

UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Institut und Poliklinik für Medizinische Psychologie

Prof. Dr. med. Dr. phil. Martin Härter

DEVELOPMENT AND PSYCHOMETRIC TESTING OF A PATIENT-REPORTED EXPERIENCE MEASURES (PREMS) CORE SET TO ASSESS PATIENT-CENTEREDNESS (PC)

Dissertation

zur Erlangung des Doktorgrades PhD
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:

Eva Christalle
aus Berlin

Hamburg 2025

Betreuer:in / Gutachter:in der Dissertation: Prof. Dr. Isabelle Scholl

Gutachter:in der Dissertation: Prof. Dr. Martin Scherer

Vorsitz der Prüfungskommission: Prof. Dr. Isabelle Scholl

Mitglied der Prüfungskommission: Prof. Dr. Martin Scherer

Mitglied der Prüfungskommission: PD Dr. Christine Blome

Datum der mündlichen Prüfung: 16.02.2026

TABLE OF CONTENTS

Summary	4
Zusammenfassung.....	6
1. Introduction.....	9
1.1. Why patient-centeredness matters	10
1.2. Definition of patient-centeredness	11
1.3. Measuring patient-centeredness	12
1.4. Patient-reported experience measures (PREMs)	13
1.5. Guidelines for developing patient-reported measures	15
1.6. Prioritization of content validity.....	16
1.7. Limitations of existing PREMs	17
1.8. Research gap	18
2. Aims and research questions	19
3. Context of this dissertation	20
3.1. Underlying scientific study	20
3.2. Key Psychometric Terms.....	22
4. Summary of publications.....	23
4.1. Publication 1.....	23
4.2. Publication 2.....	24
4.3. Publication 3.....	25
4.4. Publication 4.....	26
4.5. Publication 5.....	27

5.	Discussion	28
5.1.	Summary of results	28
5.2.	Comprehensive discussion	29
	Implications	35
6.	Conclusions	39
7.	List of abbreviations	40
8.	References	41
9.	Publications	47
9.1.	Publication 1.....	47
9.2.	Publication 2.....	56
9.3.	Publication 3.....	65
9.4.	Publication 4.....	76
9.5.	Publication 5.....	91
10.	Own contributions to the publications.....	103
11.	Danksagung	105
12.	Curriculum vitae	106
13.	Eidesstattliche Erklärung	109

SUMMARY

Patient-centeredness is increasingly recognized as a key indicator of healthcare quality. To improve patient-centeredness effectively, psychometrically sound measurement of its degree is essential. Patient-reported experience measures (PREMs) are particularly suited to capture the quality of care from the patient's perspective. However, a generic PREM that comprehensively assesses patient-centeredness is still lacking. This dissertation addresses this gap through five publications pursuing two aims: (1) to develop a generic German-language PREM comprehensively assessing patient-centeredness - the Experienced Patient-Centeredness Questionnaire (EPAT) - and (2) to select items for both modular and short versions of the EPAT and evaluate their psychometric properties.

Publication 1 is a study protocol outlining the study design that provided the framework for this dissertation and guided the investigation of all research questions.

Publication 2 addresses the definition of the construct patient-centeredness from the patients' perspective. In a Delphi study with $n = 226$ patients, all 15 dimensions of the integrative model of patient-centeredness were rated as highly relevant, and a sixteenth dimension, "patient safety," was added.

Publication 3 details the item development process based on a literature review including 48 PREMs, six focus groups with $n = 40$ patients, and $n = 10$ expert interviews. Relevance and comprehensiveness were ensured through a relevance rating with 32 experts, and comprehensibility was confirmed in 34 cognitive interviews with patients. Collectively, these steps demonstrated the content validity of the developed items.

Publication 4 describes a cross-sectional study evaluating the psychometric properties of the EPAT in a sample of $n = 2,024$ patients. Item characteristics and content considerations were used to derive the EPAT-64 - a modular version assessing all 16 dimensions of patient-centeredness with four items each. For the final version, structural validity, construct validity, and reliability were evaluated, demonstrating robust psychometric properties, although some items exhibited ceiling effects and high "does not concern me" response rates.

Publication 5 describes how the same dataset was used to select items for the EPAT-16, a short version with one item representing each of the 16 dimensions. It demonstrated good construct validity and reliability, with some limitations in structural validity and measurement invariance.

In summary, this dissertation presents the first comprehensive and generic PREM for assessing patient-centeredness and provides a best-practice example of questionnaire development that integrates theory-driven and data-driven methods to prioritize content validity throughout the entire process. The EPAT's modular and short forms offer flexible tools for both research and routine healthcare, supporting quality monitoring, intervention evaluation, and public reporting. Future research should further examine its responsiveness, test measurement invariance across diverse populations, and explore strategies for effective implementation in clinical practice.

ZUSAMMENFASSUNG

Patient:innenorientierung wird zunehmend als ein zentraler Indikator für die Qualität der Gesundheitsversorgung diskutiert. Um Patient:innenorientierung gezielt zu verbessern, ist eine psychometrisch robuste Messung unerlässlich. Patient:innenberichtete Erfahrungsmaße (PREMs) eignen sich besonders gut, um die Qualität der Versorgung aus der Perspektive der Patient:innen zu erfassen. Ein generisches PREM, das Patient:innenorientierung umfassend erfasst, fehlt jedoch bislang. Diese Dissertation adressiert diese Lücke durch fünf Publikationen, die zwei Ziele verfolgen: (1) die Entwicklung eines generischen deutschsprachigen PREMs zur umfassenden Messung von Patient:innenorientierung - den Experienced Patient-Centeredness Questionnaire (EPAT) - und (2) die Auswahl von Items für eine Modulversion und eine Kurzversion des EPAT und die Evaluation ihrer psychometrischen Eigenschaften.

Publikation 1 ist ein Studienprotokoll, das die Studie beschreibt, die den Kontext für diese Dissertation stellte und die Untersuchung aller Forschungsfragen leitete.

Publikation 2 untersucht die Definition des Konstrukts Patient:innenorientierung aus der Perspektive von Patient:innen. In einer Delphi-Studie mit $n = 226$ Patient:innen wurden alle 15 Dimensionen des integrativen Modells für Patient:innenorientierung als hochrelevant bewertet, und eine sechzehnte Dimension, „Patient:innensicherheit“, wurde ergänzt.

Publikation 3 beschreibt den Item-Entwicklungsprozess auf Basis einer Literaturrecherche mit 48 PREMs, sechs Fokusgruppen mit $n = 40$ Patient:innen sowie $n = 10$ Experteninterviews. Relevanz und Vollständigkeit wurden in einem Relevanzrating mit 32 Expert:innen sichergestellt, und die Verständlichkeit in 34 kognitiven Interviews mit Patient:innen überprüft. Zusammen zeigten diese Schritte, dass die entwickelten Items Inhaltsvalidität haben.

Publikation 4 beschreibt eine Querschnittsstudie zur psychometrischen Evaluation des EPAT in einer Stichprobe von $n = 2.024$ Patient:innen. Itemcharakteristika und inhaltliche Überlegungen wurden kombiniert, um Item für den EPAT-64 auszuwählen - eine modulare Version, die alle 16 Dimensionen der Patient:innenorientierung mit jeweils vier Items erfasst. Für die finale Version wurden strukturelle Validität, Konstruktvalidität und Reliabilität untersucht. Die Ergebnisse zeigten gute psychometrische Eigenschaften, obwohl einige Items Deckeneffekte und hohe „betrifft mich nicht“-Antwortraten aufwiesen.

Publikation 5 beschreibt, wie derselbe Datensatz zur Auswahl von Items für den EPAT-16 verwendet wurde - eine Kurzversion mit jeweils einem Item pro Dimension. Der EPAT-16 zeigte eine gute Konstruktvalidität und Reliabilität, mit einigen Einschränkungen in der strukturellen Validität und Messinvarianz.

Zusammenfassend präsentiert diese Dissertation das erste umfassende und generische PREM zur Messung von Patient:innenorientierung und liefert ein Best-Practice-Beispiel für die Fragebogenentwicklung, das theorie- und datengestützte Ansätze kombiniert, um die Inhaltsvalidität während des gesamten Prozesses zu priorisieren. Die Modulversion und Kurzversion des EPAT bieten gut anpassbare Instrumente für Forschung und Routineversorgung, z.B. für Qualitätsmonitoring, Evaluation von Interventionen und öffentliche Berichte von Ergebnissen. Zukünftige Forschung sollte die Änderungssensitivität untersuchen, die Messinvarianz über verschiedene Populationen hinweg testen und Strategien für eine effektive Implementierung in der Routineversorgung erforschen.

1. INTRODUCTION

“Quality in health and medical care [...] has two dimensions: one is objective, technical; the other is subjective and qualitative. However dazzling the technological achievements of medical science over that last fifty years, the patient’s experience of illness and medical care is at the heart of the first purpose of clinical medicine – to relieve human suffering.” [1] (page 2)

With this statement, Gerteis and her colleagues introduced the concept of patient-centeredness in their groundbreaking book *“Through the patient’s eyes: understanding and promoting patient-centered care” [1]*, a publication funded by the Picker/Commonwealth Program for Patient-Centered Care, which ultimately led to the founding of The Picker Institute [2]. Globally, increasing attention has been directed toward understanding and advancing patient-centeredness as a cornerstone of healthcare quality. This is reflected in seminal works such as the Institute of Medicine’s publication *Crossing the Quality Chasm* [3], as well as in key documents by the World Health Organization [4] and many others [5, 6]. In Germany, patient-centeredness has also gained importance in legislation, research, and education [7] for example, with the Patients’ Rights Law of 2013 [8], the National Cancer Plan launched in 2008 [9] and a funding program initiated by the German Ministry of Health in 2001 [10]. This was underlined when the German Federal Minister of Health declared patient-centeredness a guiding principle of the national healthcare system in 2019 [11]. However, to effectively promote patient-centeredness, the patient must be placed at the center of healthcare improvement strategies as well [12]. This in turn requires assessment of patient experiences so that they can inform the design of such strategies [12].

This introduction begins by outlining why patient-centeredness matters and how it is defined. Then, the measurement of patient-centeredness will be discussed, in particular how patient-centeredness can be assessed and why measuring patient-centeredness is essential. A review of the existing literature on the development of such measures follows. Finally, the research gap addressed in this dissertation will be derived.

1.1. Why patient-centeredness matters

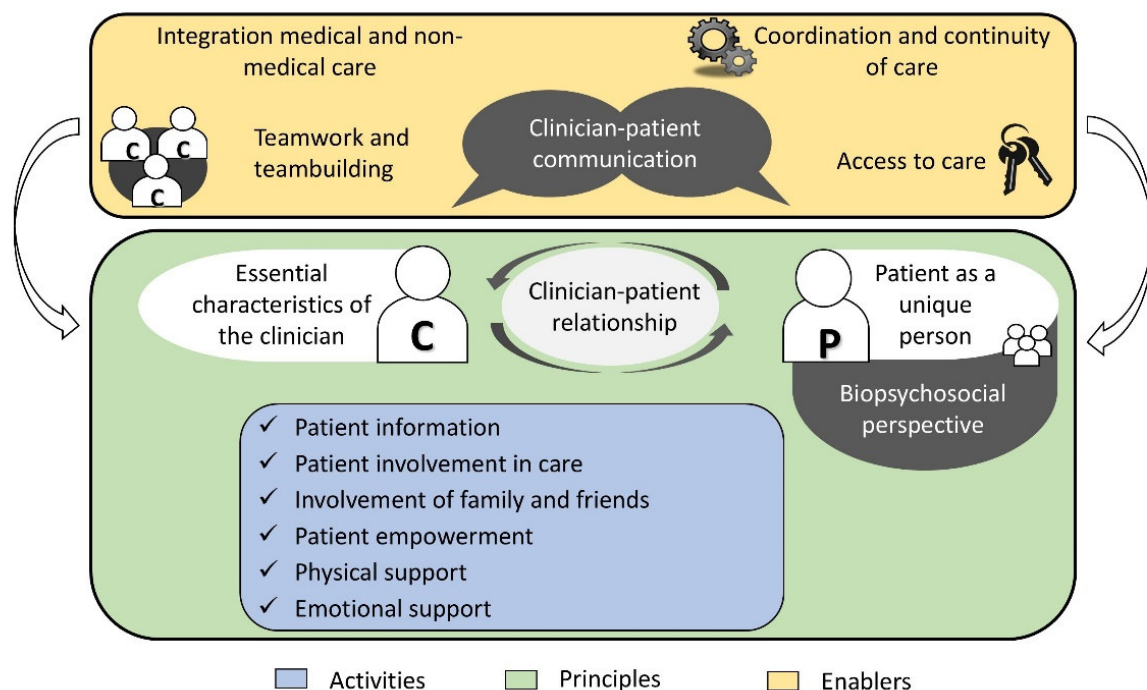
From a philosophical and ethical perspective, adopting patient-centeredness is considered “the right thing to do” [13, 14]. From a deontological point of view, this is rooted in the intrinsic moral value of behaviors and attitudes that respect patients’ autonomy, dignity, values, and preferences [13]. Such values are reflected in the principles of biomedical ethics, particularly in the respect for patient autonomy [15] - an idea well captured by the guiding statement of many patient-centeredness initiatives: *“Nothing about me without me”* [16] (title). Moreover, the fundamental goal of healthcare - to improve patients’ health and well-being - can only be fully achieved by co-designing care with patients themselves, since only patients themselves can define what “health and well-being” means to them [12]. Along similar lines, evidence-based medicine relies on patient perspectives, as meaningful evidence depends on outcomes that reflect what matters most to patients [14]. In addition, the experiential expertise of patients deserves acknowledgment [17]. Healthcare professionals bring the clinical knowledge required for accurate diagnosis and treatment planning. However, patients - being the ones living with the condition - are experts in how symptoms evolve, how illness affects their daily lives, and what they need to enhance their quality of life [17]. A comprehensive understanding - as is essential for well-informed decision-making - can only emerge through the exchange of knowledge and expertise from both parties. [13].

Another ethical argument in favor of patient-centeredness comes from a consequentialist perspective, which claims that patient-centeredness should be adopted because it leads to better outcomes [13]. For example, often when patients do not adhere to their treatment, this is due to either not understanding the instructions or not having agreed to the treatment [12]. Hence, it is not surprising that associations between patient-centeredness and treatment adherence were found [18]. Further, patient-centered communication correlates with better recovery, improved mental health, and fewer diagnostic tests and referrals [19]. In a systematic review, relationships of patient-centeredness with patient satisfaction, well-being and self-management were found [18]. However, findings on clinical outcomes were mixed [18]. This is echoed in a Cochrane review on the effects of patient-centeredness training programs, which concluded that while there is some evidence for positive effects on patient satisfaction, the evidence for improvements in health status remains limited [20]. The variability in results is often attributed to differences in intervention design and outcome selection across studies [20], but also to inconsistencies in how patient-centeredness is defined and measured [21, 22].

1.2. Definition of patient-centeredness

To advance both research and practice in patient-centeredness, a clear and consistent definition of the concept is essential. This is particularly important for the measurement of patient-centeredness, as defining the construct to be measured is the first critical step in developing valid and reliable assessment tools [23].

To address this gap, Scholl and her colleagues [24] conducted a systematic review in 2012, identifying 417 publications that included definitions of patient-centeredness. They performed a content analysis of these definitions and synthesized them into a unified framework - the integrative model of patient-centeredness - which captures the core elements across existing definitions. The resulting model comprises 15 interrelated dimensions that predominantly focus on the interaction between healthcare professionals and patients. These dimensions are grouped into three categories: activities, principles, and enablers (see Figure 1). The integrative model of patient-centeredness was further evaluated through a Delphi study involving 105 experts - including researchers, patient representatives, clinicians, and quality managers [25]. Of the 15 dimensions, eleven were rated as highly relevant by the participating experts [25].



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). To view a copy of the license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/>
Authors: Isabelle Scholl, Martin Härter, Jörg Dirmaier, Stefan Zeh, Eva Christalle, Pola Hahlweg, & Jödis Zill (University Medical Center, Hamburg, Germany)

Figure 1: Integrative model of patient-centeredness [24]

While patient representatives participated in the Delphi process [25], it remains an open question whether lay patients - those without representative roles - would identify the same dimensions as relevant. Given that a core principle of patient-centeredness is to incorporate patients' preferences and experiential knowledge, its definition should likewise reflect what truly matters to patients by actively involving them in the definition process.

1.3. Measuring patient-centeredness

Measurement is the cornerstone of evidence-based medicine [26] and is central to health research [27]. Only with valid measures can latent constructs, which are not directly observable - such as the degree of patient-centeredness - be quantified [28, 29]. Quantification, in turn, is vital for generating evidence [26], for example by assessing the status quo, investigating relationships with other constructs, monitoring change over time, or evaluating the effectiveness of interventions [14]. Therefore, delivering high-quality healthcare depends on the use of psychometrically sound measurements [26] and fostering patient-centeredness requires continuous measurement through well-designed surveys [12]. Or as Kilbourne and her colleagues summarized: "We cannot improve what we cannot measure." [30] (page 30).

Patient-centeredness can be assessed from various perspectives - by external observers, healthcare professionals, or patients themselves [31]. Observers may use rating scales or coding systems [21], which can be applied to either real-life encounters of healthcare professionals and patients or to encounters involving standardized patients [31]. While standardized patients offer high comparability across assessments, real-life encounters may have higher external validity as they reflect everyday clinical practice. However, observer-based assessments are resource-intensive, requiring specialized training, and they carry a risk of the Hawthorne effects - that is, changes in behavior due to the awareness of being observed [31]. Surveys offer a more standardized and cost-effective method for assessing patient-centeredness on a large scale [32]. These can be administered to healthcare professionals, who then self-report on their behavior. However, self-assessments are susceptible to social desirability bias [21, 31] and frequently show little correlation with patient-reported data [31].

Another approach is to ask patients directly about the extent to which their care was patient-centered. A possible limitation here is that patients may overestimate certain aspects of their care - such as the degree of involvement in decision-making processes [33]. Further, patient-reported measures must be carefully designed to differentiate between reporting actual occurrences and making subjective evaluations [14]. Despite these challenges, patient-reported assessments have proven to be the most predictive of outcomes [34]. Certain elements of patient-centeredness - such as whether patients felt respected or whether information was delivered in an understandable way - can only be evaluated by patients themselves [12, 35]. Tzelepis and her colleagues [36] argue that patient-reported measures are the most appropriate tools for assessing patient-centeredness, since only patients can reliably assess whether their values, needs, and preferences were taken into account.

1.4. Patient-reported experience measures (PREMs)

In line with this rationale, numerous patient-reported experience measures (PREMs) have been developed. PREMs focus on what patients experience during care - that is, what happened and how it happened [32, 37]. Questions focus on whether or to what extent patients experienced specific care processes [32]. In doing so, PREMs aim to capture relatively objective aspects of healthcare quality from the patient's perspective. Although patient experience and patient satisfaction are related concepts, it is critical to distinguish between them [37]. Patient satisfaction measures involve subjective judgments and are often influenced by prior expectations [37]. PREMs are process indicators evaluating interpersonal aspects of care quality, whereas patient satisfaction measures are outcome measures that reflects the result of those experiences [38]. For example, a satisfaction measure might ask, "How satisfied are you with the healthcare professional's communication?" In contrast, a PREM would focus on explicit behaviors, such as, "Did the healthcare professional make eye contact?" or "Did the healthcare professional use terms you could easily understand?" By phrasing items around such specific behaviors, it becomes easier to derive which aspects can be targeted by interventions to improve patient-centeredness.

PREMs serve a variety of purposes: they are used for benchmarking healthcare institutions [39], informing quality improvement initiatives [36], evaluating the effectiveness of interventions [32], and supporting public reporting and rankings [32, 39]. In several countries, including the USA and the UK, PREM results inform pay-for-performance schemes, where better patient experiences are linked to higher reimbursements for healthcare facilities [40]. Figure 2 illustrates various applications of PREMs identified in a scoping review by Gilmore and his colleagues [35]. These applications span all three levels of the healthcare system (macro-level, meso-level and micro-level), with most studies identified focus on improving quality of care at the micro-level [35].

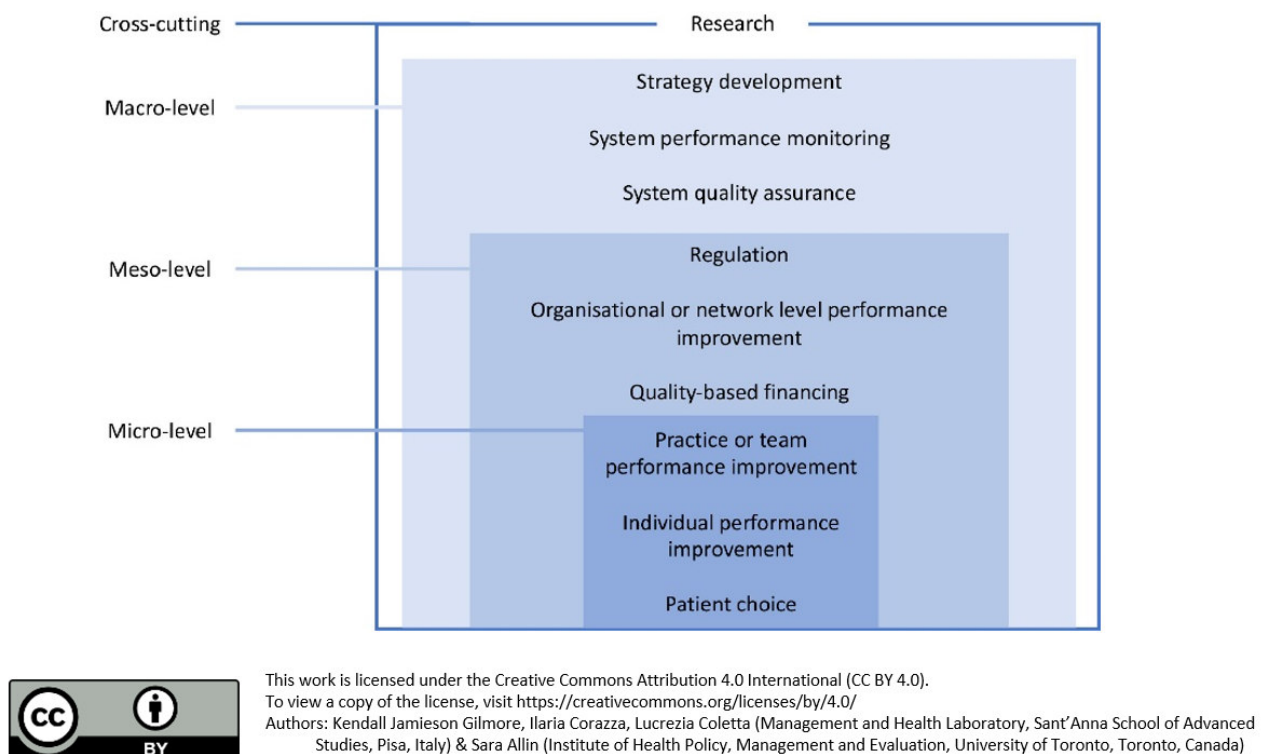


Figure 2: Uses of patient-reported experience measures by Gilmore et al. [35]

Reviews exploring the relationship between patient experiences and outcomes have found positive associations with clinical effectiveness, patient safety [41, 42], treatment adherence, and reduced healthcare utilization [42]. In addition, a randomized controlled trial, written feedback on PREM results alone did not yield improvements over a control group; however, when the feedback was supplemented by a meeting to discuss the results, positive effects were observed [43]. Further in an observational study, patient experience scores were positively correlated with how well-informed healthcare professionals were about prior PREM results [44].

