendations is evaluated with different quality indicators in Germany and the NICE coverage. A possible explanation of different quality indicators could be the different approaches are their development. To prove the efficacy of CG recommendations, quality indicators should be of sophisticated methodological goodness. The clinimetric properties of all quality indicators recommended in CGs of NICE and in Germany are unclear. This constitutes a major risk for inadequate conclusions concerning the quality of care. The measurement of quality of care for specific medical conditions requires quality indicators of a high methodological quality. An internationally accepted methodological approach does not exist at present but is necessary for the derivation of internationally comparable quality indicators in health care. The idea of Core Quality Indicator Sets could support quality measurement as it represents an opportunity to consistently measure relevant indicators of quality of care.

Acknowledgments

The authors thank Phil Alderson (National Institute for Health and Care Excellence [NICE], UK) for his advice and help regarding clinical guideline development and quality measurement processes at NICE.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Kohn L, Corrigan J, Donaldson M. *To Err Is Human: Building a Safer Health System*. Washington, DC: Institute of Medicine; 2000.

2. Schmitt J, Petzold T, Nellessen-Martens G, Pfaff H. [Prioritization and consentation of criteria for the appraisal, funding and evaluation of projects from the German Innovationsfonds: a multi-perspective Delphi study]. *Gesundheitswesen*. 2015;77:570-579.
3. Mansky T, Nimptsch U. German Inpatient Quality Indicators (G-IQI)—Qualitätsmessung in der Initiative Qualitätsmedizin. In: Kühlen R, Rink O, Zacher J, eds. *Jahrbuch Qualitätsmedizin 2010*. Berlin, Germany: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2011:17-31.
4. Busse R, Blümel M. Payment systems to improve quality, efficiency, and care coordination for chronically ill patients—a framework and country examples. In: Mas N, Wisbaum W, eds. *The Triple Aim for the Future of Health Care FUNCAS Social and Economic Studies*. Madrid, Spain: Funcas; 2015:83-100.
5. National Institute for Health Research. *Health Services and Delivery Research (HS&DR) Programme*. Southampton, England: National Institute for Health Research. <http://www.nets.nihr.ac.uk/programmes/hsdr>. Published 2015. Accessed February 14, 2018.
6. Eberlein-Gonska M. [What is evidence-based concerning quality management? Thoughts about an apparently simple question]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2011;54(2):148-153.
7. Kopp IB. Von Leitlinien zur Qualitätssicherung. *Bundesgesundheitsbl*. 2011;54:160-165.
8. Field MJLK. *Clinical Practice Guidelines: Directions for a New Program*. Washington, DC: Institute of Medicine; 1990.
9. Muche-Borowski C, Kopp I. Guideline development. *Z Herz-Thorax-Gefäßchir*. 2011;25:217-223.
10. National Institute for Health Care Excellence. *Who we are*. Date unknown. <https://www.nice.org.uk/about/who-we-are>. Accessed March 3, 2017.
11. Mainz J. Defining and classifying clinical indicators for quality improvement. *Int J Qual Health Care*. 2003;15(6):523-530.
12. Petzold T, Deckert S, Eberlein-Gonska M, Albrecht D, Schmitt J. Evidenzbasierte Qualitätsmessung als Voraussetzung für Value-based Healthcare. *Monitor Versorgungsforschung*. 2015;8(3):49-54.
13. Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. *Int J Qual Health Care*. 2003;15(suppl 1):i5-11.
14. Donabedian A. Evaluating the quality of medical care. *Milbank Q*. 2005;83(4):691-729.
15. Braspenning J, Hermens R, Calsbeek H, Westert G, Campbell S, Grol R. Quality and safety of care: the role of indicators. In: Grol R, Wensing M, Eccles M, Davis D, eds. *Improving Patient Care*. 2. London, England: BMJ Books; 2013:116-135.
16. De Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine—A Practical Guide*. New York, NY: Cambridge University Press; 2011.
17. Schmitt J, Langan S, Williams HC. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol*. 2007;120(6):1389-1398.
18. Porter ME. A strategy for health care reform—toward a value-based system. *N Engl J Med*. 2009;361:109-112.
19. Mirnezami R, Nicholson J, Darzi A. Preparing for precision medicine. *N Engl J Med*. 2012;366:489-491.
20. Schmitt J, Petzold T, Deckert S, Eberlein-Gonska M, Neugebauer EA. [Recommendations for quality indicators in German S3 guidelines: a critical appraisal]. *Gesundheitswesen*. 2014;76(12):819-826.
21. Follmann M, Nothacker M, Klinkhammer-Schalke M, Kopp I. *Entwicklung von Leitlinien basierten Qualitätsindikatoren Methodenpapier für das Leitlinienprogramm Onkologie*. <http://www.leitlinienprogramm-onkologie.de/OL/leitlinien.html>. Published 2013. Accessed February 14, 2018.
22. Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). *Programm für Nationale VersorgungsLeitlinien*. Methoden-Report. Bundesärztekammer, Kassenärztliche Bundesvereinigung, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. http://www.versorgungsleitlinien.de/methodik/pdf/nvl_methode_4.auf.pdf. Published 2010. Accessed February 14, 2018.
23. Kotter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators—a systematic review. *Implement Sci*. 2012;7:21.
24. National Institute for Health Care Excellence. *Developing NICE guidelines: the manual*. Manchester, England: National Institute for Health and Care Excellence; 2014.
25. Reiter A, Fischer B, Kotting J, Geraedts M, Jackel WH, Dobler K. [QUALIFY—a tool for assessing quality indicators]. *Z Arztl Fortbild Qualitatssich*. 2007;101(10):683-688.
26. National Institute for Health Care Excellence. *Standards and indicators*. National Institute for Health and Care Excellence. <http://www.nice.org.uk/standards-and-indicators>. Published 2015. Accessed February 14, 2018.
27. National Institute for Health Care Excellence. *Indicator process guide*. <https://www.nice.org.uk/media/default/Get-involved/Meetings-In-Public/indicator-advisory-committee/ioc-process-guide.pdf>. Published December 2017. Accessed March 13, 2017.
28. National Institute for Health Care Excellence. *Selecting and prioritising guideline and quality standard topics*. National Institute for Health and Care Excellence. 2015; <https://www.nice.org.uk/About/What-we-do/Our-programmes/NICE-guidance/NICE-guidelines/Selecting-and-prioritising-guideline-and-quality-standard-topics>. Published 2015. Accessed February 14, 2018.
29. Bennett B, Coventry E, Greenway N, Minchin M. The NICE process for developing quality standards and indicators. *Z Evid Fortbild Qual Gesundhwes*. 2014;108(8-9):481-486.
30. Joint Commission on Accreditation of Healthcare Organizations (JCAHO). *Primer on Indicator Development and Application: Measuring Quality in Health Care*. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations; 1990.
31. Bundesärztekammer Kassenärztliche Bundesvereinigung Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. *Nationale VersorgungsLeitlinie Unipolare Depression*. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. <http://www.awmf.org/uploads/tx-szleitlinien/nvl-0051-S3-Unipolare-Depression-2012-01.pdf>. Published 2012. Accessed November 30, 2013.
32. National Institute for Health Care Excellence. *Quality Standard: Depression in Adults*. London, England: National Institute for Health and Care Excellence; 2011.
33. Deutsche Dermatologische Gesellschaft. *S3-Leitlinie zur Therapie der Psoriasis vulgaris*. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. <http://www.awmf.org/uploads/tx-szleitlinien/013-0011-S3->