1.5. Guidelines for developing patient-reported measures

Taken together, PREMs are a promising method for assessing and promoting patient-centeredness. However, they require careful development and evaluation to ensure that the results are clinically meaningful and interpretable [29]. Although there is limited literature specifically focused on developing PREMs, numerous guidelines exist for the development and validation of patient-reported outcome measures (PROMs) and other health-related instruments (e.g. [23, 45]). The primary distinction between PROMs and PREMs lies in what they measure: PROMs assess outcomes - the results of healthcare - whereas PREMs assess processes - what happens during healthcare [39]. Aside from this distinction, both are self-reported questionnaires designed to capture patients' perspectives on health-related matters. Therefore, it is reasonable to generalize PROM development guidelines for the development of PREMs as well.

In a 2023 scoping review, Jung and his colleagues [23] identified 40 guidelines related to PROM development – including guidelines that are applicable to health measurements in general. A guideline recommended by Jung et al. for its comprehensiveness and described as particularly influential and highly cited [23] is the checklist developed by the CONsensus-based Standards for the selection of health status Measurement Instruments (COSMIN) initiative [46]. In addition to the checklist, the COSMIN initiative has provided a taxonomy [47], a reporting guideline [48] and a range of additional tools available through their website (www.cosmin.nl). The COSMIN taxonomy represents a major milestone in the field of healthcare measurement, addressing previous inconsistencies in terminology and definitions that had hampered the comparability of psychometric properties across instruments [47]. These definitions were developed through a scientifically rigorous Delphi study involving 43 international experts from diverse fields [47].

However, COSMIN primarily focuses on three key psychometric properties at the level of the overall questionnaire: validity, reliability, and responsiveness. It does not specifically address the characteristics of individual items, which are critical for item selection. Jung et al. [23] defined eight steps for questionnaire development: (1) Definition of the construct, (2) Choice of measurement method, (3) Selecting and formulating items, (4) Scoring issues, (5) Elaboration of the measurement model, (6) Pilot-test, (7) Field test (item reduction), and (8) Measurement properties of the final form. They recommended three guidelines for comprehensively covering all eight steps [23]: Two books by Vet et al. [27] and Streiner et al. [49], as well as a peer-reviewed paper by Swan et al. [26]. Further, comprehensive guidance specifically on item selection is provided by Boateng et al [45].

1.6. Prioritization of content validity

Among psychometric properties, the COSMIN initiative emphasizes a clear hierarchy: “Content validity is the most important measurement property of a patient-reported outcome measure (PROM) and the most challenging to assess.” [50] (page 1159). They define content validity as the degree to which all items are relevant, comprehensive, and comprehensible with respect to the target group, the construct to be measured, and the intended use of the measure [50]. Content validity is considered foundational as it influences all other measurement properties - e.g., irrelevant items may impair structural validity, and incomplete coverage can limit responsiveness [50]. To support rigorous assessment, COSMIN developed quality standards for evaluating content validity through a Delphi study involving 159 international experts [50]. Nonetheless, these standards have two limitations, which hamper usability for questionnaire development: (1) The COSMIN guideline remains vague regarding methods, often referring only to a “widely recognized or well-justified method”. This ambiguity increases subjectivity and offers limited practical guidance to researchers [50]. (2) While the checklist evaluates the quality of qualitative research, it does not assess how items are derived from qualitative data. Hence, it provides no guidance on how qualitative data can be analyzed in a way that ensures the patients’ perspective is well reflected in the items, while also taking into account the definition and aim of the target construct. Similarly, other guidelines consistently recommend incorporating qualitative data from the target population during item development to ensure content validity (e.g. [45, 51]); however, they do not specify how such data should be analyzed or translated into questionnaire items. This gap results in a lack of transparency in reporting, as evidenced by a review by Ricci and her colleagues [52] on the use of qualitative data in questionnaire development. Among the 371 studies included, 57% did not report how qualitative data were analyzed [52].

A notable strength of the COSMIN guideline is its explicit requirement to assess comprehensiveness after the final item set has been established. This is an advancement, as many guidelines recommend initial assessment followed by item reduction, without requiring a re-evaluation of content validity: Jung et al. [23] emphasize that psychometric properties should be evaluated on the final version, though this is not clearly reflected in the guidelines they recommend (e.g. [26, 27]). This is also evident in the review by Wiering and her colleagues [53]. They identified three key stages that should be informed by data collected from patients: (1) defining the outcome, (2) item generation, and (3) testing item comprehensibility. Their work underscores that content validity

requires patient involvement in these three stages of questionnaire development, yet it does not specify that content validity should also guide the item selection process. Interestingly, quite the contrast is true for literature about item selection for short form development. Here, content-driven approaches are explicitly encouraged to ensure that the content of the original instrument is adequately reflected in its short form (e.g. [54-56]).

Overall, a clearly documented best-practice example that incorporates these considerations and prioritizes content validity throughout all stages of item development and selection would provide valuable guidance for future questionnaire development.

1.7. Limitations of existing PREMs

There are several limitations regarding the current measurement of patient-centeredness with PREMs - both internationally and within the German healthcare system. Despite its importance, content validity is frequently underreported or insufficiently established. In a systematic review including all identified English-language publications on PREM development, Bull and her colleagues [39] found that only 37.5% of the included PREMs met criteria for good content validity. Further, none of the identified measures captured all dimensions of the integrative model of patient-centeredness [39], which again limits their content validity, as comprehensiveness is part of it. Similarly, a more recent review by Handley and her colleagues [57], which focused on PREMs assessing patient-centeredness in hospital settings, concluded that no existing tool provides a comprehensive assessment.

Finally, a systematic review by Mihaljevic and his colleagues [58] focused on German-language PREMs that assess patient-centeredness, compared the PREMs to the integrative model of patient-centeredness. They identified 12 PREMs that were either developed in or translated into German. Although all PREMs showed deficits in psychometric testing and results, six were recommended by the authors [58]. For generic measures, these are Hospital Consumer Assessment of Healthcare Providers and Systems survey (HCAHPS) [59], The Nordic Patient Experiences Questionnaire (NORPEQ) [60], The Picker Patient Experience Questionnaire (PPE-15) [61] and Patients' Experiences Across Health Care Sectors (PEACS) [62]. Notably, all of those are applicable to inpatient settings only, except for the PEACS, which is designed for both inpatient and outpatient settings [58]. The PEACS covers 13 of the 16 dimensions of the integrative model of patient-centeredness with 59 items and demonstrated good ratings for content validity (except comprehensiveness),

structural validity, and reliability [58]. However, data on internal consistency, measurement error, and hypothesis testing (construct validity) were lacking [58]. In conclusion, Mihaljevic and colleagues also found that none of the questionnaires provided a complete assessment [58].

Thus, no PREM - whether in the international literature or specifically within Germany - currently exists that comprehensively measures patient-centeredness and therefore demonstrates sufficient validity. The need to address this gap with a generic measure, suitable for comparisons across settings and diseases [29], is underscored by health policy and patient surveys: In 2018, the German Conference of Health Ministers (German: Gesundheitsministerkonferenz) [63] passed a resolution calling for the assessment of patient experiences as a tool for quality management across all healthcare facilities. These assessments were to be evaluated for health benefit and made publicly available as transparent and neutral patient information [11]. Nevertheless, as of 2022, German patients still reported limited access to information about the quality of care across healthcare institutions [64]. Notably, 82% of respondents indicated that the experiences of other patients were an important source of information when evaluating healthcare providers [64].

1.8. Research gap

As outlined above, measuring the degree of patient-centeredness is essential to advancing the widely supported goal of a more patient-centered healthcare system [12, 36]. In particular, well-designed, psychometrically sound PREMs hold significant potential to enhance healthcare quality and improve various outcomes [41-44]. To achieve this, content validity must be prioritized throughout the entire development process, including ensuring that the construct is comprehensively represented [50]. However, no existing PREM captures patient-centeredness in its entirety [39, 57, 58]. Specifically, there is a lack of a generic German-language PREM that can be used to respond to calls from both health policy and patients [63, 64]. As Mihaljevic et al. [58] concluded, *“developing a German-language PREM, that will cover all aspects of patient-centeredness according to Scholl et al. [...] will close important evidence gaps.”* (page 33).

To address these research gaps, the following aims and research questions were derived to guide this dissertation.

2. AIMS AND RESEARCH QUESTIONS

This dissertation pursues two aims: (1) to develop and (2) to psychometrically evaluate a PREM to assess patient-centeredness. These aims are explored through the following research questions (RQ):

Aim 1: To develop a comprehensive, German-language PREM for assessing patient-centeredness from the patients' perspective - the Experienced Patient-Centeredness Questionnaire (EPAT).

RQ1: What are the relevant dimensions of patient-centeredness from the patients' perspective?

RQ2: What items can be developed to assess these dimensions according to the international literature, patients and stakeholders?

RQ3: Do the developed items demonstrate content validity, including relevance, comprehensiveness, and comprehensibility for the target group and construct?

Aim 2: To select items for both a modular and short version of the EPAT and evaluate their psychometric properties.

RQ4: Which items show good psychometric properties and should be included in the final versions of the EPAT?

RQ5: Do the selected items demonstrate structural validity, construct validity and reliability?

3. CONTEXT OF THIS DISSERTATION

3.1. Underlying scientific study

The development and psychometric evaluation of the EPAT was conducted as part of the Assessment of Patient-Centeredness Through Patient-Reported Experience Measures (ASPIRED) study. ASPIRED was part of a junior healthcare research group funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung; BMBF). It was led by Prof. Dr. Isabelle Scholl at the Department of Medical Psychology at the University Medical Center Hamburg-Eppendorf (UKE) from October 2017 to September 2023.

The overarching aim of ASPIRED was to assess patient-centeredness from the patients' perspective. The study was carried out in three phases:

- Phase 1 focused on understanding how patients define patient-centeredness. To this end, the integrative model of patient-centeredness was translated into German lay language and then evaluated in a two-round Delphi study. Based on the results, a 16th dimension - patient safety - was added to the model.
- Phase 2 aimed to develop and psychometrically evaluate a generic PREM to assess patient-centeredness - the EPAT. This phase involved multiple methods, including a literature review, semi-structured interviews, focus groups, relevance ratings, cognitive interviews, and psychometric testing. The final instruments, the EPAT-64 and its short form, the EPAT-16, resulted from this phase.
- Phase 3 examined the feasibility of implementing the EPAT in routine care. Healthcare professionals were instructed to collect data using the EPAT on their own. Semi-structured interviews were then conducted, guided by the Consolidated Framework for Implementation Research (CFIR) [65], to explore their experiences with the implementation process.

An overview of the three study phases is provided in Figure 3.

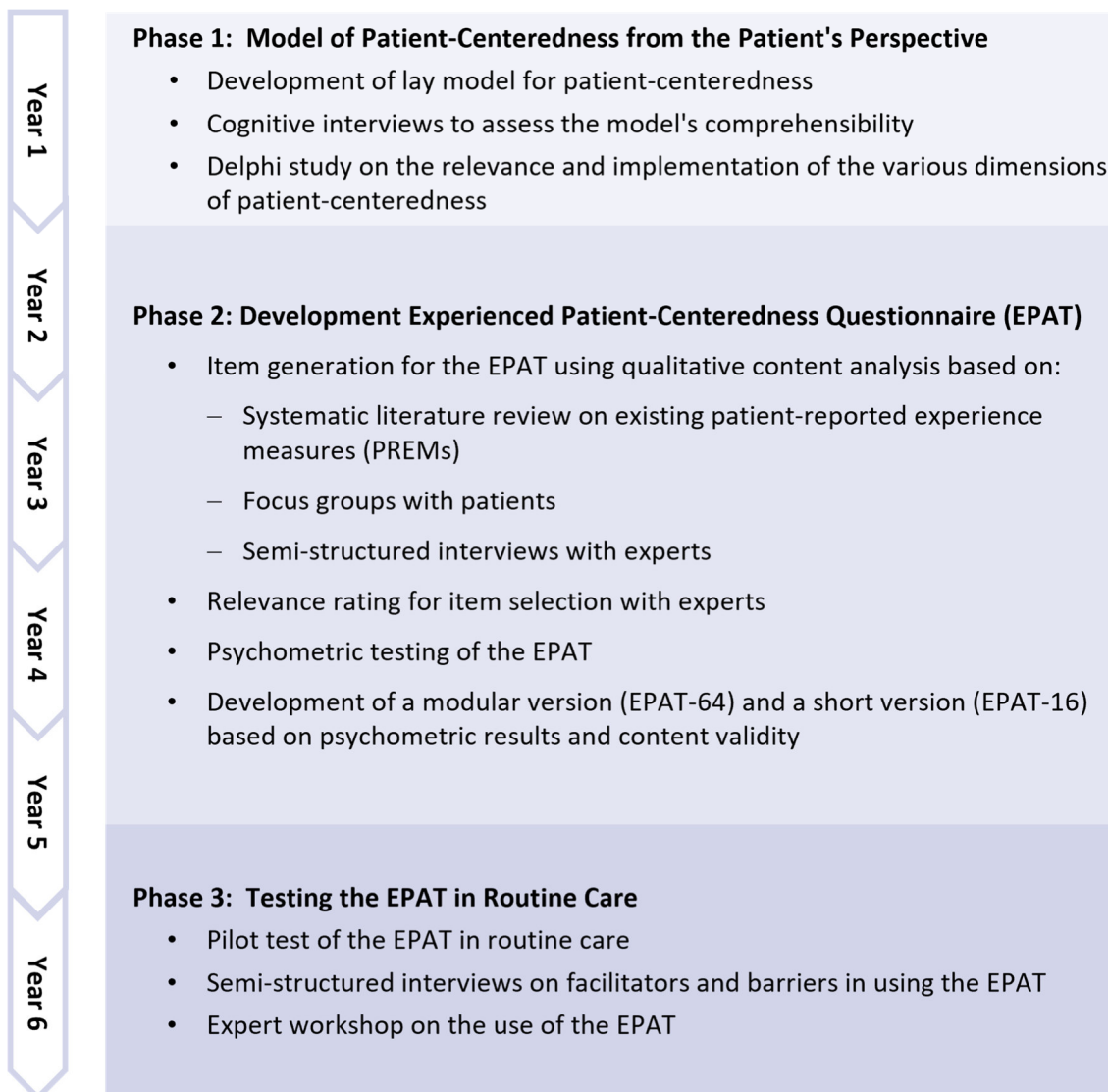


Figure 3: Overview phases of the ASPIRED study

3.2. Key Psychometric Terms

As psychometric terms are used and defined inconsistently [47], the terms applied throughout this dissertation are defined below. Table 1 summarizes psychometric properties related to the overall questionnaire, while Table 2 presents item-level characteristics used for item selection.

Table 1: Psychometric properties of the questionnaire

Psychometric property	Definition	Statistics/methods applied in this thesis
Reliability	Extent to which the measure is free from error [47]	
↳ Internal consistency	How strongly the items of a measure are interrelated [47]	Cronbachs's Alpha or McDonald's Omega
Validity	Whether the measure captures the intended construct [47]	
↳ Content validity	Whether the measure's content fully represents the construct [47] - including relevance, comprehensiveness, and comprehensibility of the measure regarding the construct and target population [50]	Relevance rating, cognitive interviews
↳ Construct validity	Whether the measure's scores behave as expected with other measures or groups [47]	Hypotheses testing
↳ Structural validity	Whether the scores adequately reflect the construct's dimensionality [47]	Confirmatory factor analysis and model fit indices
Responsiveness	Ability of the measure to detect change over time [47]	

Table 2: Item characteristics used for item selection

Item characteristic	Calculation/recommendations	Recommended range
Acceptability	Percentage of missing values, that is no reply at all [66]	-
Relevance	Percentage of 'does not concern me' replies	-
Item difficulty (assessing floor and ceiling effects)	Standardizing the mean to a range of 0 to 1 [67]	> 0.2 and < 0.8 [67]
Inter-Item-correlations	Correlation of items within a dimension [45]	> 0.3 [45]
Corrected item-total correlations	Correlation of an item with the sum score of the remaining items of that dimension [45]	> 0.3 [45]

4. SUMMARY OF PUBLICATIONS

This dissertation focusses on the first and second phase of the ASPIRED study. Further details on the overall study are provided in Publication 1, which outlines the study protocol. Publication 2 describes how a definition of patient-centeredness from the patients' view was derived. Publication 3 describes the development process of the EPAT. Publication 4 addresses the development of the EPAT-64 through item selection and psychometric evaluation, while Publication 5 focuses on the EPAT-16.

4.1. Publication 1

Reference: Christalle, E.*, Zeh, S.*, Hahlweg, P., Kriston, L., Härter, M., & Scholl, I. (2018). Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open*, 8(10), e025896.

(*Shared first authorship)

Aims: This study protocol outlines the Assessment of patient-centeredness through patient-reported experience measures (ASPIRED) study, which has three aims: (1) Complement the integrative model of patient-centeredness with the patients' perspective and assess the relevance of its dimensions, (2) to develop and psychometrically test a German PREM to assess patient-centeredness, and (3) to test the feasibility of implementing this PREM in routine care.

Methods and analysis: Using a mixed-methods design, this study will be carried out in three phases. In phase 1, a model of patient-centeredness will be translated into lay language, its comprehensibility tested, and then be evaluated by N=200 patients in a Delphi study regarding the relevance of its dimensions. Further, patients will have the possibility to add new content to the model if any important aspects are missing. In phase 2, items for a PREM will be developed based on a literature review, key informant interviews, and focus groups with patients. Items will be selected in a relevance rating, and then tested for comprehensibility using cognitive interviews. The preliminary questionnaire will be tested in a sample of N=2000 patients treated in inpatient and outpatient settings for cancer, cardiovascular diseases, mental health disorders, or musculoskeletal disorders. Items of the final questionnaire will be selected based on the results of this psychometric evaluation. In phase 3 the resulting questionnaire will be pilot tested in routine care, and semi-structured interviews will be used to explore experiences with its implementation. Additionally, an expert workshop will be conducted to explore strategies for facilitating implementation.

4.2. Publication 2

Reference: Zeh, S., Christalle, E., Hahlweg, P., Härter, M., & Scholl, I. (2019). Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ Open*, 9(12), e031741.

Relevant research questions: RQ1

Aims: This study aimed to evaluate the perceived relevance and current implementation of patient-centeredness from the perspective of chronically ill patients in Germany.

Methods: The integrative model of patient-centeredness was translated into lay language and its comprehensibility tested through cognitive interviews. A two-round Delphi study was conducted with patients recruited via community-based strategies across Germany. In round one, participants anonymously rated the relevance and implementation of the 15 patient-centeredness dimensions and provided free-text comments to suggest improvements or missing aspects. In round two, they received a summary of the initial results and repeated the ratings. Eligible participants were adults (18+ years) with one or more of the following chronic conditions: cancer, cardiovascular disease, mental illness, or musculoskeletal disorder.

Results: Of the 226 patients who completed the first round, 214 participated in the second. Based on their comments, a 16th dimension - "patient safety" - was added. All dimensions were rated as highly relevant but only partially implemented. The most dimensions identified as most relevant were "patient safety," "access to care," and "patient information."

Conclusion: This study is among the first to explore how patients define patient-centeredness. In a previous Delphi study with experts, only eleven dimensions of the integrative model of patient-centeredness were rated as relevant. In contrast, patients regarded all dimensions as highly relevant. Moreover, a sixteenth dimension, "patient safety", was added. While the original definition included this aspect within the dimension of "physical support", patients considered it so critical that it warranted recognition as an independent dimension, which they subsequently rated as the most relevant. These insights offer a valuable foundation for enhancing patient-centeredness and for developing a patient-reported experience measure that assesses all aspects of patient-centeredness relevant to patients.

4.3. Publication 3

Reference: Christalle, E., Zeh, S., Hahlweg, P., Kriston, L., Härter, M., Zill, J., & Scholl, I. (2022). Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—A best practice example for generating patient-reported measures from qualitative data. *Health Expectations*, 25(4), 1529-1538.

Relevant research questions: RQ2 and RQ3

Aim: This study aimed to develop and select items for the EPAT, a generic PREM that captures patient-centeredness from the perspective of adult patients in both inpatient and outpatient settings across Germany. Additionally, it seeks to establish a best-practice model for developing patient-reported measures based on qualitative data.

Methods: A mixed-methods design was used across three phases: (1) preparation, (2) item generation, and (3) item selection including content validity testing. Items were developed based on the previously developed definition of patient-centeredness, which guided qualitative content analysis of data from focus groups, key informant interviews, and a literature review. Subsequently, items were complemented and selected in a relevance rating and comprehensibility tested in cognitive interviews. Participants included patients with chronic conditions (cancer, cardiovascular diseases, mental disorders, musculoskeletal disorders) and healthcare experts (clinicians, researchers, and patient representatives).

Results: Data from six focus groups (N=40 patients), ten expert interviews, and 48 identified PREMs from the literature were analyzed. A total of 152 initial items were included in a relevance rating (N=32 experts), resulting in 125 items covering 16 dimensions. Based on 34 cognitive interviews, only minor revisions were needed to ensure comprehensibility.

Conclusions: The thorough development process integrating data from literature and experts with a strong focus on integrating the target group in all steps, ensures high content validity. In particular, the different steps ensured that all items are relevant, comprehensive and comprehensible to the target group and construct, thus fulfilling the criteria for content validity based on the COSMIN definition [50]. Moreover, transparent reporting offers the first methodological step-by-step guide for developing patient-reported measures based on qualitative data.

4.4. Publication 4

Reference: Christalle, E., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J. M., Härter M., Bokemeyer C., Gallinat J., Gebhardt C., Magnussen C., Müller V., Schmalstieg-Bahr K., Strahl A., Kriston L., Scholl, I. (2024). Through the patients' eyes: psychometric evaluation of the 64-item version of the Experienced Patient-Centeredness Questionnaire (EPAT-64). *BMJ Quality & Safety*, 34(10), 643-656.

Relevant research questions: RQ4 and RQ5

Aim: This study aims to evaluate psychometric properties of the EPAT and derive the final 64-item version (EPAT-64).

Methods: This is a cross-sectional study, which included adult patients receiving treatment for cardiovascular diseases, cancer, musculoskeletal diseases, and mental disorders in German inpatient and outpatient centers. Several item characteristics were used to select four items for each of the 16 dimensions: (1) proportion of missing values, (2) proportion of "does not concern me" responses, (3) item difficulty, (4) inter-item-correlations, (5) item-total-correlations, and (6) the items content. For the resulting 64 items, structural validity was assessed by confirmatory factor analysis (CFA), reliability by McDonald's Omega and construct validity by examining correlations with general health status and patient satisfaction. CFA was further applied to test measure invariance between the four health conditions.

Results: Analyzing responses of N=2,024 patients (n=1,092 outpatients and n=931 inpatients), items were selected separately for outpatient and inpatient settings, thus resulting in two versions of the EPAT-64. All items of the final EPAT-64 demonstrated high acceptance and acceptable item-total correlations. For approximately 80% of the items, few patients selected "does not concern me." In the outpatient version in particular, seven items showed low applicability, with more than 50% of respondents selecting "does not concern me." Most items had difficulties within the recommended range, although around 30% exhibited ceiling effects. CFA supported a bifactor model, where each item was associated with both a general factor and a specific dimension-related factor. Reliability was high for the general factor with McDonald's Omega but varied strongly across specific dimensions. Construct validity was confirmed by the absence of correlations between general health status with any of the factors and a strong correlation of the general factor with patient satisfaction. Models to test measure invariance could not be estimated.

Conclusions: Psychometric evaluation demonstrated robust properties of the EPAT-64. It offers high flexibility by allowing users to assess individual dimensions through four-item modules based on their needs and resources. The EPAT-64 can be used for quality improvement and the evaluation of interventions aimed at enhancing patient-centeredness across diverse healthcare settings.

4.5. Publication 5

Reference: Christalle, E., Kriston, L., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J.M., Härter, M. & Scholl, I. (2025). Development and Psychometric Testing of EPAT-16: A Short and Valid Measure for Patient-Centeredness From the Patient's Perspective. *Health Expectations*, 28(3), e70296.

Relevant research questions: RQ4 and RQ5

Background: This study aims to select items for the short version of the EPAT - the EPAT-16 - and to evaluate its psychometric properties.

Methods: This cross-sectional study included adult inpatients and outpatients in Germany receiving treatment for cardiovascular diseases, cancer, musculoskeletal conditions, or mental health disorders. To ensure balanced representation of all 16 patient-centeredness dimensions, one item per dimension was selected based on its content relevance to the dimension's definition, item difficulty, the proportion of "does not concern me" responses, and item-total correlations within that dimension. Structural validity was assessed using confirmatory factor analysis (CFA), reliability was evaluated via McDonald's Omega, and construct validity was tested by examining correlations between the EPAT-16 total score and measures of general health and patient satisfaction. Further, measure invariance between health conditions was tested by CFA.

Results: The EPAT-16 demonstrated good acceptability and item-total correlations. Approximately two-thirds of the items demonstrated acceptable difficulty levels, while the remaining third exhibited ceiling effects. In the inpatient sample, all items had "does not concern me" rates below 30%, whereas in the outpatient sample, three items exceeded this rate. While CFA indicated a slight misfit for the unidimensional model, the average variance explained was high, and McDonald's Omega indicated strong reliability. All hypotheses regarding construct validity were supported. Tests showed limitations regarding measure invariance.

Conclusions: The EPAT-16 exhibited solid psychometric properties in the development sample. It offers a practical alternative for assessing overall patient-centeredness when using the full EPAT-64 is not feasible, making it particularly suitable for routine healthcare applications such as constant monitoring and quality improvement. As a generic tool, the EPAT-16 is applicable across various medical settings and thus suitable for public reporting. Future studies should evaluate its psychometric performance in diverse and independent patient populations to confirm the stability and generalizability of the results across different cultural and clinical contexts.

5. DISCUSSION

The first aim was to develop items for the EPAT - a generic PREM to comprehensively assess patient-centeredness from the patients' perspective. This was followed by the second aim to select items for the final questionnaire and validate its psychometric properties.

5.1. Summary of results

Publication 1 [68] presents the study design of the ASPIRED study, within which this dissertation was conducted. It outlines the methodological framework used to address all research questions and highlights how patient input was integrated across all phases of EPAT development - from construct definition and item generation to item selection.

Publication 2 [69] addresses RQ1 through a two-round Delphi study involving 214 patients. Participants were asked to evaluate the relevance of the 15 dimensions of the integrative model of patient-centeredness [24] and to suggest any additional relevant topics. All 15 dimensions were rated as highly relevant by patients, along with an added sixteenth dimension, "patient safety."

Publication 3 [70] focuses on RQ2 and RQ3. In response to RQ2, 152 items were developed based on 48 identified PREMs from the literature, six focus groups with 40 patients, and interviews with ten experts. Relevance and comprehensiveness were ensured through a relevance rating conducted with 32 experts, while comprehensibility was tested in 34 cognitive interviews with patients. Consequently, addressing RQ3, the resulting pool of 125 items demonstrated content validity.

Publication 4 [71] describes the item selection process and psychometric evaluation of the EPAT-64. All previously developed EPAT items were tested in a sample of 2,024 patients. Regarding RQ4, four items were selected for each of the 16 dimensions of patient-centeredness based on item characteristics such as item-total correlations and content relevance. In doing so, two versions were created - one for outpatient and one for inpatient settings. In response to RQ5, structural validity was supported with a bifactor model showing the strongest fit, hypotheses to test construct validity were confirmed, and reliability was high for the general factor, although considerable variance was observed across specific dimensions.

Publication 5 [72] used the same sample to address RQ4 and RQ5 for the short version, EPAT-16. Regarding RQ4, one item per dimension was selected, resulting in a 16-item short version applicable to both outpatient and inpatient settings. In response to RQ5, all hypotheses testing construct validity were confirmed, and the unidimensional factor demonstrated high reliability. With respect to structural validity, confirmatory factor analyses indicated a slight misfit, suggesting the need for further investigation.

5.2. Comprehensive discussion

The primary outcome of this dissertation is a set of PREMs - the EPAT-64 and EPAT-16 - developed to assess patient-centeredness from the patients' perspective. The following section discusses the development process and key characteristics of the EPAT in relation to the existing literature.

5.2.1. Measurement of patient-centeredness

This dissertation advances the measurement of patient-centeredness by providing a comprehensive definition of the construct, enriched by patients' perspectives. Based on feedback from patients in the Delphi study, a 16th dimension - "patient safety" - was added to the existing integrative model of patient-centeredness. Although "safe care" had been subsumed under the dimension "physical support" in the integrative model of patient-centeredness [24], the frequency and emphasis of patient comments on safety aspects warranted recognition as a distinct dimension. This adjustment was further justified when "patient safety" was rated the most important dimension in the second Delphi round. Patients rated all 16 dimensions as highly relevant, in contrast to findings from a Delphi study by Zill and her colleagues [25]. Despite using the same rating scale, they found that experts considered only 11 of the original 15 dimensions to be highly relevant. This discrepancy underscores the importance of involving the target population in defining the construct. Had expert ratings alone guided development, a substantial portion of the construct would have been excluded. In contrast, the Delphi study with patients clearly demonstrated that assessment of all 16 dimensions is required to achieve comprehensiveness and thus content validity.

Overall, the EPAT-64 and EPAT-16 demonstrated commendable psychometric properties. However, limitations were observed regarding item difficulty, with approximately 30% of items showing slight ceiling effects [67], thereby reducing the instrument's ability to differentiate between high levels of patient-centeredness. In addition, some items, particularly in the outpatient sample, exhibited high

rates of “does not concern me” responses. While this suggests limited applicability in some contexts, it does not inherently justify exclusion. For example, the item “When I had pain, I was helped quickly”- used to assess the dimension “physical support” in all EPAT-versions - was marked as not applicable by 54.7% of outpatients, compared to only 19.4% of inpatients. The item was retained regardless, because pain relief is a core aspect of “physical wellbeing”, and for those experiencing pain, the item is highly relevant. Nonetheless, in populations where pain is uncommon, alternative or more representative items may be required.

While testing construct validity, the correlations of patient satisfaction with the general factor of the EPAT-64 as well as the total of the EPAT-16 greatly exceeded the 0.5 threshold (range 0.728 to 0.833), warranting further discussion. While PREMs are intended to capture more objective aspects of care, the relationship between patient experience and patient satisfaction is known to be “complexly related” [37] (page 1). As discussed in the introduction, patient-centeredness has been linked to higher patient satisfaction [18, 20]. On one hand, patient satisfaction may influence patient experience ratings due to the inherently subjective nature of self-reporting [38]; on the other, positive experiences are known to increase patient satisfaction, which is considered an outcome of those experiences [38]. The bifactor structure of the EPAT-64 may help disentangle these constructs, since the specific dimensions of the EPAT-64 were modeled to be orthogonal to the general factor and therefore showed no correlations to patient satisfaction. Hence, the general factor - strongly correlated with patient satisfaction - may capture a more subjective, overall impression of care, while the specific dimensions - modeled as uncorrelated with the general factor - may reflect more objective and distinct elements of patient-centeredness.

Compared to the PEACS [62] - the only generic PREM available in German for assessing patient-centeredness across inpatient and outpatient settings, as identified by Mihaljevic et al. [58] - the EPAT is more comprehensive, covering all 16 dimensions of patient-centeredness, whereas the PEACS addresses only 13. Furthermore, the PEACS did not include a relevance rating; instead, items were primarily selected based on item difficulty and principal component analyses [62]. Consequently, the content validity of the PEACS is questionable, as there is no evidence demonstrating that its items are both relevant and comprehensive. In contrast, the PEACS showed satisfactory results for construct validity and test–retest reliability [62], the latter of which has not yet been evaluated for the EPAT.

It is important to note that whether a generic measure of patient-centeredness is the most suitable option depends on the specific measurement aims and context [73]. Generic measures such as the EPAT have the advantage of being applicable across a wide range of fields and populations [29]. This makes them particularly valuable for population-level assessments, for example, to inform health policy decisions across diverse disease areas [29]. Moreover, they facilitate implementation in outpatient and inpatient settings that treat patients with various conditions, allowing the same questionnaire to be administered without first verifying individual diagnoses. In general, generic measures can enhance the generalizability and interpretability of results to broader contexts [29]. However, such measures may include items that are not relevant to specific patient populations [29], as illustrated by the high “does not concern me” rates observed for some EPAT items in certain subgroups [71, 72]. Conversely, they may omit items that are highly relevant for only particular patient groups [29], thereby reducing content validity for these populations [73]. Moreover, focusing on a specific context often allows for the development of shorter, more targeted questionnaires, which helps to reduce respondent burden [73]. Therefore, when research targets a specific patient population, it may be more appropriate to use condition- or context-specific PREMs. Several reviews have examined PREMs developed for specific contexts, including those for surgical patients [74], mental health care [75], emergency care [76] or adult inpatient settings [77]. In the German healthcare context, the work of the Institute for Quality Assurance and Transparency in Health Care (Institut für Qualitätssicherung und Transparenz im Gesundheitswesen; IQTIG) is particularly noteworthy. The institute, commissioned by the Federal Joint Committee (Gemeinsamer Bundesausschuss; G-BA), is responsible for developing quality indicators tailored to specific healthcare contexts [78]. Since 2022, these indicators have also included the assessment of patient experiences [79], with five PREMs already in use (e.g., for patients with schizophrenia or those undergoing percutaneous coronary intervention) and three additional PREMs currently under development [80].

5.2.2. Questionnaire development with focus on content validity

The development process of the EPAT, together with its transparent, open-access reporting, provides a best-practice example for future developments of patient-reported measures. Although many guidelines emphasize the importance of incorporating qualitative data from the target population to ensure content validity (e.g. [45, 50, 51]), they frequently describe item development only vaguely as being “based on” qualitative data. In many studies on questionnaire development, detailed documentation of coding strategies, item derivation, or the mapping of items to the construct is lacking [52]. Applying established methods of formal qualitative data analysis, combined with reporting standards such as the COnsolidated criteria for REporting Qualitative research (COREQ) guideline [81] - as was done in the EPAT development [70] - can substantially enhance methodological rigor and transparency. Furthermore, patients were involved in all three development stages identified by Wiering et al. [53], distinguishing the EPAT from many other instruments. In their review, Wiering et al. found that 25.9% of PROMs involved no patient participation at any stage, and only 6.7% engaged patients in all three [53].

Moreover, this dissertation adhered to COSMIN’s recommendation to prioritize content validity [50], including during the item selection phase - an approach still uncommon in current practice. While COSMIN explicitly advises assessing comprehensiveness in the final item set, most psychometric guidelines treat content validity as an early step, after which item reduction is driven primarily by statistical criteria (e.g. [26, 27, 45]). Boateng et al. [45], for example, argue that item selection should retain only items strongly related to the construct. While this approach may improve reliability by increasing internal consistency, it also risks semantic redundancy [82] and may inadvertently narrow the construct’s scope by reducing content coverage [55]. Furthermore, cut-off values used for item selection are often arbitrary, and recommendations vary considerably - for example, Boateng et al. suggest deleting all items with inter-item correlations below 0.3 [45], whereas de Vet and colleagues recommend deletion if correlations are below 0.2 or above 0.9 [27].

In contrast, literature on short-form development explicitly supports content-driven item selection. Diamantopoulos and his colleagues [82] recommend that at least two researchers assess item redundancy and ensure that all relevant facets of the construct remain represented. In the EPAT item selection process, five experts evaluated each item’s relevance to the construct definition and its added content compared with already selected items. Similarly, Smith and his colleagues [55] emphasize that content coverage should be reviewed by multiple experts and that dimensions

should be proportionally represented. In the EPAT, this principle was applied by evaluating each dimension separately and selecting an equal number of items per dimension. Smith et al. [55] also propose considering item content alongside item-total correlations - a practice that was followed and extended in EPAT development through the inclusion of additional statistical indicators. Yet, existing recommendations call for a formalized expert rating of content importance [54], which was not applied during the EPAT's item selection process. Nonetheless, the discussion-based approach adopted in this dissertation - evaluating multiple item characteristics in structured team meetings - offers the advantage of real-time exchange of perspectives and may serve as a pragmatic, time-efficient alternative, since even expert ratings typically require discussion and interpretation to finalize item selection.

In summary, the methods proposed in this dissertation are not the only valid approach to ensuring content validity. However, the EPAT development process demonstrates that consistently prioritizing content validity across all stages of development is both feasible and effective in practice.

5.2.3. Strengths and limitations

This section presents the overarching strengths and limitations of this dissertation as a whole. The specific strengths and limitations of each individual publication are discussed within the respective articles [68-72].

A key strength of this dissertation is its consistent and comprehensive focus on content validity across all stages of development. First, patients were involved at all three stages of the development process outlined by Wiering et al. [53]: (1) patient input informed the definition of the construct, (2) items were generated based on patient focus groups, and (3) item comprehensibility was tested in cognitive interviews with patients. Second, all development steps were anchored in a well-established conceptual model of patient-centeredness [24], which was further enriched by the patients' perspective. The model served not only as the analytical framework for qualitative data analysis during item development, but also as a guiding framework for item selection, thereby integrating both data-driven and theory-driven approaches. Relevance and comprehensibility were established prior to item reduction and are not influenced by deleting items. However, by selecting items for each dimension based on their importance to the dimension's definition, comprehensiveness was maintained - ensuring the overall content validity of the final EPAT. Third,

the study adhered to transparent and detailed reporting practices, following existing guidelines (e.g., [48, 81]) and enhancing them through additional documentation to support replicability. This transparent, step-by-step reporting serves as a best practice example for other researchers developing patient-reported measures. Finally, the EPAT is highly accessible to all users, aligning with current efforts to promote open science. All related publications are available in open access, and the different versions of the questionnaire can be freely downloaded at www.uke.de/epat, under a Creative Commons license.

Despite these strengths, several limitations should be acknowledged. First, although the recommendations of Wiering et al. [53] regarding patient involvement were followed, the role of patients was limited to that of research participants. Future research could benefit from patient and public involvement in the research process, for example as advisors, partners or decision makers [83]. Integrating their experiential expertise may further enhance the relevance and impact of the research [84]. Second, while item selection was guided by discussions based on the construct definition to maintain representativeness of its content [55], patients were not involved in this step. Stronger evidence of content validity could have been obtained by involving the target group in the item selection process and by having them assess the content validity of the final item set [23, 85]. Third, while the inclusion of four disease groups enabled subgroup analyses and reflected a broad range of healthcare experiences, the generalizability of the findings to other patient populations remains uncertain. Measurement invariance across these four groups could not be confirmed in the present studies [71, 72]. Hence, while the EPAT was designed as a generic measure instrument, the ability to compare scores across disease groups and settings might be limited.

Implications

This dissertation described the development and psychometric evaluation of the first German PREM to assess patient-centeredness. Based on the findings, implications can be derived in three areas: (1) future research on the measurement of patient-centeredness, (2) psychometric methods in questionnaire development, and (3) practical applications for the use of the EPAT in routine healthcare. The following sections discuss each of these in turn.

5.2.4. Research implications for measuring patient-centeredness

Several implications for future research arise from this dissertation, both in terms of further evaluating the EPAT and adapting it to other contexts.

Several questions remain regarding the psychometric properties of the EPAT. Due to the one-time-point study design, no data are currently available on test–retest reliability or responsiveness. Demonstrating responsiveness is essential for evaluative purposes, such as assessing the effectiveness of interventions [27]. Insights into the EPAT’s responsiveness may be gained from the Hamburg “Life After Cancer Program,” a hybrid effectiveness–implementation study focused on the development, implementation, and evaluation of a structured survivorship program. The EPAT will be used as a secondary outcome. Moreover, since psychometric properties can change due to item reduction [55], validation of the final versions of both the EPAT-64 and EPAT-16 is needed [26]. Ideally, both versions should be tested simultaneously in the same sample to examine their shared variance [55].

Open questions regarding the measurement model remain. In particular, the bifactor model of the EPAT-64 and its potential to partition variance based on subjective and objective ratings should be explored in other populations. In addition, measurement invariance should be further investigated to see to what extent EPAT scores are comparable between different patient populations and settings. Overall, psychometric properties are not transitive, but are valid only for the particular population [27]. Hence, before applying the EPAT to different populations, it is crucial to test its psychometric properties in this population. Adaptation to specific populations or care settings may also be required. For example, in the CarePreg study, the EPAT was adapted to assess person-centeredness in healthcare and social support services for people experiencing unintended pregnancy [86].

Additionally, the EPAT has been translated into English using a team-based approach [87]. This is a key step toward enabling international use and comparative studies. However, before this version can be widely implemented, further research is needed to evaluate its comprehensibility and cross-cultural validity. In general, translation, cultural adaptation, and validation would substantially broaden applications of the EPAT and improve representativeness by including patients who are not proficient in German.

Lastly, to achieve a comprehensive understanding of patient-centeredness, triangulation from three perspectives - patients, healthcare professionals, and observers - is recommended [21]. The EPAT could be adapted for self-assessment or observer ratings. A first step was taken in a study which evaluated a patient-centered communication skills training for nurses and adapted the EPAT-16 as a self-assessment tool for healthcare professionals [88].

5.2.5. Research implications for psychometric methods

This dissertation offers a best practice example for the development of patient-reported measures with a strong focus on content validity. By transparently reporting item development from qualitative data and consistently involving patients throughout all phases, Publication 3 serves as a practical example for future questionnaire development. Future instrument development can use this methodology as a blueprint, adapting it to suit different constructs and target populations. By following these steps, content validity is ensured through (1) the consistent integration of the target group's perspective and (2) the systematic consideration of the construct's definition. Even if researchers choose to adapt rather than fully adopt the procedure, transparent reporting on where and why adaptations were made will contribute to improved transparency about item development methods and hopefully spark more debate about available strategies, their advantages, and limitations. As more such examples emerge, they can inform updates to existing questionnaire development guidelines. In particular, they can help clarify what constitutes "appropriate methods" [50] as described in the COSMIN standards for assessing content validity.

This dissertation also provides a pragmatic and feasible example of how content validity can be incorporated into the item selection process - simply by considering the content of the items and definition of the target construct alongside with statistical results [71, 72]. Future questionnaire

development should adopt this integrative approach, rather than relying solely on statistical cut-offs, to ensure that the full content of the construct is retained in the final instrument.

5.2.6. Practice implications

Measuring patient experiences is not enough - the results must be actively used to improve the quality of healthcare [89]. The EPAT provides a valuable tool to inform and evaluate efforts aimed at strengthening patient-centeredness. The next step is to explore how this can be achieved in practice, beginning with the implementation of continuous EPAT assessments into routine care. Implementation should be tailored to each institutional context, as healthcare processes and infrastructures vary widely - and so do the most suitable strategies for integrating a questionnaire. Feedback from both healthcare professionals and patients is therefore essential to ensure that implementation is both feasible and accepted by all stakeholders [89]. Encouragingly, successful examples of PREM implementation are increasingly reported in the scientific literature, including implementation via an automated data collection system [90] as well as tailored and cocreated implementation strategies [91, 92]. Standardized tools such as the Patient-Reported Measures Implementation Assessment Tool (PRMIAT) [93] can be used to evaluate the acceptability and perceived value of integrating PREMs into routine care.

Once the EPAT is embedded in everyday workflows, attention should shift to how its results are applied. For example, results could be presented to clinical teams as a basis for joint reflection on potential areas for improvement [94]. To deepen the understanding of PREM results, qualitative data could also be incorporated [12, 95]. Patient narratives, for instance, can contextualize quantitative results and provide richer insights within a mixed-methods approach [96].

The EPAT can be used for several applications as outlined in the introduction. It offers high flexibility—through the brevity of the EPAT-16 on the one hand, and the modular design of the EPAT-64 on the other. The EPAT-64 comprises 16 modules with four items each, corresponding to the 16 dimensions of the integrative model of patient-centeredness. Instead of administering all 64 items, users can select a subset of modules relevant to their specific measurement objectives. This flexibility makes the EPAT a practical instrument for continuous assessment in routine care, such as regular feedback or quality monitoring. Its transfer into practice is further facilitated by the Creative Commons license, which allows all individuals and institutions to use the EPAT—regardless of financial resources or whether the facility operates on a commercial or non-commercial basis.

Developed within the German healthcare system and designed as a generic instrument applicable in both inpatient and outpatient settings, the EPAT is particularly suited to fulfilling national calls for patient experience measurement across healthcare institutions as part of transparent quality reporting [63]. Such applications empower patients by providing the information they consider most relevant for evaluating the quality of care across facilities [64].