- Psoriasis-vulgaris-Therapie-01.pdf. Published 2011. Accessed November 30, 2013.
34. Feldman SR, Fleischer AB Jr, Reboussin DM, et al. The self-administered psoriasis area and severity index is valid and reliable. *J Invest Dermatol*. 1996;106(1):183-186.
 35. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)—a simple practical measure for routine clinical use. *Clin Exp Dermatol*. 1994;19(3):210-216.
 36. National Institute for Health Care Excellence. *Quality standard: psoriasis*. London, England: National Institute for Health and Care Excellence; 2013.
 37. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care*. 2002;11:358-364.
 38. Marshall M, Campbell SM. Introduction to quality assessment in general practice. In: Marshall M, Campbell SM, Hacker J, Roland MO, eds. *Quality Indicators for General Practice: A Practical Guide for Health Professionals and Managers*. London, England: Royal Society of Medicine; 2002:1-6.
 39. Department of Health. *New care objectives for improving our health and healthcare*. Date unknown. <https://www.gov.uk/government/news/new-care-objectives-for-improving-our-health-and-healthcare>. Accessed March 13, 2017.
 40. Schmitt J, Petzold T, Eberlein-Gonska M, Neugebauer EA. [Requirements for quality indicators. The relevance of current developments in outcomes research for quality management]. *Z Evid Fortbild Qual Gesundheitswes*. 2013;107(8):516-522.
 41. Schmitt J, Langan S, Stamm T, Williams HC. Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. *J Invest Dermatol*. 2011;131(3):623-630.
 42. Comet Initiative. *Core outcome measurement database 2014*. Date unknown. <http://www.comet-initiative.org/studies/search>. Accessed February 14, 2018.
 43. Williamson P, Altman D, Blazeby J, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. *J Health Serv Res Policy*. 2012;17(1):1-2.
 44. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol*. 1998;25(2):198-199.
 45. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.
 46. Comet Initiative. *A walk through core outcome sets—useful references for core outcome set developers*. Liverpool, Spain: Comet Initiative. <http://www.comet-initiative.org/resources/coreresourcepack>. Published 2016. Accessed February 14, 2018.
 47. Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012;13(1):1-8.
 48. Schmitt JAC, Spuls PI, Thomas KS, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol*. 2015;135:24-30.
 49. Mitchell LG, Goldenberg NA, Male C, et al. Definition of clinical efficacy and safety outcomes for clinical trials in deep venous thrombosis and pulmonary embolism in children. *J Thromb Haemost*. 2011;9(9):1856-1858.
 50. Sinha I, Gallagher R, Williamson PR, Smyth RL. Development of a core outcome set for clinical trials in childhood asthma—a survey of clinicians, parents and young people. *Trials*. 2012;13:103.
 51. Mullis R, Barber J, Lewis M, Hay E. ICF core sets for low back pain: do they include what matters to patients? *J Rehabil Med*. 2007;39(5):353-357.
 52. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine*. 1998;23(18):2003-2013.
 53. Brouwers MKM, Browman GP, Cluzeau F, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *Can Med Assoc J*. 2010;182:E839-E842.
 54. Ärztliches Zentrum für Qualität in der Medizin. *Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI): Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, Ärztliches Zentrum für Qualität in der Medizin*. <http://www.leitlinien.de/leitlinienmethodik/leitlinienbewertung/delbi>. Published 2008. Accessed February 14, 2018.