Finally, the generic design of the EPAT allows its use not only at the micro level (e.g., within individual healthcare institutions), which has been the primary focus of most PREM applications to date [35]. The EPAT can also be applied in broader contexts, such as population-representative surveys assessing the overall state of patient-centeredness in Germany. This approach has already been demonstrated in the “TK-Monitor Patientensicherheit 2024”, where EPAT items related to patient safety were adapted for telephone interviews [97]. Furthermore, a review by Jesus and his colleagues [98] identified several studies that successfully used PREMs to evaluate the effectiveness of quality improvement initiatives at the meso and macro levels, highlighting the potential for system-wide application of PREMs.

6. CONCLUSIONS

Psychometrically sound measurement is essential for improving patient-centeredness. To ensure measurement quality, content validity must be prioritized throughout all stages of questionnaire development. Until now, no instrument existed that comprehensively assessed patient-centeredness from the patients' perspective. This dissertation addressed this gap by developing the Experienced Patient-Centeredness Questionnaire (EPAT), a comprehensive and generic measure of patient-centeredness.

Several conclusions can be drawn from this dissertation:

- The EPAT is the first comprehensive German-language PREM to assess patient-centeredness from the patients' perspective.
- The EPAT-64 and its short form, the EPAT-16, demonstrated good psychometric properties. The EPAT is a generic measure suitable for both inpatient and outpatient settings.
- Transparent reporting of the item generation process offers a feasible best practice example of how patient-reported measures can be developed using qualitative data.
- The EPAT item selection process demonstrated that combining data-driven and theory-driven approaches is a viable strategy for achieving content validity in the final questionnaire.
- Further psychometric research is needed to investigate measurement invariance and responsiveness of the EPAT across diverse patient populations and clinical settings.
- Future studies should explore strategies for implementing the EPAT in routine healthcare practice to support continuous quality improvement.

7. LIST OF ABBREVIATIONS

ASPIRED	Assessment of Patient-Centeredness through patient-reported experience measures
BMBF	Bundesministerium für Bildung und Forschung
CFIR	Consolidated Framework for Implementation Research
COREQ	COnsolidated criteria for REporting Qualitative research guideline
COSMIN	COnsensus-based Standards for the selection of health status Measurement Instruments
EPAT	Experienced Patient-Centeredness Questionnaire
G-BA	Gemeinsamer Bundesausschuss
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
PREM	Patient-reported experience measure
RQ	Research questions
UKE	University Medical Center Hamburg-Eppendorf

8. REFERENCES

1. Gerteis M, Edgman-Levitan S, Daley J, Delbanco TL. Through the patient's eyes: understanding and promoting patient-centered care. San Francisco: Jossey-Bass; 1993.
2. Picker. Our history [webpage]. Available from: <https://picker.org/who-we-are/our-history/> [Accessed: 01.03.2025]
3. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: The National Academies Press; 2001.
4. World Health Organization, Organisation for Economic Co-operation and Development, and The World Bank. Delivering quality health services - A global imperative for universal health coverage. Geneva; 2018. Contract No.: 02.03.2025. Available from: <https://iris.who.int/bitstream/handle/10665/272465/9789241513906-eng.pdf> [Accessed: 02.03.2025]
5. Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. The Lancet global health. 2018;6(11):e1196-e252.
6. National Academies of Sciences E, Medicine, Health, Medicine D, Board on Health Care S, Board on Global H, et al. The National Academies Collection: Reports funded by National Institutes of Health. Crossing the Global Quality Chasm: Improving Health Care Worldwide. Washington (DC): National Academies Press (US) Copyright 2018 by the National Academy of Sciences. All rights reserved.; 2018.
7. Hahlweg P, Bieber C, Brütt AL, Dierks M-L, Dirmaier J, Donner-Banzhoff N, et al. Moving towards patient-centered care and shared decision-making in Germany. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 2022;171:49-57.
8. Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten. Jahrgang 2013; Teil I Nr. 9. Available from: https://www.bgbl.de/xaver/bgbl/start.xav?start=///*%5B@attr_id=%27bgbl113s0277.pdf%27%5D#/switch/tocPane?_ts=1758897277759 [Accessed: 26.09.2025]
9. Bundesministerium für Gesundheit. Nationaler Krebsplan - Handlungsfelder, Ziele, Umsetzungsempfehlungen und Ergebnisse. 2008. Available from: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/N/Nationaler_Krebsplan/Informationspapier_Nationaler_Krebsplan.pdf [Accessed: 26.09.2025]
10. Härter M, Müller H, Dirmaier J, Donner-Banzhoff N, Bieber C, Eich W. Patient participation and shared decision making in Germany—history, agents and current transfer to practice. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 2011;105(4):263-70.
11. Bundesministerium für Gesundheit. Spahn: "Patientenorientierung ist das Leitbild für das gesamte Gesundheitswesen" [webpage]. 2019. Available from: <https://www.bundesgesundheitsministerium.de/gesundheitskompetenz/ministerinterview.html> [Accessed: 02.03.2025]
12. Edgman-Levitan S, Schoenbaum SC. Patient-centered care: achieving higher quality by designing care through the patient's eyes. Israel Journal of Health Policy Research. 2021;10:1-5.
13. Duggan PS, Geller G, Cooper LA, Beach MC. The moral nature of patient-centeredness: is it "just the right thing to do"? Patient education and counseling. 2006;62(2):271-6.
14. Epstein RM, Street RL. The values and value of patient-centered care. Annals Family Med; 2011. p. 100-3.
15. Beauchamp TL. The principles of biomedical ethics as universal principles. Islamic Perspectives on the Principles of Biomedical Ethics: Muslim Religious Scholars and Biomedical Scientists in Face-to-Face Dialogue with Western Bioethicists: World Scientific; 2018. p. 91-119.
16. Delbanco T, Berwick DM, Boufford JJ, Ollenschläger G, Plamping D, Rockefeller RG. Healthcare in a land called PeoplePower: nothing about me without me. Health expectations. 2001;4(3):144-50.

17. Karazivan P, Dumez V, Flora L, Pomey M-P, Del Grande C, Ghadiri DP, et al. The patient-as-partner approach in health care: a conceptual framework for a necessary transition. *Academic Medicine*. 2015;90(4):437-41.
18. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Medical Care Research and Review*. 2013;70(4):351-79.
19. Stewart M, Brown JB, Donner A, McWhinney IR, Oates J, Weston WW, et al. The impact of patient-centered care on outcomes. *Fam Pract*. 2000;49(9):796-804.
20. Lewin SA, Skea ZC, Entwistle V, Zwarenstein M, Dick J. Interventions for providers to promote a patient-centred approach in clinical consultations. *The Cochrane database of systematic reviews*. 2001(4):CD003267.
21. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Social science & medicine*. 2000;51(7):1087-110.
22. Zill JM, Scholl I, Härter M, Dirmaier J. Evaluation of dimensions and measurement scales in patient-centeredness. *Patient Preference and Adherence*. 2013:345-51.
23. Jung A, Challoumas D, Pagels L, Armijo-Olivo S, Braun T, Luedtke K. Guidelines for the development and validation of patient-reported outcome measures: a scoping review. *BMJ Evidence-Based Medicine*. 2024;29(6):363-73.
24. Scholl I, Zill JM, Härter M, Dirmaier J. An integrative model of patient-centeredness—a systematic review and concept analysis. *PloS one*. 2014;9(9):e107828.
25. Zill JM, Scholl I, Härter M, Dirmaier J. Which dimensions of patient-centeredness matter?—Results of a web-based expert delphi survey. *PloS one*. 2015;10(11):e0141978.
26. Swan K, Speyer R, Scharitzer M, Farneti D, Brown T, Woisard V, et al. Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. *Frontiers in Psychology*. 2023;14:1225850.
27. De Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine: a practical guide*: Cambridge university press; 2011.
28. Atkinson MJ, Lennox RD. Extending basic principles of measurement models to the design and validation of patient reported outcomes. *Health and Quality of Life Outcomes*. 2006;4(1):65.
29. Cano SJ, Hobart JC. The problem with health measurement. *Patient preference and adherence*. 2011:279-90.
30. Kilbourne AM, Beck K, Spaeth-Rublee B, Ramanuj P, O'Brien RW, Tomoyasu N, et al. Measuring and improving the quality of mental health care: a global perspective. *World psychiatry*. 2018;17(1):30-8.
31. Epstein RM, Franks P, Fiscella K, Shields CG, Meldrum SC, Kravitz RL, et al. Measuring patient-centered communication in patient–physician consultations: theoretical and practical issues. *Social science & medicine*. 2005;61(7):1516-28.
32. Beattie M, Murphy DJ, Atherton I, Lauder W. Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Systematic reviews*. 2015;4(1):1-21.
33. Mertz K, Eppler S, Yao J, Amanatullah DF, Chou L, Wood KB, et al. Patient perceptions correlate weakly with observed patient involvement in decision-making in orthopaedic surgery. *Clinical Orthopaedics and Related Research®*. 2018;476(9):1859-65.
34. Hudon C, Fortin M, Haggerty JL, Lambert M, Poitras M-E. Measuring patients' perceptions of patient-centered care: a systematic review of tools for family medicine. *The Annals of Family Medicine*. 2011;9(2):155-64.
35. Gilmore KJ, Corazza I, Coletta L, Allin S. The uses of patient reported experience measures in health systems: a systematic narrative review. *Health policy*. 2023;128:1-10.
36. Tzelepis F, Sanson-Fisher RW, Zucca AC, Fradgley EA. Measuring the quality of patient-centered care: why patient-reported measures are critical to reliable assessment. *Patient preference and adherence*. 2015:831-5.
37. Bull C. Patient satisfaction and patient experience are not interchangeable concepts. *International Journal for Quality in Health Care*. 2021;33(1):mzab023.

38. Larson E, Sharma J, Bohren MA, Tunçalp Ö. When the patient is the expert: measuring patient experience and satisfaction with care. *Bulletin of the World Health Organization*. 2019;97(8):563.
39. Bull C, Byrnes J, Hettiarachchi R, Downes M. A systematic review of the validity and reliability of patient-reported experience measures. *Health services research*. 2019;54(5):1023-35.
40. Bull C, Callander EJ. Current PROM and PREM use in health system performance measurement: still a way to go. *Patient Experience Journal*. 2022;9(1):12-8.
41. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ open*. 2013;3(1):e001570.
42. Anhang Price R, Elliott MN, Zaslavsky AM, Hays RD, Lehrman WG, Rybowski L, et al. Examining the role of patient experience surveys in measuring health care quality. *Medical Care Research and Review*. 2014;71(5):522-54.
43. Reeves R, West E, Barron D. Facilitated patient experience feedback can improve nursing care: a pilot study for a phase III cluster randomised controlled trial. *BMC Health Serv Res*. 2013;13:259.
44. Murante AM, Vainieri M, Rojas D, Nuti S. Does feedback influence patient - professional communication? Empirical evidence from Italy. *Health Policy*. 2014;116(2-3):273-80.
45. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health*. 2018;6:149.
46. Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf [Accessed: 17.05.2021]
47. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*. 2010;63(7):737-45.
48. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of life research*. 2021;30:2197-218.
49. Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*: Oxford University Press, USA; 2015.
50. Terwee CB, Prinsen CA, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of life research*. 2018;27:1159-70.
51. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. *Optometry and Vision Science*. 2007;84(8):663-74.
52. Ricci L, Lanfranchi J-B, Lemetayer F, Rotonda C, Guillemin F, Coste J, et al. Qualitative methods used to generate questionnaire items: a systematic review. *Qualitative health research*. 2019;29(1):149-56.
53. Wiering B, de Boer D, Delnoij D. Patient involvement in the development of patient-reported outcome measures: a scoping review. *Health Expectations*. 2017;20(1):11-23.
54. Schipolowski S, Schroeders U, Wilhelm O. Pitfalls and challenges in constructing short forms of cognitive ability measures. *Journal of Individual Differences*. 2014.
55. Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. *Psychological assessment*. 2000;12(1):102.
56. Ziegler M, Kemper CJ, Kruijen P. *Short scales—Five misunderstandings and ways to overcome them*. Hogrefe Publishing; 2014.

57. Handley SC, Bell S, Nembhard IM. A systematic review of surveys for measuring patient-centered care in the hospital setting. *Medical care*. 2021;59(3):228-37.
58. Mihaljevic AL, Doerr-Harim C, Kalkum E, Strunk G. Measuring patient centeredness with German language Patient-Reported Experience Measures (PREM)—A systematic review and qualitative analysis according to COSMIN. *PLOS ONE*. 2022;17(11):e0264045.
59. Squires A, Bruyneel L, Aiken LH, Van den Heede K, Brzostek T, Busse R, et al. Cross-cultural evaluation of the relevance of the HCAHPS survey in five European countries. *International Journal for Quality in Health Care*. 2012;24(5):470-5.
60. Groene O, Arah OA, Klazinga NS, Wagner C, Bartels PD, Kristensen S, et al. Patient experience shows little relationship with hospital quality management strategies. *PloS one*. 2015;10(7):e0131805.
61. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *International Journal for Quality in Health Care*. 2002;14(5):353-8.
62. Noest S, Ludt S, Klingenberg A, Glassen K, Heiss F, Ose D, et al. Involving patients in detecting quality gaps in a fragmented healthcare system: development of a questionnaire for Patients' Experiences Across Health Care Sectors (PEACS). *International Journal for Quality in Health Care*. 2014;26(3):240-9.
63. Gesundheitsministerkonferenz. Beschlüsse der GMK 20.06.2018 - 21.06.2018 [webpage]. Available from: <https://www.gmkonline.de/Beschluesse.html?id=698&jahr=2018> [Accessed: 08.04.2025]
64. Ärzteblatt D. Deutsche fühlen sich über Qualität im Gesundheitswesen schlecht informiert [webpage]. 2022. Available from: <https://www.aerzteblatt.de/news/deutsche-fuehlen-sich-ueber-qualitaet-im-gesundheitswesen-schlecht-informiert-9d6d57e2-3ffd-4820-ad21-9e779e276c9a> [Accessed: 02.03.2025]
65. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation science*. 2009;4(1):50.
66. Sepucha KR, Matlock DD, Wills CE, Ropka M, Joseph-Williams N, Stacey D, et al. "It's valid and reliable" is not enough: critical appraisal of reporting of measures in trials evaluating patient decision aids. *Medical Decision Making*. 2014;34(5):560-6.
67. Döring N, Bortz J. *Forschungsmethoden und Evaluation*. Wiesbaden: Springerverlag. 2006.
68. Christalle E, Zeh S, Hahlweg P, Kriston L, Härter M, Scholl I. Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ open*. 2018;8(10):e025896.
69. Zeh S, Christalle E, Hahlweg P, Härter M, Scholl I. Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ open*. 2019;9(12):e031741.
70. Christalle E, Zeh S, Hahlweg P, Kriston L, Härter M, Zill J, et al. Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—A best practice example for generating patient-reported measures from qualitative data. *Health Expectations*. 2022.
71. Christalle E, Zeh S, Führes H, Schellhorn A, Hahlweg P, Zill JM, et al. Through the patients' eyes: psychometric evaluation of the 64-item version of the Experienced Patient-Centeredness Questionnaire (EPAT-64). *BMJ Quality & Safety*. 2024:bmjqs-2024-017434.
72. Christalle E, Kriston L, Zeh S, Führes H, Schellhorn A, Hahlweg P, et al. Development and Psychometric Testing of EPAT-16: A Short and Valid Measure for Patient-Centeredness From the Patient's Perspective. *Health Expectations*. 2025;28(3):e70296.
73. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Medical care*. 1989;27(3):S217-S32.

74. Darwish M, Nandy S, Willis S, Coulson J, Withers K, Bosanquet DC. Surgical patient-reported experience measures and qualitative experience studies: systematic review. *BJS open*. 2025;9(1):zrae142.
75. Fernandes S, Fond G, Zendjidian XY, Baumstarck K, Lançon C, Berna F, et al. Measuring the patient experience of mental health care: a systematic and critical review of patient-reported experience measures. *Patient preference and adherence*. 2020;2147-61.
76. Male L, Noble A, Atkinson J, Marson T. Measuring patient experience: a systematic review to evaluate psychometric properties of patient reported experience measures (PREMs) for emergency care service provision. *International Journal for Quality in Health Care*. 2017;29(3):314-26.
77. Kang Y, Guan T, Chen X, Zhang Y. Patient-Reported Experience Measures in Adult Inpatient Settings: A Systematic Review. *Journal of Nursing Management*. 2024;2024(1):5166392.
78. Heller K-D, Hoffmann R, Kladny B. Qualität und Krankenhausplanung – planungsrelevante Qualitätsindikatoren des IQTiG. *Krankenhausstrukturgesetz*. 2016;05(05):458-60.
79. IQTiG - Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Start der ersten Patientenbefragung der gesetzlichen Qualitätssicherung im QS-Verfahren QS PCI am 1. Juli 2022 [webpage]. 2022. Available from: <https://iqtig.org/aktuelles/news/start-der-ersten-patientenbefragung-der-gesetzlichen-qualitaetssicherung-im-qs-verfahren-qs-pci-am-1-juli-2022-1/> [Accessed: 15.07.2025]
80. IQTiG - Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. QS-Verfahren - Befragungen [webpage]. Available from: <https://iqtig.org/qs-verfahren/hintergrund/durchfuehrung-von-befragungen/> [Accessed: 15.07.2025]
81. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*. 2007;19(6):349-57.
82. Diamantopoulos A, Sarstedt M, Fuchs C, Wilczynski P, Kaiser S. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*. 2012;40(3):434-49.
83. Smits D-W, Van Meeteren K, Klem M, Alsem M, Ketelaar M. Designing a tool to support patient and public involvement in research projects: the Involvement Matrix. *Research involvement and engagement*. 2020;6(1):30.
84. Shen S, Doyle-Thomas KA, Beesley L, Karmali A, Williams L, Tanel N, et al. How and why should we engage parents as co-researchers in health research? A scoping review of current practices. *Health Expectations*. 2017;20(4):543-54.
85. Mokkink LB, Terwee CB, de Vet HC. COSMIN: Consensus-based standards for the selection of health status measurement instruments. *Encyclopedia of quality of life and well-being research*: Springer; 2024. p. 1445-8.
86. Zill JM, Lindig A, Reck LM, Scholl I, Härter M, Hahlweg P. Assessment of person-centeredness in healthcare and social support services for women with unintended pregnancy (CarePreg): protocol for a mixed-method study. *BMJ open*. 2022;12(9):e066939.
87. Harkness J, Pennell BE, Schoua-Glusberg A. Survey questionnaire translation and assessment. *Methods for testing and evaluating survey questionnaires*. 2004:453-73.
88. Lindig A, Mielke K, Frerichs W, Cöllen K, Kriston L, Härter M, et al. Evaluation of a patient-centered communication skills training for nurses (KOMPAT): study protocol of a randomized controlled trial. *BMC nursing*. 2024;23(1):2.
89. Coulter A, Locock L, Ziebland S, Calabrese J. Collecting data on patient experience is not enough: they must be used to improve care. *BMJ*. 2014;348:g2225.
90. Chakhunashvili A, Blommengren A, Kullberg A. Implementation of Automated PREM Process to Better Capture Patients' Overall Experience of Care Services at Karolinska University Hospital. *The International Journal of Health Planning and Management*. 2025.

91. Van Rooijen M, Van Dijk-De Vries A, Lenzen S, Dalemans R, Moser A, Beurskens A. How to foster successful implementation of a patient reported experience measurement in the disability sector: an example of developing strategies in co-creation. *Research Involvement and Engagement*. 2021;7(1):45.
92. van Rooijen M, van Dijk-de Vries A, Lenzen S, Dalemans R, Moser A, Beurskens AJ. Implementation of a patient-reported experience measure in a Dutch disability care organization: A process evaluation of cocreated tailored strategies. *Health Expectations*. 2023;26(1):132-45.
93. Amat-Fernandez C, Pardo Y, Ferrer M, Bosch G, Lizano-Barrantes C, Briseño-Díaz R, et al. Evaluating the implementation of PROMs and PREMs in routine clinical care: co-design of tools from the perspective of patients and healthcare professionals. *Health and Quality of Life Outcomes*. 2025;23(1):15.
94. De Rosis S, Cerasuolo D, Nuti S. Using patient-reported measures to drive change in healthcare: the experience of the digital, continuous and systematic PREMs observatory in Italy. *BMC Health Serv Res*. 2020;20(1):315.
95. Tsianakas V, Maben J, Wiseman T, Robert G, Richardson A, Madden P, et al. Using patients' experiences to identify priorities for quality improvement in breast cancer care: patient narratives, surveys or both? *BMC health services research*. 2012;12(1):271.
96. Locock L, Robert G, Boaz A, Vougioukalou S, Shuldham C, Fielden J, et al. Using a national archive of patient experience narratives to promote local patient-centered quality improvement: an ethnographic process evaluation of 'accelerated' experience-based co-design. *Journal of health services research & policy*. 2014;19(4):200-7.
97. Müller H, Müller B, Schwappach D. TK-Monitor Patientensicherheit: Erlebte Patientensicherheit aus Sicht der Bevölkerung. Eine bevölkerungsrepräsentative Befragung zum Stand der Sicherheit in der medizinischen Versorgung. 2024.
98. Jesus TS, Lee D, Zhang M, Stern BZ, Struhar J, Heinemann AW, et al. Organizational and Service Management Interventions for Improving the Patient Experience With Care: Systematic Review of the Effectiveness. *The International Journal of Health Planning and Management*. 2025.

9. PUBLICATIONS

9.1. Publication 1

Reference: Christalle, E.*, Zeh, S.*, Hahlweg, P., Kriston, L., Härter, M., & Scholl, I. (2018). Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open*, 8(10), e025896.

(*Shared first authorship)

BMJ Open Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study

Eva Christalle, Stefan Zeh, Pola Hahlweg, Levente Kriston, Martin Härter, Isabelle Scholl

To cite: Christalle E, Zeh S, Hahlweg P, *et al.* Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open* 2018;**8**:e025896. doi:10.1136/bmjopen-2018-025896

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-025896>).

EC and SZ contributed equally.

Received 7 August 2018

Revised 1 September 2018

Accepted 12 September 2018



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Correspondence to

Dr Isabelle Scholl;
i.scholl@uke.de

ABSTRACT

Introduction The impact of patient centredness (PC) in healthcare has grown over the years. However, conceptualisations of PC are heterogeneous. Existing patient-reported measures of PC have shown inconsistencies and shortcomings. This impedes the comparison of results across studies. To foster PC, it is important to know which dimensions matter most to patients and to be able to measure its current extent from the patients' perspective. This study aims (1) to assess relevance of dimensions of PC from the patients' perspective, (2) to develop and psychometrically test a core set of patient-reported experience measures (PREMs) assessing PC and (3) to investigate the feasibility of implementation of this core set in routine healthcare.

Methods and analysis A mixed-methods approach will be used. In phase 1, 200 patients will assess the relevance of the dimensions of PC in a Delphi study using a plain language description. In phase 2, the core set of PREMs will be developed through literature reviews, focus groups, key informant interviews and content validity ratings. The core set will be tested psychometrically in a cross-sectional study with 2000 inpatient and outpatients with different chronic conditions (ie, cancer, cardiovascular diseases, mental disorders and musculoskeletal disorders). In phase 3, the feasibility of implementation of the core set will be assessed through semistructured interviews with healthcare practitioners after piloting in routine care. Furthermore, an expert workshop will be held on how to foster implementation.

Ethics and dissemination The study will be carried out in accordance to the latest version of the Helsinki Declaration of the World Medical Association and principles of good scientific practice. The study was approved by the Ethics Committee of the Medical Association Hamburg, Germany (study ID: PV5724). The study results will be disseminated in scientific journals and through collaboration partners and plain language press releases.

INTRODUCTION

The concept of patient centredness (PC) has been widely discussed over the last years. Research has shown that PC is positively associated with patient satisfaction, well-being

Strengths and limitations of this study

- No exhaustive patient-reported measure of patient centredness (PC) in Germany yet.
- Combination of quantitative and qualitative methods to ensure a comprehensive understanding of patients' perspectives.
- Dual purpose: development of a patient-reported measure of PC and investigation of the feasibility of its implementation.
- Extensive inclusion of patients' perspectives throughout development and evaluation of the measure.
- Limitation to four disease groups and geographical restriction to the metropolitan area of Hamburg, Germany.

and self-management,^{1–3} which are especially relevant in chronic disease management. Furthermore, the US Institute of Medicine declared patient-centred care to be one of six goals to achieve high-quality healthcare in their landmark publication 'Crossing the Quality Chasm' at the beginning of the 21st century.^{4 5} In Germany, specific federal funding programmes were initiated in order to foster research on high-quality care.⁶ Since 2013, the Patient Rights Law in Germany also stresses the importance of PC by outlining the right to informed decisions, comprehensive and comprehensible patient information and decisions based on the partnership of the patient and the clinician.⁷ Similar developments have been reported in the UK, Canada, Australia and several other countries.⁸

Unfortunately, a lack of conceptual clarity and inconsistencies in the use of the term complicate investigating and implementing PC. Several conceptual models have been developed,^{9–13} making it difficult to compare research results. Scholl and colleagues aimed to close this gap by developing an integrative

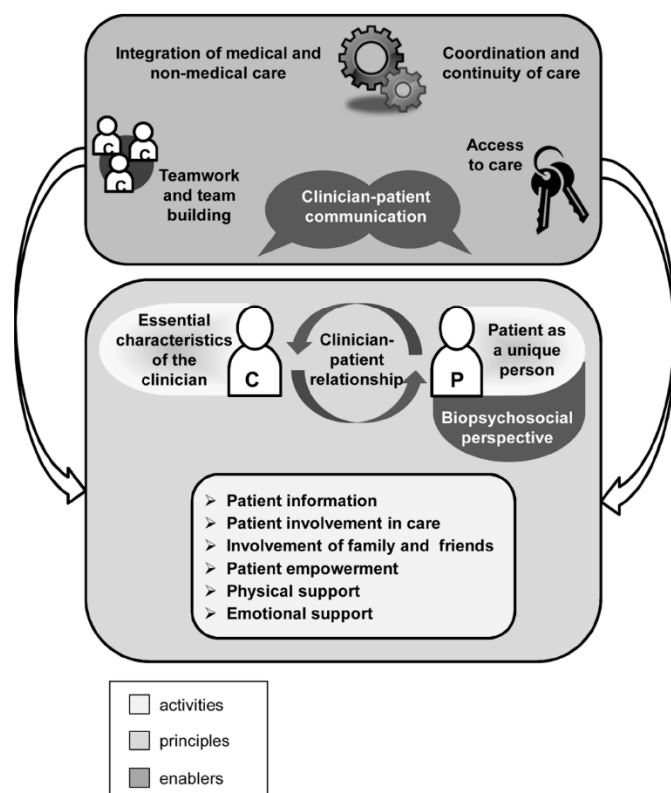


Figure 1 Integrative model of patient centredness.

generic model highlighting 15 dimensions of PC (see figure 1).¹⁴ In our research, we use this definition of PC. The model is based on a systematic review and was validated by assessing the relevance in a Delphi study conducted with n=71 international experts.¹⁵ However, in this expert survey, patients' perspectives on clarity and relevance of the dimensions were not represented.¹⁵ A stronger focus on the patients' perspective on PC has been called for internationally.^{16 17} If we want to change healthcare towards being more patient-centred, it is crucial to know which dimensions matter most to patients.^{16 17} In order to foster and implement PC in routine practice, it is necessary to assess how patients actually experience the extent of realised PC.¹⁷⁻¹⁹

Patient-reported experience measures (PREMs) assess patients' experiences with processes in healthcare and could be a useful way forward regarding the measurement of PC. Contrary to patient-reported outcome measures (PROMs), PREMs did not yet receive considerable attention,^{20 21} especially in Germany. The measurement of satisfaction, which is a frequently used PROM, has traditionally produced ceiling effects. PREMs on the other hand aim to collect more information on whether specific aspects of healthcare delivery took place or not making them therefore more objective.²² Psychometrically sound PREMs are especially important for PC, as patients are often the best or only source to assess to which extent healthcare is patient-centred.²³ Similar to PROMs, standardised and validated PREMs can be used in multiple ways: as performance, benchmarking or quality improvement measures

in health systems or healthcare provider organisations, in clinical trials, as well as in routine clinical practice for informing patients and clinicians.^{24 25} Many countries and institutions have highlighted the importance of assessing patients' experiences. For example, in the UK, Scandinavian countries and the USA, patients' experiences are used to guide reimbursement of hospitals.²⁶⁻²⁹ Despite the large amount of international literature on measurement of PC, available measures display substantial variation due to disparity of underlying conceptualisations of PC.^{9 30} Furthermore, systematic reviews on the measurement of different dimensions of PC have revealed flaws in existing measures and a lack of psychometrically sound measures in German language.³¹⁻³⁵ For example, measures were too long or psychometrically insufficient for valid comparisons and for a precise description of PC.³¹⁻³⁵ Thus, it is important to develop and psychometrically test a core set of PREMs applicable to the German healthcare context. This will allow assessing the degree of PC of current German healthcare for the first time. Additionally, those results could be used internationally enriching the status quo of research of PC from the patients' perspective.

Objectives

This study aims (1) to assess content and relevance as well as the current implementation of dimensions of PC from the patients' perspective, (2) to develop and psychometrically test a core set of PREMs assessing PC and (3) to investigate the feasibility of implementation of the developed core set of PREMs in routine healthcare.

METHODS AND ANALYSIS

Study population

PC has been described as especially important in the large and heterogeneous population of patients with chronic conditions.^{36 37} The study will therefore focus on the perspectives of patients with chronic conditions from four disease groups: cancer, cardiovascular diseases, mental disorders and musculoskeletal disorders. Adult patients with conditions from the respective disease groups who have experienced treatment (outpatient and/or inpatient) in the past will be included. Patients with severe cognitive impairment and without sufficient command of the German language will be excluded. No other inclusion or exclusion criteria will be applied.

Study design and data analyses

This 5-year mixed-methods study comprises three phases. Each phase examines one of the objectives described above. Figure 2 gives an overview.

Phase 1: assessment of patients' views on dimensions of PC

This phase aims at identifying which dimensions of PC are relevant to patients.

Methods, procedures and instruments

As a preparation for the relevance assessment, the first step will be to translate the integrative model of PC¹⁴

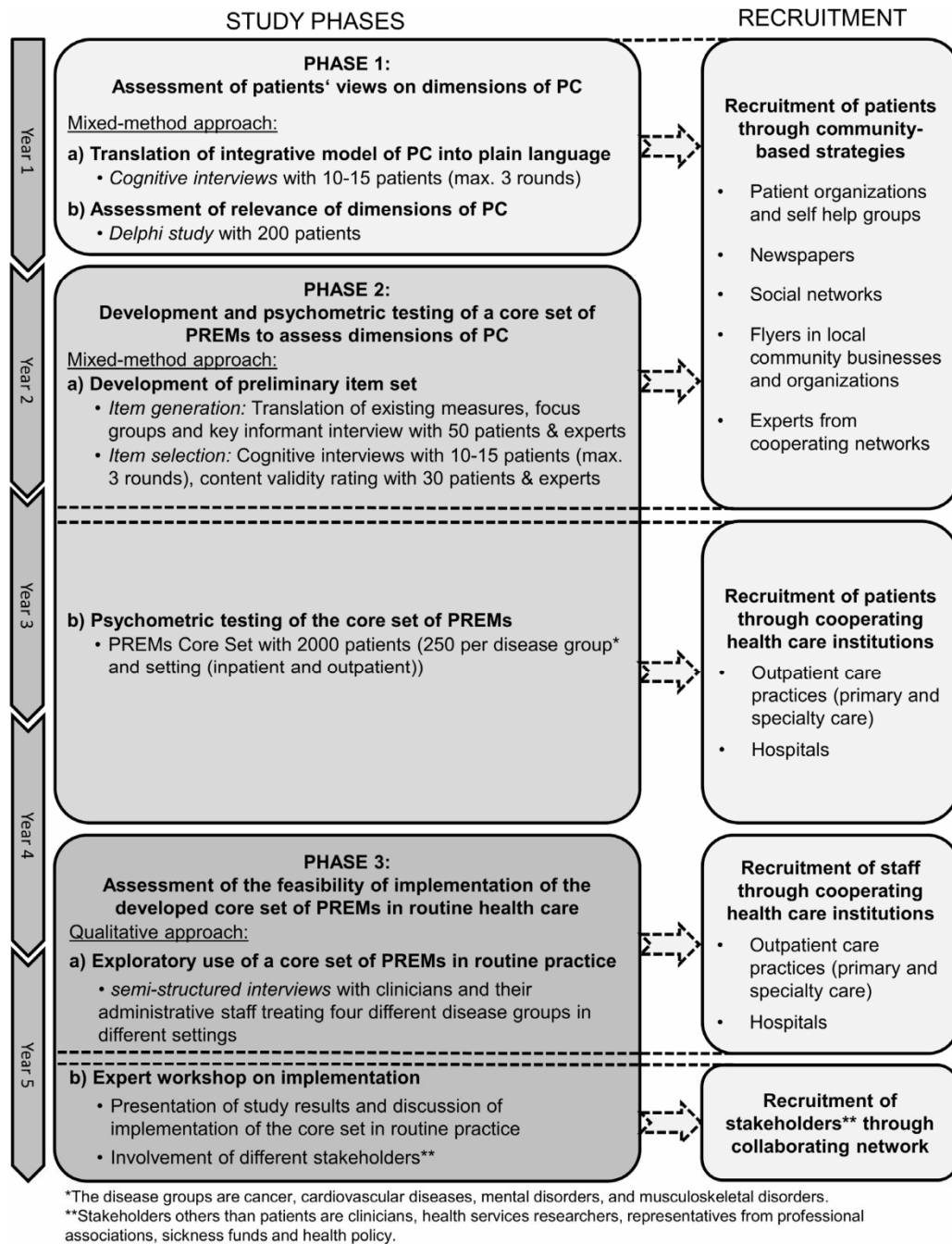


Figure 2 Overview of study design and patient recruitment strategies. PC, patient centredness; PREMs, patient-reported experience measures.

into a plain language version to make it more easily comprehensible for patients. Two researchers will independently translate the integrative model of PC into a plain language version. Next, they will generate a consensus version through discussion with two other members of the research team. Then, we will use cognitive interviewing with patients to test comprehensibility of the plain language model. Participants will be asked to read the plain language model step by step and to paraphrase in their own words what they have read. The cognitive interviews will be audio-recorded, transcribed and analysed independently by two researchers

regarding misunderstandings as well as expressed difficulties with the text and suggestions for improvement. Based on the results, the plain language version will be revised in a group discussion between three researchers. Depending on the extent of revision necessary, a further round of cognitive interviews might be conducted, with a maximum of three rounds.

To assess the patient-reported relevance of the different dimensions of PC, a Delphi study will be conducted. Relevance in this case means the subjective importance of the different dimensions from the patients' view. In addition, we will assess to what extent those different

aspects are already implemented in the German health-care system according to the patients' subjective view. In a first round, participating patients will receive the revised plain language model as well as a questionnaire to rate the relevance and implementation of each dimension. In a second round, they will receive the median ratings of all participants as well as their own ratings from the first round. They will again be asked to rate the relevance and implementation. This procedure makes it possible to foster exchange and consensus within a large group of patients. The anonymity of the Delphi method helps to give all opinions an equivalent weight without dominance of certain persons, as might be the case in the personal contact within focus groups.^{38 39}

For all patient participants, we will collect demographic data (eg, age, gender) with a short questionnaire. Furthermore, participants of the cognitive interviews will be asked to answer the European Health Literacy Questionnaire HLS-EU-Q16. This questionnaire was developed and used in The European Health Literacy Project (HLS-EU) in eight different European countries.^{40 41} The German version was tested with n=4952 respondents. It showed an internal consistency of Cronbach's $\alpha=0.9$. The questionnaire is used, because health literacy might influence the ability to comprehend the plain language version. Thus, we want to make sure that we included some patients with a low health literacy as well.

Recruitment/sample size

Patients for phase 1 will be recruited through a range of community-based strategies (including media campaigns using newspaper announcements and flyers, as well as recruitment through patient organisations throughout Germany). This approach was chosen, as it has been shown to be less resource intensive than recruitment through healthcare institutions.^{42–44}

For each round of cognitive interviews, approximately 10 to 15 participants are expected to be necessary to reach theoretical saturation.^{45 46}

For the Delphi study, we aim to collect data from a total sample of 200 participants. This allows for calculating descriptive parameters with high precision for the total sample and warrants a sufficient precision of estimates for disease subgroups as well. To reach 200 participants in the second round, more participants need to be included in the first round. Former Delphi studies showed dropout rates from round 1 to round 2 ranging from 10% to 30%.^{15 47–49} We expect a rather small dropout rate of approximately 15%, as patients will be reminded personally, and patients who participate in both rounds will receive a financial compensation. We therefore aim to recruit 236 patients for the first round.

Data analysis

We will analyse all sociodemographic data, health literacy scores and ratings from the Delphi survey descriptively (eg, frequencies, means, SD) for all participants as

well as for subgroups according to the diseases of the participants.

Phase 2: development and psychometric testing of a core set of PREMs to assess dimensions of PC

Phase 2 aims to develop a core set of PREMs and test its psychometric properties. We will describe both steps separately in the following paragraphs.

Phase 2a: Development of preliminary item set

Methods, procedures and instruments

Based on the relevance ratings of phase 1, items for the core set of PREMs will be developed using multiple sources. First, we will conduct a literature search to identify existing international PREMs on PC. This search will include an electronic search as well as scanning of existing reviews on measures that assess dimensions of PC.^{32 34 35 50} We will translate suitable items into German using a team translation process. This process consists of the following steps: (1) translation from original language to German by two independent translators, (2) suggestion of consensus version by third person and (3) final consensus team discussion. Furthermore, we will conduct focus groups and key informant interviews with patients as well as with experts (eg, clinicians, health services researchers, representatives from professional associations, sickness funds and health policies). Participants will be asked to generate items, which are suitable to measure PC based on the model from phase 1.

All generated items will be tested using cognitive interviews with patients and revised if necessary to ensure that all items are comprehensible for the target population. Again, in this step, we will use the HLS-EU-Q16 to check the variance of health literacy within our sample.

In a second step, the number of items will be reduced if necessary. Experts and patients will be asked to rate the relevance of all generated items for measuring PC. In addition, they will be asked whether any important questions are missing. This allows to establish content validity by ensuring that all items are relevant as well as that they are reflecting PC comprehensively.⁵¹ Items rated as not relevant or less relevant will be dropped.

Recruitment/sample size

We plan to include 50 participants (patients and experts) in the focus groups and key informant interviews for item generation. For cognitive interviewing, we will recruit between 10 and 15 patients per round, with a maximum of three rounds. Content validity will be assessed by 5 to 10 experts as well as 5 to 10 patients, which corresponds to common recommendations.^{46 52 53}

Again, we will recruit patients via community-based strategies in the Hamburg area. Experts other than patients will be invited from our network of collaborating project partners and from the large regional Network for Healthcare Research Hamburg (HAM-NET; <http://ham-net.de/en/>).

Data analysis

The key informant interviews and focus groups will be recorded, transcribed and analysed via qualitative content analysis.⁵⁴ Two researchers will conduct the analyses. Regular meetings with a third researcher will be held to ensure that coding schemes and analyses are comparable.

For content validity rating, a four-point Likert scale (1=not relevant, 2=somewhat relevant, 3=quite relevant and 4=very relevant) will be used.^{46 52 53} We will examine frequencies and median of the content validity ratings. Furthermore, we will calculate the Content Validity Index^{46 52 53} to explore inter-rater agreement of the relevance ratings for each item.

Phase 2b: psychometric testing

Methods, procedures and instruments

After the development of the preliminary item set, its psychometric characteristics will be tested. Patients will be asked to fill in a survey. The survey will include the developed items as well as selected validated measures that test related constructs to establish construct validity. For example, we will use the first item of the German version of 12-Item Short Form Survey⁵⁵ to explore discriminant validity. This item explores the general perception of health status and is expected to show a low correlation <0.3 .²⁷ For convergent validity, we will for example assess satisfaction with the treatment by using the German version of the 8-item Client Satisfaction Questionnaire (ZUF-8).⁵⁶ For satisfaction, we expect a high correlation >0.5 .²⁷

The decision whether further measures will be used to test construct validity will be part of the preparation for phase 2b and can only be determined once the content of the item set emerges from the data. The core set of PREMS will be revised according to the results of phase 2b. We will delete items according to common thresholds in floor and ceiling effects, intraclass correlations or Kappa and loadings in factor analysis (eg, items with more than 20% of the replies at one end of the scale and items with intraclass correlations or Kappa coefficients below 0.40⁵⁷).

Recruitment/sample size

Patients will be recruited in healthcare institutions in the metropolitan area of Hamburg, Germany (eg, academic cancer centre, academic heart centre, psychosomatic clinic, primary care practices). Consecutive sampling will be used in participating outpatient and inpatient healthcare institutions of primary and secondary care. This way, only patients that are currently being treated in a specific healthcare institution will be included in this study phase.

To examine the factorial structure of items and to perform analyses of associations among variables with sufficient precision and power, an approximate sample size of 250 participants per group will be targeted. Overall, there will be eight groups as we collect data from each of the four disease groups in two settings (inpatient and outpatient). Therefore, we aim for a

total sample size of 2000 participants. This allows for the use of maximum likelihood or weighted least squared estimators in the investigation of moderately complex models and for the detection of weak but meaningful correlations ($r=0.20$) with an alpha error probability of $\alpha=0.05$ and a power of 0.90. The exact methods of assessment, statistical analyses and exact necessary sample sizes will be determined once phase 2a is completed, and the exact number of items and measures of the core set has been decided on.

Data analysis

We will calculate several psychometric indices. We will examine internal consistency using Cronbach's α or the Kuder Richardson 20 coefficient (depending on the response format of items). For further item reduction and assessment of structural validity, we will analyse item characteristics (discrimination and difficulty) and perform exploratory factor analysis. To assess construct validity, we will calculate correlations. We will use t-tests and Mann-Whitney U-tests (depending on the distribution of the outcome) to assess discriminant validity. We will examine floor and ceiling effects using descriptive analyses by comparing the distribution of the observed values to the theoretically possible range of values. As a proxy of acceptability, we will examine if there are systematic patterns of missing data or low response rates.⁵⁸ We do not examine retest reliability and inter-rater reliability, as we do not expect the measured construct to remain stable over time or between people.

Phase 3: assessment of the feasibility of implementation of the developed core set of PREMs in routine healthcare

In phase 3, we aim to investigate the feasibility of implementation of the developed core set of PREMs in routine healthcare in a phase of exploratory use.

Methods, procedures and instruments

Collaborating healthcare organisations will be offered to use the final core set of PREMs in their routine practice. As a first step, onsite meetings with staff of participating healthcare institutions will be held. Those have the purpose to come to a shared understanding on how to best administer the core set of PREMs within the current clinical workflow of the respective institution. In these meetings, we will, for example, discuss who should give the core set to the patients, at what time point exactly and how the replies should be collected. In addition, the participating institutions will be able to decide whether they want to receive the results of the PREMs core set in order to use them as performance feedback. These onsite meetings will be followed by a phase of exploratory use of several weeks. After this phase of exploratory use, we will conduct interviews. Specifically, we will ask the participating staff how they organised the implementation and whether they made adaptations to better fit the setting requirement. Furthermore, we will ask which barriers and facilitators they experienced and which steps need to be

taken in order to further use the PREMs core set in their routine care.

As a second step of phase 3, we will conduct an expert workshop on implementation of the core set of PREMs in routine care. This workshop aims at (1) presenting the results of the psychometric testing and the assessment of PC in routine practice, and (2) discussing opportunities for the implementation of the core set in routine practice after testing it in the phase of exploratory use. This discussion aims to bring together the results from the semistructured interviews in the pilot phase as well as the knowledge and experience of the participating experts. It will cover both the implementation at individual healthcare organisations as well as the implementation as a potential quality indicator for purposes of accountability and quality improvement at the health system level.²⁵

Recruitment/sample size

For the pilot test, we aim to interview 10 to 15 clinicians and administrative staff from our collaborating healthcare organisations where the core set of PREMs was tested.

We plan to conduct the workshop with 25 to 30 experts. This will allow inviting a range of experts with different backgrounds and areas of expertise (eg, clinicians, health services researchers, representatives from professional associations, sickness funds and health policy). At the same time, it is a sample size feasible to allow the realisation of discussions in small groups and in the full group. They will be invited from our network of collaborating project partners and HAM-NET.

Data analysis

The interviews and focus groups will be recorded, transcribed and analysed via qualitative content analysis.⁵⁴ The analyses will be conducted by two researchers. Regular meetings with a third researcher will be held to ensure that coding schemes and analyses are comparable.

Patient and public involvement

Patients were not involved in the development of the research questions. However, during the development of the research proposal and prior to submission to the funding agency, we obtained collaboration agreements with several federal and regional patient organisations in order to secure field access and feasibility. Patient organisations were approached for that purpose. All gave us positive feedback on the study aims, acknowledging that more research is needed on patients' experiences related to patient-centred care. Thus, patient organisations will support recruitment of study participants by disseminating advertisement for study participation. No individual patient will be involved in recruitment and conduct of the study. All study participants and every interested person in the public will have the possibility to read and download regular project updates and study results on the project website (<http://www.ham-net.de/de/projekte/projekt-aspired.html>).

Software

We will use the software IBM SPSS Statistics (IBM Corp., Armonk, New York, USA) for most quantitative data analyses. Additionally, we will perform structural equation modelling with Mplus (Muthén & Muthén, Los Angeles). To support the transcription of focus groups and interviews, we will use the software F4 transkript (dr. dresing & pehl GmbH, Marburg). We will use the software MAXQDA (VERBI GmbH, Berlin) to support all qualitative data analyses.

ETHICS AND DISSEMINATION

Ethical and safety considerations

The study will be carried out according to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good scientific practice will be respected. Study participation is voluntary and no foreseeable risks for participants result from the participation in this study. Participants will be fully informed about the aims of the study, data collection and the use of collected data. Written informed consent will be sought prior to participation. Preserving principles of data sensitivity, data protection and confidentiality requirements will be met.

Dissemination plan

During the course of the study, a biannual newsletter will be made available to study participants, collaboration partners and the interested public to inform them about the progress of the study and its results. This newsletter will be made available for download on the project website. The results of the project will be published in scientific journals. To ensure high accessibility, we aim to publish our work in open access journals, if feasible. Furthermore, dissemination of the project's results will also be part of the implementation workshop of phase 3, which will include different stakeholders. Additionally, the results will be presented at relevant national and international conferences. Finally, the resulting core set of PREMs assessing PC will be made available for download free of charge under a Creative Commons License.

Acknowledgements We thank our student assistants Sophia Schulte and Tanja Kloster who supported us in preparing the manuscript. We are thankful for the support of many national and regional patient organisations: Bundesselbsthilfeverband für Osteoporose, Deutsche Fibromyalgie Vereinigung, Deutsche Gesellschaft für Prävention und Rehabilitation von Herz-Kreislauferkrankungen, Frauenselbsthilfe nach Krebs, Haus der Krebs-Selbsthilfe, Kontakt- und Informationsstellen für Selbsthilfegruppen in Hamburg, Landesverband Psychiatrie-Erfahrener, Nationale Kontakt- und Informationsstelle zur Anregung und Unterstützung von Selbsthilfegruppen und Rheuma-Liga. Furthermore, we would like to thank our collaborating health care institutions: Landesverband Hamburg der Deutschen Gesellschaft für Psychosomatische Medizin und Ärztliche Psychotherapie, OptiMedis AG, RehaCentrum Hamburg/Klinikum Bad Bramstedt, Schön Klinik Eilbek and the University Medical Center Hamburg-Eppendorf (Center for Health Care Research, Department of General Practice/Primary Care, Department of Psychiatry and Psychotherapy, Taskforce Patient-Centeredness, University Cancer Center and University Heart Center). Finally, we would like to thank further individuals for their collaboration and support: Paul Barr (Dartmouth College, USA), Debra de Silva (The Evidence Center), Marie-Luise Dierks (Hannover Medical School), Ed Harding (The Health Policy

Partnership), David Klemperer (Ostbayerische Technische Hochschule Regensburg), Christoph Kranich (Consumer Association Hamburg), Jürgen Matzat (Kontaktstelle für Selbsthilfegruppen der Deutschen Arbeitsgemeinschaft Selbsthilfegruppen e.V.), Hardy Müller (Wissenschaftliches Institut der TK für Nutzen und Effizienz im Gesundheitswesen) and Edmund Neugebauer (University Witten/Herdecke).

Contributors EC and SZ contributed to the specification of the study design and wrote the first draft of the manuscript in equal shares. IS, PH, LK and MH conceptualised and designed the study, wrote the grant proposal and obtained funding. They also critically revised the manuscript for important intellectual content. IS is the responsible primary investigator of the project. All authors gave final approval of the version to be published and agree to be accountable for the work.

Funding This work is funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung–BMBF) with grant number 01GY1614.

Disclaimer The funder reviewed and approved the study design during the grant application process, but had no role in decision to publish or preparation of the manuscript.

Competing interests None declared.

Patient consent Not required.

Ethics approval Ethics Committee of the Medical Association Hamburg, Germany (study ID: PV5724).

Provenance and peer review Not commissioned; peer reviewed for ethical and funding approval prior to submission.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev* 2013;70:351–79.
- Robinson JH, Callister LC, Berry JA, et al. Patient-centered care and adherence: definitions and applications to improve outcomes. *J Am Acad Nurse Pract* 2008;20:600–7.
- Castro EM, Van Regenmortel T, Vanhaecht K, et al. Patient empowerment, patient participation and patient-centeredness in hospital care: a concept analysis based on a literature review. *Patient Educ Couns* 2016;99:1923–39.
- Berwick DM. A user's manual for the iom's 'quality chasm' report. *Health Aff* 2002;21:80–90.
- Leavitt M. Medscape's response to the institute of medicine report: Crossing the quality chasm: A new health system for the 21st century. *MedGenMed* 2001;3:233–5.
- Härter M, Dirmaier J, Scholl I, et al. The long way of implementing patient-centered care and shared decision making in Germany. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:46–51.
- Bundestag. Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten: Bundesgesetzblatt, 2013:277–82. 2013. https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&start=//.
- Härter M, Moumjid N, Cornuz J, et al. Shared decision making in 2017: International accomplishments in policy, research and implementation. *Z Evid Fortbild Qual Gesundhwes* 2017;123–124:1–5.
- Epstein RM, Franks P, Fiscella K, et al. Measuring patient-centered communication in patient-physician consultations: theoretical and practical issues. *Soc Sci Med* 2005;61:1516–28.
- Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000;51:1087–110.
- Ouwens M, Hermens R, Hulscher M, et al. Development of indicators for patient-centred cancer care. *Support Care Cancer* 2010;18:121–30.
- Stewart MA, Brown JB, Weston W, et al. *Patient-centered medicine - transforming the clinical method*. Abingdon: Radcliffe Medical Press, 2003.
- van Dulmen S. Patient-centredness. *Patient Educ Couns* 2003;51:195–6.
- Scholl I, Zill JM, Härter M, et al. An integrative model of patient-centeredness - a systematic review and concept analysis. *PLoS One* 2014;9:e107828.
- Zill JM, Scholl I, Härter M, et al. Which dimensions of patient-centeredness matter? - Results of a web-based expert delphi survey. *PLoS One* 2015;10:e0141978.
- Bensing J, Rimondini M, Visser A. What patients want. *Patient Educ Couns* 2013;90:287–90.
- Hudon C, Fortin M, Haggerty J, et al. Patient-centered care in chronic disease management: a thematic analysis of the literature in family medicine. *Patient Educ Couns* 2012;88:170–6.
- Luxford K. What does the patient know about quality? *Int J Qual Health Care* 2012;24:439–40.
- Härter M, Müller H, Dirmaier J, et al. Patient participation and shared decision making in Germany - history, agents and current transfer to practice. *Z Evid Fortbild Qual Gesundhwes* 2011;105:263–70.
- Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167–7.
- Rolfson O, Rothwell A, Sedrakyan A, et al. Use of patient-reported outcomes in the context of different levels of data. *J Bone Joint Surg Am* 2011;93 :66–71.
- Garratt AM, Bjaertnes ØA, Krogstad U, et al. The outpatient experiences questionnaire (opeq): Data quality, reliability, and validity in patients attending 52 norwegian hospitals. *Qual Saf Health Care* 2005;14:433–7.
- Anhang Price R, Elliott MN, Zaslavsky AM, et al. Examining the role of patient experience surveys in measuring health care quality. *Med Care Res Rev* 2014;71:522–54.
- Nelson EC, Eftimovska E, Lind C, et al. Patient reported outcome measures in practice. *BMJ* 2015;350:g7818.
- Patwardhan A, Spencer CH. Are patient surveys valuable as a service-improvement tool in health services? An overview. *J Healthc Leadersh* 2012;4:33–46.
- Department of Health. The NHS Outcomes Framework 2011 / 12. *Framework*, 2010:56.
- Olteal S, Garratt A, Bjertnes Ø, et al. The NORPEQ patient experiences questionnaire: data quality, internal consistency and validity following a Norwegian inpatient survey. *Scand J Public Health* 2007;35:540–7.
- Giordano LA, Elliott MN, Goldstein E, et al. Development, implementation, and public reporting of the HCAHPS survey. *Med Care Res Rev* 2010;67:27–37.
- Szablowski KM. Hospital Value-Based Purchasing (VBP) program: measurement of quality and enforcement of quality improvement. *Conn Med* 2014;78:49–51.
- McCormack LA, Treiman K, Rupert D, et al. Measuring patient-centered communication in cancer care: a literature review and the development of a systematic approach. *Soc Sci Med* 2011;72:1085–95.
- Barr PJ, Scholl I, Bravo P, et al. Assessment of patient empowerment--a systematic review of measures. *PLoS One* 2015;10:e0126553.
- Müller E, Zill JM, Dirmaier J, et al. Assessment of trust in physician: a systematic review of measures. *PLoS One* 2014;9:e106844.
- Scholl I, Frerichs W, Christalle E, et al. Assessment of patient information needs: a systematic review of measures. *PLoS One* 2015;10:e0126553.
- Scholl I, Koelwijn-van Loon M, Sepucha K, et al. Measurement of shared decision making - a review of instruments. *Z Evid Fortbild Qual Gesundhwes* 2011;105:313–24.
- Zill JM, Christalle E, Müller E, et al. Measurement of physician-patient communication-a systematic review. *PLoS One* 2014;9:e112637.
- Holman H, Lorig K. Patients as partners in managing chronic disease. Partnership is a prerequisite for effective and efficient health care. *BMJ* 2000;320:526–7.
- McMillan SS, Kendall E, Sav A, et al. Patient-centered approaches to health care: a systematic review of randomized controlled trials. *Med Care Res Rev* 2013;70:567–96.
- Häder M. *Delphi-Befragungen: Ein Arbeitsbuch*: Springer VS, 2014.
- Boulkedid R, Abdoul H, Loustau M, et al. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011;6:e20476.
- Pelikan J, Ganahl K. Die europäische Gesundheitskompetenz-Studie: Konzept, Instrument und ausgewählte Ergebnisse. 2017.
- Jordan S, Hoebel J. Gesundheitskompetenz von erwachsenen in deutschland. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2015;58:942–50.
- Brewster WR, Anton-Culver H, Ziegas A, et al. Recruitment strategies for cervical cancer prevention study. *Gynecol Oncol* 2002;85:250–4.
- Brill SE, El-Emir E, Allinson JP, et al. Community-based recruitment of patients with COPD into clinical research. *Thorax* 2014;69:951–2.



44. Davey R, Edwards SM, Cochrane T. Recruitment strategies for a clinical trial of community-based water therapy for osteoarthritis. *Br J Gen Pract* 2003;53:315–7.
45. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. *Public Opin Q* 2007;71:287–311.
46. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. 2015.
47. Grotorex J, Dexter T. An accessible analytical approach for investigating what happens between the rounds of a Delphi study. *J Adv Nurs* 2000;32:1016–24.
48. Rognstad S, Brekke M, Fetveit A, et al. The Norwegian General Practice (NORGE) criteria for assessing potentially inappropriate prescriptions to elderly patients. A modified Delphi study. *Scand J Prim Health Care* 2009;27:153–9.
49. Nowack M, Endrikat J, Guenther E. Review of delphi-based scenario studies: quality and design considerations. *Technol Forecast Soc Change* 2011;78:1603–15.
50. De Silva D, 2014. Helping measure person-centred care: a review of evidence about commonly used approaches and tools used to help measure person-centred care www.health.org.uk/helpingmeasurepcc (accessed 12 Feb 2018).
51. 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;19:539–49.
52. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health* 2006;29:489–97.
53. Lynn MR. Determination and quantification of content validity. *Nurs Res* 1986;35:382–5.
54. Schreier M. Varianten qualitativer Inhaltsanalyse: ein wegweiser im dickicht der Begrifflichkeiten. *Forum Qual Sozialforsch* 2014;15.
55. Wirtz MA, Morfeld M, Glaesmer H, et al. Konfirmatorische Prüfung der Skalenstruktur des SF-12 Version 2.0 in einer deutschen bevölkerungs-repräsentativen Stichprobe. *Diagnostica* 2018;64:84–96.
56. Kriz D, Nübling R, Steffanowski A, et al. Patientenzufriedenheit in der stationären Rehabilitation: psychometrische reanalyse des zuf-8 auf der basis multizentrischer stichproben verschiedener indikation. *Zeitschrift für Medizinische Psychol* 2008;17:67–79.
57. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000;81:S15–S20.
58. Sepucha KR, Matlock DD, Wills CE, et al. "It's Valid and Reliable" Is Not Enough: critical appraisal of reporting of measures in trials evaluating patient decision aids. *Med Decis Making* 2014;34:560–6.

9.2. Publication 2

Reference: Zeh, S., **Christalle, E.**, Hahlweg, P., Härter, M., & Scholl, I. (2019). Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ Open*, 9(12), e031741.

BMJ Open Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study

Stefan Zeh , Eva Christalle , Pola Hahlweg, Martin Härter, Isabelle Scholl

To cite: Zeh S, Christalle E, Hahlweg P, *et al.* Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ Open* 2019;9:e031741. doi:10.1136/bmjopen-2019-031741

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-031741>).

Received 21 May 2019

Revised 13 November 2019

Accepted 27 November 2019

ABSTRACT

Objective Patient-centredness (PC) has particularly grown in relevance in health services research as well as in politics and there has been much research on its conceptualisation. However, conceptual work neglected the patients' perspective. Thus, it remains unclear which dimensions of PC matter most to patients. This study aims to assess relevance and current degree of implementation of PC from the perspective of chronically ill patients in Germany.

Methods We conducted a Delphi study. Patients were recruited throughout Germany using community-based strategies (eg, newspapers and support groups). In round 1, patients rated relevance and implementation of 15 dimensions of PC anonymously. In round 2, patients received results of round 1 and were asked to re-rate their own results. Participants had to have at least one of the following chronic diseases: cancer, cardiovascular disease, mental disorder or musculoskeletal disorder. Furthermore, patients had to be at least 18 years old and had to give informed consent prior to participation.

Results 226 patients participated in round 1, and 214 patients in round 2. In both rounds, all 15 dimensions were rated highly relevant, but currently insufficiently implemented. Most relevant dimensions included 'patient safety', 'access to care' and 'patient information'. Due to small sizes of subsamples between chronic disease groups, differences could not be computed. For the other subgroups (eg, single disease vs multi-morbidity), there were no major differences.

Conclusion This is one of the first studies assessing PC from patients' perspective in Germany. We showed that patients consider every dimension of PC relevant, but currently not well implemented. Our results can be used to foster PC healthcare delivery and to develop patient-reported experience measures to assess PC of healthcare in Germany.

INTRODUCTION

Over the last years, the concept of patient-centredness (PC) has been widely discussed. More and more, the terms person-centredness^{1 2} and people-centredness³ also appear in the literature, either being used synonymously to PC or converging around central ideas. Its relevance in politics, research

Strengths and limitations of this study

- No rating of dimensions of patient-centredness from the patients' perspective available yet.
- Extensive inclusion of patients' perspectives within the Delphi study.
- No restriction to a specific regional area through nation-wide community-based strategies.
- Possibility for participants to give written feedback in order to enhance quality of given answers.
- Limitation to only four chronic disease groups.

and healthcare practice has been growing constantly.⁴⁻⁶ In Germany, the Law on Patients' Rights⁵ mandates certain aspects of PC care such as the right for comprehensible and comprehensive patient information. However, there is no national recommendation for the implementation of PC care in Germany yet. Many studies showed a positive association of PC with relevant outcomes such as patient satisfaction, well-being and adherence,⁷ which are especially relevant in chronic disease management. Nevertheless, there has been a lack of conceptual clarity regarding the use of the term of PC.⁸⁻¹³ Scholl and colleagues¹² closed this gap by developing an integrative generic model derived from literature extracting 15 dimensions of PC (see figure 1). This model was validated by assessing the views of different healthcare stakeholders on its relevance and clarity.¹⁴ Although some representatives of patient organisations participated in this assessment, the perspective of the recipients of PC care, the patients themselves, was under-represented. In summary, without the patients' perspective, the conceptualisation of PC is still missing its most crucial aspect: the assessment and evaluation of PC in healthcare from the view of patients.¹⁵⁻¹⁷

There are only few international studies investigating the patients' assessment of PC,^{18 19} without any published studies from Germany.



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Correspondence to

Stefan Zeh; s.zeh@uke.de

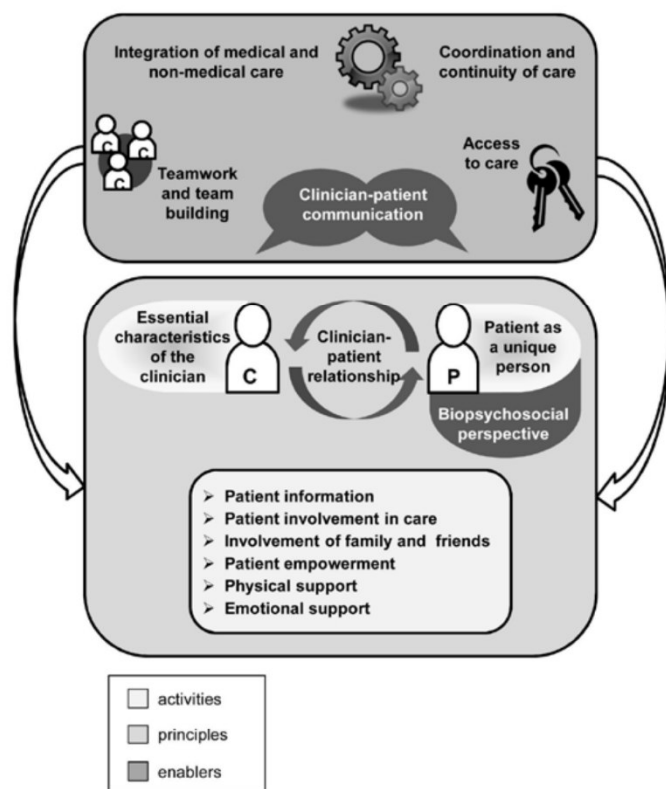


Figure 1 Integrative model of patient-centredness.

Thus, we do not know what patients consider relevant about PC, nor how patients view the current degree of implementation of PC in the German healthcare system. It is a system in which healthcare coverage is universal, mandatory and offers extensive services. Government delegates decision-making powers to self-regulated organisations of payers and providers. While outpatient care is mainly delivered in one-physician or two-physician practices on a fee-for-service payment model, hospitals mainly deliver inpatient care on case-based payment model.^{20 21} Yet, in order to foster PC in Germany, it is crucial to know what aspects of PC are relevant to patients and how well these aspects are already implemented, as it allows prioritising aspects especially relevant or not yet implemented.

Therefore, this study aims to close this gap by assessing relevance and current degree of implementation of different dimensions of PC from the patients' perspective in Germany. Assessing relevance and current degree of implementation of different aspects of PC can therefore be seen as the first steps in achieving a more PC healthcare.

This Delphi study was conducted as part of the study 'Assessment of patient-centeredness through patient-reported experience measures (ASPIRED).²² Within the study ASPIRED, a patient-reported experience measure will be developed to foster PC care.

METHODS

Study design

We used a Delphi study, comparable with the expert validation study mentioned earlier.¹⁴ A Delphi study consists of several rounds. In each round after the first, participants

get their own ratings as well as the median ratings of all participants in the previous round. This procedure generates the possibility for participants to share their opinion from different regional locations without the threat of some participants dominating the discussion.²³ It aims to reach high group consensus through a structured and anonymous group discussion.²⁴ Like Zill and colleagues¹⁴ or van Rijssen and colleagues,²⁵ we planned to perform two rounds prior to conducting the first round. However, if we had seen major discrepancies after two rounds, we would have performed another round, if necessary.

Preparation of the Delphi study: adaptation of the model of PC using cognitive interviews

We based our Delphi study on the integrative model of PC.^{12 14} In the first step, the scientific model was adapted into a version comprehensible to patients (first version). Adaptation was done by group discussions in our research team (SZ, EC, PH, IS) in nine rounds. To make sentences easy to understand, we used dictionaries and recommendations on adaptation of wording and grammar.²⁶ We then tested comprehensibility of the new version through cognitive interviews with patients. We recruited 10 patients through several community-based strategies (see section Delphi study) in order to get a heterogeneous sample regarding demographic variables (eg, age, first language, educational level). Participants had to have at least one of the following chronic diseases: cancer, cardiovascular disease, mental disorder or musculoskeletal disorder. In cognitive interviews, we handed out the first adapted version described earlier. Participants were asked to read each dimension step by step and explain in their own words how they understood the description. All interviews were audio-recorded, transcribed and analysed independently by two researchers (SZ, EC). According to the results, the model was revised (second version) in subsequent group discussions (SZ, EC, IS, PH).

Delphi study

Participants received the adapted integrative model of PC (second version) and were asked to rate relevance and current implementation from their perspective. For both aspects, a nine-point Likert scale (1=not relevant/not well implemented; 9=very relevant/well implemented) was given. In addition, we asked an open-ended question about whether any relevant aspects were missing in the model and whether they had any further comments (eg, regarding comprehensibility of descriptions). Based on answers to those open-ended questions, we adapted the model again (third version). In the second round of the Delphi study, patients received their own ratings from the first round again. Furthermore, we gave them the median rating of other participants in comparison. Following, we then asked participants to repeat their rating. For all dimensions with adapted descriptions, participants were also asked to rate whether the description worsened or improved with the adaptation (from 1=much worse to 9=much better).

For both rounds, participants could choose whether they wanted to participate online or via mail. Both versions were identical in terms of response format (see online supplementary appendix A1 for the mail-based version with adaptations (highlighted) made after round 1). For the online version, we used LimeSurvey.²⁷ The study was conducted from June 2018 to August 2018. For each round, participants had 3 weeks to reply and there were 4 weeks between the two rounds. After completion of both rounds, participants received a compensation of 25€. Conducting the Delphi study as well as reporting of methods and results were based on the recommendations of Hasson and colleagues²⁴ as well as of Jünger and colleagues.²⁸

Study population and recruitment

We aimed to include 200 patients in the study. We expected a drop-out rate of 15% between rounds 1 and 2. Thus, we aimed to include 240 in the first round. We recruited patients through different community-based strategies: (1) notices in local stores in Hamburg, (2) dissemination through national and regional patient organisations and networks, (3) Facebook groups of different towns and (4) two national newspapers (ZEIT and Bild). Participants could reach out to SZ and EC with the contact information provided on the notices and in emails. Since no member of the research team had direct contact with participants, the possibility of having influenced the participants' judgements was reduced. Participants were included if they gave their informed consent by signing the informed consent letter (mail version) or by choosing 'yes' in a drop-down menu after a question whether they consent (online version), were 18 years or older and if they reported at least one disease from one of

the aforementioned chronic disease groups. In figure 2, an overview illustrates the procedure.

Patient and public involvement

Patients were not involved in the development of the research questions. However, during the development of the research proposal and prior to submission to the funding agency, we obtained collaboration agreements with several federal and regional patient organisations in order to secure field access and feasibility. Patient organisations were approached for that purpose. All gave us positive feedback on the study aims, acknowledging that more research is needed on patients' experiences related to PC care. Thus, patient organisations supported recruitment of study participants by disseminating advertisement for study participation. No individual patient was involved in recruitment and conduct of the study. All study participants and every interested person in the public have the possibility to read and download regular project updates and study results on the project website (<http://www.ham-net.de/de/projekte/projekt-aspired.html>).

Data analyses

For the ratings, median, mean and SD were calculated for all participants using SPSS V.21.²⁹ Demographic data were analysed accordingly, except for nominal items where we calculated frequencies. Consensus was examined by calculating distributions of answers per tertile (1–3; 4–6; 7–9) in comparison with Zill *et al.*¹⁴ We compared different subgroups due to potential differences in using the health-care system including age groups (18–30 (young); 31–65 (middle aged); over 65 (older)), way of participation (mail vs online) and morbidity (single disease vs multi-morbidity). Open answers were analysed, summarised by

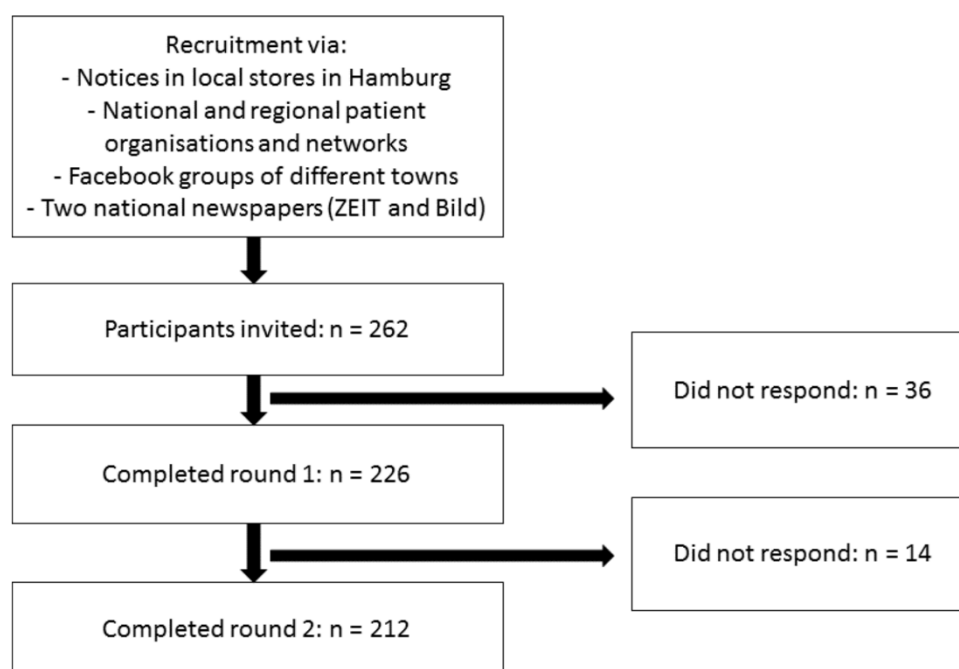


Figure 2 Flow chart for the Delphi study.

two researchers and the student assistant (SZ, EC, Tanja Kloster) and reviewed in a group discussion (SZ, EC, IS).

RESULTS

Adaptation of the integrative model into a comprehensible version

Within cognitive interviews, each participant understood generally well the patient version of PC (first version). Some participants wanted easier terms (eg, treatment ('Behandlung') instead of healthcare ('Gesundheitsversorgung')). We tested two versions: version A was not directed towards patients ('Patients can contact their healthcare provider easily'), and version B was directed towards patients ('You can contact your healthcare provider easily'). Nine out of 10 participants preferred personal pronouns (version B); thus, we adapted accordingly.

Delphi study

Characteristics of patients from the Delphi study

Two hundred and sixty-two patients, who reacted to the community-based recruitment strategies, fulfilled the inclusion criteria and were invited to participate via letter or online in the first round of the Delphi study after having expressed interest in study participation. In the first round 226 patients participated (response rate=86.26%), and in the second round 214 patients participated (final response rate=81.7%). Thus, the final sample consisted of 214 patients (drop-out rates: from invitation to first round: n=36 (13.7%); from invitation to second round: n=48 (18.3%); from first to second round: n=12 (5.3%)). Mean age was 52 years (normally distributed), and two-thirds were women. Further information on participants' demographic and clinical characteristics is displayed in table 1 and in online supplementary table S1.

Results of round 1

In table 2, results from the first round of the ratings of relevance and degree of current implementation are shown. Regarding ratings of relevance, 13 out of 15 dimensions showed the maximum median of 9. The dimensions 'patient as a unique person', 'clinician-patient relationship', 'clinician-patient communication', 'access to care' and 'physical support' showed highest means (range from 8.46 to 8.31). Consensus was defined as the degree to which patients rated dimensions in agreement and was predefined as reached if it was higher than 50%. It means that consensus was reached when at least 50% of the patients rated dimensions to be very or not very relevant (tertile from 7 to 9 or 1 to 3, respectively). Every dimension showed consensus of >50% for tertile 7-9 (ie, majority of patients rated within the highest tertile). Regarding the current degree of implementation, 3 out of 15 dimensions showed a median of 6 ('clinician-patient relationship', 'clinician-patient communication' and 'physical support'). All other dimensions showed a median of 5. The dimensions with the highest means were the same as those with the highest medians. Regarding consensus,

Table 1 Demographic and clinical characteristics of participants of the first round (n=226)

Characteristic	Frequency
Age (in years)	M=51.79 (SD=16.46)
Sex	
Female	149 (65.9%)
Male	76 (33.6%)
First language	
German	219 (96.9%)
Other	5 (2.2%)
Chronic disease group (multiple answers possible)	
Cancer	36 (15.3%)
Cardiovascular disease	90 (39.8%)
Mental disorder	117 (51.8%)
Musculoskeletal disorder	114 (50.4%)
Others*	149 (67.4%)
Residence	
Hamburg	59 (26.1%)
Baden-Wuerttemberg	13 (5.8%)
North-Rhine Westphalia	23 (10.2%)
Rhineland-Palatinate	3 (1.3%)
Saarland	3 (1.3%)
Saxony	13 (5.8%)
Saxony-Anhalt	4 (1.8%)
Schleswig-Holstein	10 (4.4%)
Thuringia	1 (0.4%)
Bavaria	22 (9.7%)
Berlin	14 (6.2%)
Brandenburg	11 (4.9%)
Bremen	4 (1.8%)
Hesse	14 (6.2%)
Mecklenburg-Western	19 (8.4%)
Pomerania	13 (5.8%)
Lower Saxony	
Health status†	M=2.49 (SD=0.77)

*Additional to one of the four groups mentioned earlier. Other chronic diseases included, for example, diabetes (n=12), sleep apnea (n=5) and hypothyroidism (n=4). Please see online supplementary S1 for a comprehensive overview of other diseases. †Health status was assessed with the item 'How would you assess your health status in general?' ranging from 1=excellent to 5=bad. M, mean.

every dimension had a rating of >50% for the tertile from 4 to 6. In general, every dimension of round 1 was rated as very relevant and as insufficiently implemented. Furthermore, patients were united in their rating because of the high consensus.

Sixty participants responded to the open-ended questions. Comments referred to additional aspects, including time issues within healthcare (eg, 'Neither nurses, nor

Table 2 Results for the assessment of relevance and implementation from the first round

	Relevance		Implementation	
	Median	M (SD)	Median	M (SD)
1. PC characteristics of healthcare providers	9	8.46 (0.90)	5	5.30 (1.44)
2. Trustful relationship	9	8.41 (0.94)	6	5.61 (1.72)
3. Uniqueness of each patient	9	8.31 (1.11)	5	5.21 (1.74)
4. Consideration of personal circumstances	9	8.15 (1.23)	5	4.77 (1.86)
5. Appropriate communication	9	8.31 (0.97)	6	5.56 (1.74)
6. Integration of additional healthcare elements	8	7.34 (1.81)	5	4.56 (1.99)
7. Teamwork of healthcare providers	9	8.29 (1.18)	5	5.08 (1.90)
8. Access to care	9	8.31 (0.99)	5	4.88 (2.08)
9. Good planning of care	9	8.30 (1.04)	5	5.25 (1.92)
10. Personally tailored information	9	8.27 (1.12)	5	5.10 (1.87)
11. Collaboration as equal partners and involvement in decision-making	9	8.25 (1.11)	5	5.15 (2.04)
12. Involvement of family and friends	7	6.92 (2.16)	5	4.91 (2.21)
13. Empowerment of patients	9	8.20 (1.13)	5	5.23 (1.97)
14. Support of physical well-being	9	8.31 (0.98)	6	5.58 (1.95)
15. Support of mental well-being	9	8.12 (1.35)	5	4.83 (1.98)

M, mean; PC, patient-centredness.

doctors have enough time for the patients'), financial aspects ('without additional costs for people living in poverty') and medication plans ('discussing medication plans is too fast' or 'more information about side effects'). We revised descriptions after the first round according to the most frequent comments (third version). German descriptions of dimensions used in round 2, in which changes from the second to third version are highlighted, are listed in online supplementary appendix A1.

A few patients stated that patient safety could be improved by better coordination of medication plans. Furthermore, patient safety was already depicted as part of the dimension 'physical support' in the model.¹⁴ Due to comments, we revised the model to add 'patient safety' as a separate 16th dimension.

Results of round 2

In round 2, most ratings were very similar to the first round and medians did not change (see table 3). The added dimension of patient safety was also considered relevant but currently insufficiently implemented. The adaptations we made based on the feedback were rated as 'better than before' (all means over 7). This time, most relevant dimensions were 'patient safety', 'teamwork and team building', 'access to care', 'patient information' and 'essential characteristics of clinicians'. 'Clinician-patient communication', 'clinician-patient relationship' and 'physical support' again reached the best ratings regarding implementation. Additionally, the newly added dimension 'patient safety' also reached equally high ratings.

Subgroup analyses

In every chronic disease group, distribution of patients participating via mail or online was similar (approximately 45:55 ratio, respectively). There was no major age

difference between mail-based and online-based participation (mean age (mail): 54.83 years (SD=17.85); mean age (online): 49.35 years (SD=14.88)). Due to small sizes of subsamples (eg, only six patients indicated to have cancer), differences in ratings depending on chronic disease groups could not be analysed (ie, via analysis of variance). Neither multi-morbidity (at least two chronic disease groups indicated compared with only one), nor type of participation (mail vs online) did affect assessment of relevance and implementation by patients. That is, there were no differences between means bigger than 1. The differences of means above 1 were found within different age groups. The implementation of 'access to care' was rated lower by young patients in comparison with older patients (mean=4.37 for young vs mean=5.38 for older patients). Implementation of 'patient information' was also rated lower by young patients (mean=4.70 for young vs mean=5.71 for older patients).

DISCUSSION

In our study, we have assessed relevance and current degree of implementation of different dimensions of PC in the German healthcare system from the patients' perspective. The findings show that patients consider every dimension of PC very relevant but currently only mediocly implemented. Furthermore, there were no evident factors influencing these results (eg, way of participation or multi-morbidity).

These results point out the relevance to ask patients when investigating PC. Comparing this assessment with the experts' ratings of the same dimensions from the same model,¹⁴ differences between expert and patient ratings become evident. For example, experts did not

Table 3 Results for the assessment of adaptation, relevance and implementation of the second round

	Adaptation (round 1 to round 2)			Relevance			Implementation		
	Median	M	SD	Median	M	SD	Median	M	SD
1. PC characteristics of healthcare providers	–	–	–	9	8.36	1.30	5	5.27	1.38
2. Trustful relationship	8	7.64	1.59	9	8.18	1.34	6	5.47	1.44
3. Uniqueness of each patient	–	–	–	9	8.30	1.24	5	5.17	1.50
4. Consideration of personal circumstances	8	7.14	1.80	9	8.07	1.36	5	4.83	1.61
5. Appropriate communication	–	–	–	9	8.33	1.12	6	5.62	1.54
6. Integration of additional healthcare elements	–	–	–	8	7.46	1.67	5	4.61	1.81
7. Teamwork of healthcare providers	–	–	–	9	8.45	1.04	5	5.08	1.75
8. Access to care	8	7.79	1.59	9	8.38	1.09	5	4.81	1.83
9. Good planning of care	8	7.28	1.82	9	8.18	1.28	5	5.07	1.73
10. Patient safety	–	–	–	9	8.56	0.94	6	5.61	1.64
11. Personally tailored information	8	7.86	1.43	9	8.37	1.02	5	5.03	1.82
12. Collaboration as equal partners and involvement in decision-making	–	–	–	9	8.32	1.12	5	5.25	1.74
13. Involvement of family and friends	–	–	–	7	6.85	2.01	5	4.92	1.91
14. Empowerment of patients	8	7.59	1.41	9	8.15	1.23	5	5.21	1.77
15. Support of physical well-being	8	7.40	1.74	9	8.18	1.26	6	5.55	1.71
16. Support of mental well-being	–	–	–	9	8.25	1.33	5	4.87	1.85

M, mean; PC, patient-centredness.

rate each dimension as very relevant but patients did (eg, involvement of family and friends which was less relevant to experts but patients considered it relevant). This discrepancy is relevant for development of strategies to foster implementation of PC in Germany. If only the experts' perspective is known, implementation efforts might not match patients' needs. Consequently, a more PC care can only be achieved by taking into account patients' perspectives.

Furthermore, the study reveals that while patients in Germany find PC extremely relevant, they do not find it well implemented in the German healthcare system. Thus, interventions to foster dimensions of PC on the micro level of care (eg, tools for clinicians and patients) could be developed and evaluated in future research. Furthermore, the results can guide the development of patient-reported experience measures. On the federal level, the results indicate that the German healthcare system could benefit from stronger health policy directives in regard to the delivery of PC care.

Strengths and limitations

Major strengths of this study are the diverse sample from all over Germany and the possibility to anonymously generate a group consensus by performing a Delphi study. By doing so, ratings of each individual were not affected by dominant people in group discussions, and therefore are more likely to be unbiased. Furthermore, another strength of this study are the response rates

indicating little drop out from round 1 to round 2. This could be ascribed to the internal motivation of participating patients willing to ameliorate healthcare services in Germany. Another explanation could be the financial compensation which they received only after having completed both rounds. However, although the notices were spread out widely, a participation bias (ie, only people with a higher interest in the topic would contact us) could have occurred by this method of recruitment. Furthermore, there was little variation within ratings of relevance of the dimensions of PC. This could indicate a lack of specification in the descriptions. Patients might have been unable to differentiate between subaspects of a single description. Thus, they might have rated the whole dimension as relevant even if they considered only some parts relevant enough to give a high score. Therefore, additional qualitative approaches (eg, by conducting semi-structured interviews) could help to enrich findings and further explain results by specifying relevant aspects instead of broad dimensions. Additionally, as our study focused on adding the patients' perspective to the concept of PC care, we based our study on the integrative model of PC, which integrates over 400 definitions and models published between 1969 and 2012.¹² We did not take into account further conceptual work published since (eg, the framework developed by Santana and colleagues³⁰) which can be considered a limitation of this study. Furthermore, it is unclear if the results are generalisable to other

countries. Since healthcare systems differ substantially between countries (eg, role of government, coverage, organisation of the delivery system, lengths of stay),^{20 31} patients' perspective on PCC might also differ. Further research is needed to evaluate the patients' perspective in other countries. Finally, another limitation of our study was the relatively low level of co-production with patients. As patients are recipients of PCC, this study could have benefitted from a higher level of patients' participation.³²

Future studies

Regarding future studies or specific recommendations, policy makers or other stakeholders, interested in fostering PC healthcare, should therefore not only focus on the ratings of relevance of experts but also consider the patients' rating. In their opinion, every dimension of PC is relevant. These results should not be neglected. When implementing PC care, a wholesome perspective is therefore needed.

CONCLUSIONS

In summary, our study provides the patients' perspective on PC by assessing relevance and current degree of implementation and therefore enriches current conceptualisations of PC. We showed that patients consider every dimension of PC relevant, but currently not well implemented. Furthermore, these results can be used for developing comprehensive questionnaires such as patient-reported experience measures regarding PC care. Regarding the status quo, especially in Germany, we therefore still need to make efforts since patients do not receive the PC healthcare they consider so relevant. Thus, these data can be used to deduce healthcare policies which tackle the insufficient implementation of PC from the patients' perspective.

Acknowledgements We thank our student assistant Tanja Kloster and our research intern Lara Dreyer for supporting us preparing the manuscript and transcribing the interviews.

Contributors IS is the responsible principle investigator of the study. IS, PH and MH were involved in planning and preparation of the study. SZ and EC recruited participants and collected data. SZ analysed the data. SZ, IS, EC and PH interpreted the results. SZ wrote the first draft of the manuscript. IS, EC, PH and MH critically revised the manuscript for relevant intellectual content. All authors gave final approval of the version to be published and agree to be accountable for the work.

Funding This study is funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) with the grant number 01GY1614.

Disclaimer The funder had no role in decision to publish or preparation of the manuscript.

Competing interests PH gave one scientific presentation on shared decision-making during a lunch symposium, for which she received compensation and travel compensation from GlaxoSmithKline GmbH (pharmaceutical company) in 2018.

Patient consent for publication Not required.

Ethics approval The study was carried out according to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good scientific practice were respected. The study had been approved by the Ethics Committee of the Medical Association Hamburg (study ID: PV5724) prior to the start of data collection. Study participation was voluntary and no foreseeable risks for participants resulted from the participation in this study. Participants were fully

informed about the aims of the study, data collection and the use of collected data. Written informed consent was obtained prior to participation. Preserving principles of data sensitivity, data protection and confidentiality requirements were met.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Stefan Zeh <http://orcid.org/0000-0002-1748-3155>

Eva Christalle <http://orcid.org/0000-0001-7813-5800>

REFERENCES

- 1 Louw JM, Marcus TS, Hugo JFM. Patient- or person-centred practice in medicine? - A review of concepts. *Afr J Prim Health Care Fam Med* 2017;9:e1-7.
- 2 Håkansson Eklund J, Holmström IK, Kumlin T, et al. 'Same same or different?' A review of reviews of person-centered and patient-centered care. *Patient Educ Couns* 2019;102:3-11.
- 3 World Health Organization (WHO). WHO framework on integrated people-centred health services. Available: <https://www.who.int/servicedeliverysafety/areas/people-centred-care/en/>
- 4 Härter M, Dirmmaier J, Scholl I, et al. The long way of implementing patient-centered care and shared decision making in Germany. *Z Evid Fortbild Qual Gesundheitswes* 2017;123-124:46-51.
- 5 Bundesrat. Gesetzesbeschluss des deutschen Bundestages. Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten (Patientenrechtegesetz), 2013. Available: http://www.bmjd.de/SharedDocs/Downloads/DE/pdfs/Gesetze/Gesetz_zur_Verbesserung_der_Rechte_von_Patientinnen_und_Patienten.pdf;jsessionid=1185B669A9CD4A090FDE13ABA75890D2.1_cid297?__blob=publicationFile [Accessed 04 Dec 2014].
- 6 Baumgart J. Ambivalentes Verhältnis. *Dtsch Arztebl* 2010;107:2554-7.
- 7 Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev* 2013;70:351-79.
- 8 Epstein RM, Franks P, Fiscella K, et al. Measuring patient-centered communication in patient-physician consultations: theoretical and practical issues. *Soc Sci Med* 2005;61:1516-28.
- 9 Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000;51:1087-110.
- 10 Ouwens M, Hermens R, Hulscher M, et al. Development of indicators for patient-centred cancer care. *Support Care Cancer* 2010;18:121-30.
- 11 Stewart MA, Brown JB, Weston W, et al. *Patient-centered medicine—transforming the clinical method*. Abingdon: Radcliffe Medical Press, 2003.
- 12 Scholl I, Zill JM, Härter M, et al. An integrative model of patient-centeredness—a systematic review and concept analysis. *PLoS One* 2014;9:e107828.
- 13 van Dulmen S. Patient-centredness. *Patient Educ Couns* 2003;51:195-6.
- 14 Zill JM, Scholl I, Härter M, et al. Which dimensions of patient-centeredness matter? - Results of a web-based expert Delphi survey. *PLoS One* 2015;10:e0141978.
- 15 Bensing J, Rimondini M, Visser A. What patients want. *Patient Educ Couns* 2013;90:287-90.
- 16 Hudon C, Fortin M, Haggerty J, et al. Patient-centered care in chronic disease management: a thematic analysis of the literature in family medicine. *Patient Educ Couns* 2012;88:170-6.
- 17 Luxford K. What does the patient know about quality? *Int J Qual Health Care* 2012;24:439-40.
- 18 Maassen EF, Schrevel SJC, Dedding CWM, et al. Comparing patients' perspectives of 'good care' in Dutch outpatient psychiatric services with academic perspectives of patient-centred care. *J Ment Health* 2017;26:84-94.
- 19 Yasein NA, Shakhatreh FM, Shroukh WA, et al. A Comparison between patients' and residents' perceptions of patient centeredness and communication skills among physicians working at Jordan University Hospital. *Open J Nurs* 2017;7:698-706.
- 20 Busse R, Riesberg A, World Health Organization. *Health care systems in transition: Germany*. Copenhagen: WHO Regional Office for Europe, 2004.



- 21 Mossialos E, Djordjevic A, Osborn R. *International profiles of health care systems*, 2017.
- 22 Christalle E, Zeh S, Hahlweg P, *et al.* Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open* 2018;8:e025896–13.
- 23 Boulkedid R, Abdoul H, Loustau M, *et al.* Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011;6:e20476.
- 24 Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32:1008–15.
- 25 van Rijssen LB, Gerritsen A, Henselmans I, *et al.* Core set of patient-reported outcomes in pancreatic cancer (COPRAC): an international Delphi study among patients and health care providers. *Ann Surg* 2019;270:158–64.
- 26 Bundesministerium für Arbeit und Soziales. *Leichte Sprache - Ein Ratgeber*, 2014.
- 27 Limesurvey GmbH. An open source survey tool /LimeSurvey GmbH, Hamburg, Germany.
- 28 Jünger S, Payne SA, Brine J, *et al.* Guidance on conducting and reporting Delphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review. *Palliat Med* 2017;31:684–706.
- 29 IBM SPSS statistics for windows; 2015.
- 30 Santana MJ, Manalili K, Jolley RJ, *et al.* How to practice person-centred care: a conceptual framework. *Health Expect* 2018;21:429–40.
- 31 Eurostat. *Hospital discharges and length of stay for inpatient and curative care*, 2018: 1–16. <http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>
- 32 Sacristán JA, Aguarón A, Avendaño-Solá C, *et al.* Patient involvement in clinical research: why, when, and how. *Patient Prefer Adherence* 2016;10:631–40.

9.3. Publication 3

Reference: Christalle, E., Zeh, S., Hahlweg, P., Kriston, L., Härter, M., Zill, J., & Scholl, I. (2022). Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—A best practice example for generating patient-reported measures from qualitative data. *Health Expectations*, 25(4), 1529-1538.

Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—A best practice example for generating patient-reported measures from qualitative data

Eva Christalle MSc, Research Assistant  | Stefan Zeh MSc, Research Assistant  |
Pola Hahlweg PhD, Senior Research Assistant, Head of Research Group  |
Levente Kriston PhD, Professor  | Martin Härter PhD, MD, Professor, Director  |
Jödis Zill PhD, Senior Research Assistant, Head of Research Group  |
Isabelle Scholl PhD, Professor, Head of the Outpatient Clinic 

Department of Medical Psychology,
University Medical Center Hamburg-
Eppendorf, Hamburg, Germany

Correspondence

Eva Christalle, MSc, Research Assistant,
Department of Medical Psychology,
University Medical Center Hamburg-
Eppendorf, Martinistraße 52, 20246
Hamburg, Germany.
Email: e.christalle@uke.de

Funding information

Bundesministerium für Bildung und
Forschung, Grant/Award Number: 01GY1614

Abstract

Introduction: To effectively foster patient-centeredness (PC), it is crucial to measure its implementation. So far, there is no German measure to assess PC comprehensively. The aim of this study is to develop and select items for the Experienced Patient-Centeredness (EPAT) Questionnaire, a patient-reported experience measure (PREM). The EPAT intends to assess PC from the perspective of adult patients treated for different chronic diseases in inpatient and outpatient settings in Germany. Furthermore, we aim at providing a best-practice example for developing PREMs from qualitative data.

Methods: The development process comprised a three-phase mixed-method design: (1) preparation, (2) item generation and (3) item selection and testing of content validity. We generated items using qualitative content analysis based on information from focus groups, key informant interviews and literature search. We selected items using relevance rating and cognitive interviews. Participants were patients from four chronic disease groups (cancer, cardiovascular disease, mental disorder, musculoskeletal disorder) and healthcare experts (e.g., clinicians, researchers, patient representatives).

Results: We conducted six focus groups with a total of 40 patients, key informant interviews with 10 experts and identified 48 PREMs from international literature. After team discussion, we reached a preliminary pool of 152 items. We conducted a relevance rating with 32 experts and 34 cognitive interviews with 21 patients. We

Jödis Zill and Isabelle Scholl shared senior authorship.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Health Expectations* published by John Wiley & Sons Ltd.

selected 125 items assessing 16 dimensions of PC and showed high relevance and comprehensibility.

Conclusions: The EPAT questionnaire is currently undergoing psychometric testing. The transparent step-by-step report provides a best practice example that other researchers may consider for developing PREMs. Integrating literature and experts with a strong focus on patient feedback ensured good content validity. The EPAT questionnaire will be helpful in assessing PC in routine clinical practice in inpatient and outpatient settings for research and quality improvement.

Patient or Public Contribution: Patients were not involved as active members of the research team. While developing the funding proposal, we informally reached out to several patient organizations who all gave us positive feedback on the study aims, thereby confirming their relevance. Those patient organizations endorsed the funding proposal with formal letters of support and supported recruitment by disseminating advertisements for study participation.

KEYWORDS

mixed methods, patient-centeredness, patient-reported experience measure, qualitative data, questionnaire development

1 | INTRODUCTION

During the last decades, the importance of patient-centeredness (PC) in healthcare and its positive effects on outcomes has been increasingly discussed in the research.^{1–3} PC has been incorporated in healthcare agendas and laws by international organizations^{4,5} and national initiatives.⁶ As there was inconsistent use of the term PC, Scholl et al.⁷ integrated different descriptions of PC into a comprehensive framework. The resulting integrative model of PC included 15 dimensions, which were grouped into principles, enablers and activities.⁷ Later, a 16th dimension, patient safety, was added based on a Delphi study with patients.⁸

To effectively foster PC, it is crucial to first assess its status quo. By measuring the degree of implementation of PC, it is possible to identify areas for improvement and evaluate interventions aimed to increase PC.⁹ As PC focuses on patients' experiences, it is highly relevant to assess PC from their perspective. Furthermore, improvement of patient experiences is part of the triple and quadruple aim for enhancing the healthcare system.^{10,11} Patient-reported experience measures (PREMs) can be used to assess patients' experiences by asking about specific processes and actions in healthcare delivery.¹² They are different from satisfaction measures, which encompass a more subjective judgement based on individual expectations.¹² PREMs can be used for quality improvement, benchmarking as well as research and evaluation of interventions.^{12,13}

There are several instruments that measure different dimensions of PC.^{14–19} To our knowledge there is no German PREM that assesses all 16 dimensions of a PC. Neither did Bull et al.¹³ identify any PREMs in the German language nor any PREM reflecting all dimensions from the integrative model of PC in their systematic review.

We have identified a range of guidelines for developing patient-reported measures.^{20–22} Those guidelines agree on the importance of including the target group (i.e., patients in general) in the item development process as well as using qualitative techniques by conducting focus groups (FGs) or interviews with the said target group. Yet, there is a lack of guidelines on the exact process of deriving items from qualitative data.

Therefore, the aim of this paper was twofold. First, we aimed to report the item development and selection process of the Experienced Patient-Centeredness (EPAT) questionnaire, a PREM to assess PC from the patients' perspective. Second, we aimed to provide a best-practice example for developing patient-reported measures from qualitative data.

2 | METHODS

2.1 | Study design

The study was conducted as part of the ASPIRED project (Assessment of Patient-Centeredness through PREMs), a mixed-methods study to develop and psychometrically test a measure to assess PC from the patients' perspective.²³ The development process of the measure consisted of three main phases: (1) preparation, (2) item generation and (3) item selection and testing of content validity, which were based on the guidelines and recommendations for developing patient-reported measures.^{20–22} The Consolidated criteria for reporting qualitative research (COREQ) can be found in Appendix S1.

An overview of the study design and methods is given in Figure 1.

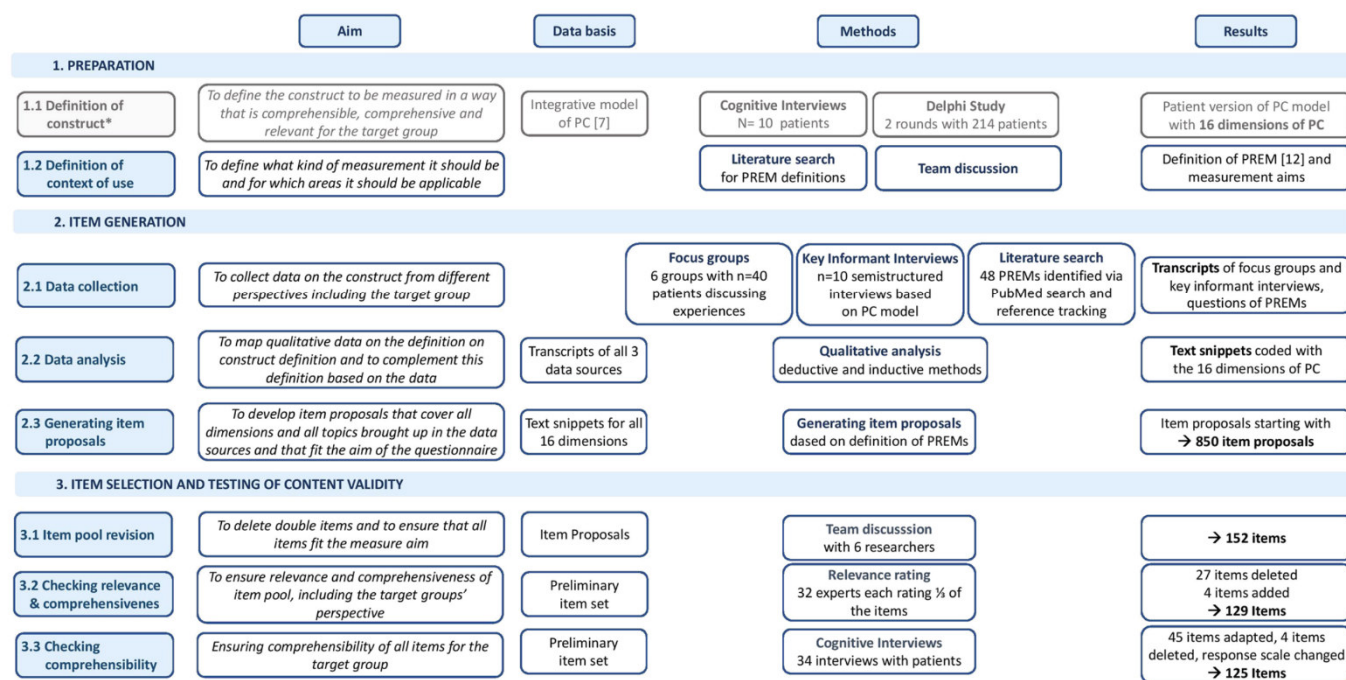


FIGURE 1 Overview of methods and results of the development of the EPAT questionnaire. *Step 1.1 is not reported in detail as it is already published by Zeh et al. [8]. EPAT, Experienced Patient-Centeredness; PC, patient-centeredness; PREM, patient-reported experience measure

2.2 | Sample

For cognitive interviews and FGs, patients were recruited through community-based strategies, that is, postings in supermarkets and cooperating self-help groups and patient organizations. As PC is particularly important for patients with chronic diseases, we focussed on four groups of chronic diseases: cancer, cardiovascular diseases, mental disorders or musculoskeletal disorders. We chose these four groups as they are common and show a high diversity in their course of the disease between and within the disease groups (e.g., constant severity in high blood pressure, episodic courses in depression or remission for long periods in cancer) as well as in their treatments (e.g., surgery for heart diseases, chemotherapy for cancer, psychotherapy for mental disorders). Due to this diversity, we expect that the questionnaire will be generalizable to other diseases. We used a convenience sample including all patients who wanted to participate, self-reported inclusion criteria were: age (18 years or older) and diagnosis (at least one disease of the four disease groups). All participants gave written consent.

Experts were selected as a convenience sample from the professional network of the research team, for example, collaboration partners, conference contacts, other research projects on PC or by recommendation from experts. The sample included clinicians, healthcare researchers, patient representatives, hospital quality managers and representatives of sickness funds. Experts were contacted via e-mail or phone and invited to participate in the key informant interviews and/or relevance rating.

2.3 | Step 1: Preparation

To develop items for a measure, the first step is to define the construct that should be measured and the context of use.^{20,24,25}

2.3.1 | Step 1.1: Definition of construct

The integrative model of PC was used as a construct for item development.^{7,8} It characterizes PC as follows: Patients have good access to continuous and interdisciplinary care. Healthcare professionals (HCPs) consider the patients as unique persons taking on a biopsychosocial perspective. They establish a good relationship and treat the patients with empathy and respect. The HCPs communicate well and integrate medical and nonmedical care. Patients get tailored information, are empowered to take care of their own health and are involved in treatment decisions. They receive safe care as well as physical and emotional support. For more details on the model and the methods, see Zeh et al.⁸

2.3.2 | Step 1.2: Definition of context of use

We searched for definitions of the term PREM and selected the following by Beattie et al.¹²: 'The emphasis is on asking patients whether or not, or how often, they have experienced certain care processes'.¹² In our understanding, the term 'care processes' encompasses all kinds of behaviours by the HCPs in line with the

definition by Bull et al.¹³ that PREMs 'are tools that capture "what" happened during an episode of care, and "how" it happened from the perspective of the patient'. Further, we decided that the measure should be (1) usable for all four chronic disease groups (i.e., cancer, cardiovascular disease, mental disorder, musculoskeletal disorder), (2) suitable for inpatient and outpatient healthcare delivery for adult patients and (3) applicable for research (e.g., evaluation of interventions) and routine clinical practice (e.g., for quality improvement). As the measure should be used to provide feedback to outpatient and inpatient centres, we focused on the interaction between patients and HCPs and excluded the system level (e.g., interaction with health insurances, conditions due to health policy regulations).

2.4 | Step 2: Item generation

2.4.1 | Step 2.1: Data collection

A combination of deductive (based on literature and existing measures) and inductive methods (generation of items from the responses of individuals) is recommended for developing items for patient-reported measures.²⁰ Therefore, we used three different sources: (1) a *literature search* to identify existing PREMs on PC, (2) FGs with patients and (3) *key informant interviews* with experts.

(1) The *literature search* was performed in PubMed with the following search strategy: PREM[Title/Abstract] OR (patient-reported experience measures[Title/Abstract]) OR (patient-reported experience measure[Title/Abstract]). No limitations were applied. All references resulting from this search were screened by one researcher (E. C.) and all measures identified were included. Furthermore, references to articles that were included after the first screening and references of a review on PREMs for hospitals¹² were screened for further measures.

(2) In the FGs, we shortly introduced the project, and that our aim was to develop a questionnaire assessing PC. We asked patients to reflect on their own healthcare experiences and to write down three notes to each question: 'What was good in your healthcare?' and 'What should have been done differently in your healthcare?'. Notes were collected and discussed in the group. To focus on specific processes, we asked participants to describe which specific actions the HCPs showed or should have shown in their opinion. FGs had a duration of 120 min (including a break of 15 min).

Focus groups (FGs) were conducted in German at the Department of Medical Psychology of the University Medical Center Hamburg-Eppendorf (UKE). The FGs were moderated by S. Z. with support from E. C. In addition, I. S. participated in the first group and J. Z. in the last as additional moderators. J. Z. and I. S. (both female) are psychologists, licensed psychotherapists and postdoctoral senior researchers with extensive experience in qualitative research methods. E. C. (female) and S. Z. (male) are junior researchers and PhD candidates. They had training in qualitative research methods and experience with conducting semi-structured interviews. All had several years of experience as researchers in the field of PC. FG

participants did not know the researchers before. All researchers introduced themselves, their professional background and personal motivation as well as the aim of the study.

(3) We conducted semi-structured *key informant interviews* with a duration of 30–45 min. All key informant interviews were conducted by S. Z. without the presence of third persons. They took place either at the department, at the workplace of the participant or via telephone. We instructed participants to read the model of PC⁸ step by step and describe for each dimension what specific actions by the HCPs are needed to fulfil the respective dimension. Further, participants were to suggest questions that could be used to assess each dimension in a questionnaire.

The guides for the FGs and key informant interviews were designed in German by S. Z. and E. C. with feedback from J. Z. or I. S. They were piloted with student research assistants and research interns. The English translations can be found in Appendix S2 (FGs) and Appendix S3 (key informant interviews).

2.4.2 | Step 2.2: Data analysis

FGs and key informant interviews were audio-recorded, transcribed, and anonymized. No field notes were taken. All analyses were carried out using the transcripts. Neither transcripts nor findings were returned to participants for feedback and no repeat interviews were carried out.

All sources (identified measures and said transcripts) were imported into MaxQDA,²⁶ a software specifically designed for qualitative analysis. For analysis of FGs, key informant interviews and identified measures, we used qualitative content analysis based on Mayring.^{27,28} First, we generated a coding system based on the model of PC⁸ (deductive approach). This included 16 codes corresponding to each of the 16 dimensions of the PC. Further, we created subcodes based on the description of each dimension in the model. We further created new codes whenever new topics emerged from the data (inductive approach).

We analysed all data sources by two of three coders (E. C., J. Z. or S. Z.). The first FG as well as the first key informant interview were analysed independently by two researchers (E. C. and S. Z., and J. Z. and S. Z., respectively). All further data sources were coded by one researcher and a second researcher checked the codings and noted cases of disagreement. Then, the two researchers discussed these codings to reach consent. We discussed the codings of every FG/interview/batch of measures before coding the next one to allow for the development of our coding system while ensuring consent on the usage and meaning of every code. No major or minor themes were identified as this is not part of qualitative content analysis based on Mayring and we did not aim to prioritize themes.^{27,28}

2.4.3 | Step 2.3: Generating items

For each dimension, the text fragments that had been marked by the respective codes and subcodes were collected in one document.

Then, one researcher (E. C., J. Z. or S. Z.) used MaxQDA²⁶ to develop items based on the text fragments and a second researcher complemented the items. To focus on specific actions of HCPs, we started items with 'The HCPs did/have/showed...' whenever possible. The result of this step was a pool of items for each dimension.

2.5 | Step 3: Item selection and testing of content validity

In the next steps, items were selected. During this process, we checked the item pool for content validity, that is whether the items are 'relevant, comprehensive and comprehensible with respect to the construct of interest and target population'.²⁹

2.5.1 | Step 3.1: Item pool revision

As text fragments were often coded with more than one dimension, they were also included more than once in the generation of items. This led to many double items. First, two researchers (E. C. and S. Z.) revised the item pool by sorting the items and merging similar items. Second, iterative group discussions with other researchers followed (P. H., L. K., I. S., J. Z.). These included (1) ensuring that all items were sorted into the best fitting dimension according to the model of PC, (2) deleting all items that did not assess aspects of PC or did not match the context of use, (3) further merging of similar items to eliminate duplicates and (4) refining the wording of the items.

2.5.2 | Step 3.2: Checking relevance and comprehensiveness

We conducted a relevance rating with experts (clinicians, healthcare researchers, patient representatives and hospital quality managers) using an online survey via LimeSurvey.³⁰ We included experts with a professional background as well as with a patient background as recommended by COSMIN.²⁵ Due to a large number of items, every expert rated one-third of the item pool. The items were equally distributed according to the profession of the experts. For each item, experts were asked to rate two questions: 'How relevant is this item for patients?' (response scale 0 = not relevant, 1 = a little bit relevant, 2 = quite relevant, 3 = very relevant) and 'How feasible is it for patients to assess this item?' (response scale 0 = not feasible, 1 = feasible, but not easy, 2 = rather easy, 3 = easy). To check for comprehensiveness, we added a free text box for each dimension asking whether any important items were missing.

We established item deletion criteria in an exploratory approach after data collection using mean ratings and the content validity index (CVI). The CVI of each item is the percentage of experts who rated an item as 'quite relevant' or 'very relevant'.³¹ In the literature, deletion of items with a CVI of less than 0.7 or 0.8 is recommended.³¹ We decided against such a high threshold as it would lead to the deletion

of items that were not upmost but still quite relevant. Further, there is a discrepancy in recommendations that at least twice as many items as in the final questionnaire should be included in psychometric testing,²⁰ and that each subscale needs to comprise at least four items.³² We decided to delete an item, if (1) the item was rated by at least half of the experts as not or hardly relevant (corresponding to a CVI of 0.5 or lower), (2) the item had mean feasibility to assess rating below 2 or (3) for subscales that still had more than four items, items with a mean relevance rating below 3 were deleted.

Finally, items were added after discussing the free text box comments in our team.

2.5.3 | Step 3.3: Checking comprehensibility

The resulting item pool was checked for comprehensibility in cognitive interviews with patients. To minimize the cognitive load of the questionnaire, we decided to use the same response scale for all items instead of changing it for different items. We translated a 4-point scale describing frequencies into German (0 = never, 1 = sometimes, 2 = usually, 3 = always). This response scale is used for example in the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).³³

Participants in cognitive interviews were instructed to think of their last inpatient admission or outpatient appointment. The interviewer (S. Z.) asked participants to read the questions step by step and think aloud about how they would reply. Further, he asked them to explain certain words or phrases using their own words and about their preference regarding different versions of certain items. The interview guide was piloted with student research assistants and research interns and can be found in Appendix S4. Cognitive interviews were conducted by S. Z. without the presence of a third person. Some of the participants already knew the interviewer from the FGs, others met S. Z. for the first time. The interviews took place at the Department of Medical Psychology of the University Medical Center Hamburg-Eppendorf (UKE) and were done in German.

Cognitive interviews were audio-recorded, transcribed and anonymized and lasted between 45 and 60 min. The interviewer took field notes about the proposed changes directly after the interviews. Then, another researcher (E. C.) listened to the audio recordings, rated the comprehensibility of each item on a 3-point scale (good, unclear, insufficient) and took notes on suggestions for item revisions made by the participants. The results and suggestions for item revisions were discussed by two researchers (S. Z., E. C.), consulting a third person if needed (I. S.). For each item, at least three cognitive interviews were conducted per round. If any of those interviews showed problems with the comprehensibility of an item or if any improvements were suggested, the corresponding item was adjusted and tested in another round. Only when all interviews in a round showed good comprehensibility of an item, we stopped testing it. Based on this procedure, we conducted two rounds of cognitive interviews as in the second round no items showed a further need for adjustments.

Participants who were willing to participate in a second interview round were allowed to do so as the recruitment process was hastened. In the second interview round, either completely new items were presented or items that were modified after the first round. The two interviews were at least 4 weeks apart.

3 | RESULTS

An overview of the results is given in Figure 1.

3.1 | Sample

We conducted six FGs with a total of 40 patients and 34 cognitive interviews with 21 patients (some participated in two interviews). We have no data on how many patients refused to participate. In cognitive interviews, there were no drop-outs. In FGs, five patients did not show up without giving reasons.

Sample characteristics of the patients are shown in Table 1.

For key informant interviews, 12 experts were invited of which 10 participated (two declined due to competing demands). There was no dropout. Mean age was 54.2 (standard deviation: 7.2, range: 44–65) and 80% were male. Five were physicians, two were working as head of a health facility/professional association and there was one support group coordinator, researcher and health insurance representative each.

Overall, 32 experts participated in the relevance rating. Each third of the items was rated by 8–13 experts from different professional backgrounds. The distribution of different professions over the items is reported in Table 2.

3.2 | Step 2: Item generation

We identified 48 internationally existing PREMs in the literature search (see Appendix S5 for references). Main settings were inpatient care ($n = 7$), primary care ($n = 5$) and other outpatient settings ($n = 4$). Most have been developed in Europe ($n = 29$) and North America ($n = 17$).

Based on the model, 16 codes corresponding to the 16 dimensions and 98 subcodes were derived (deductive approach). Based on the qualitative content analysis, we did not find any additional dimensions of PC but introduced 44 new subcodes (inductive approach). As we used convenience sampling, full theoretical saturation was not aimed for and was difficult to reach. Based on the last FG only two more codes were introduced. For a full overview of our coding tree and to know how many codes were derived from each source, please refer to Appendix S6. For text examples with the derived items and the full first item pool, please refer to Appendix S7. Overall, we developed 850 items, ranging from 15 items for the dimension 'support of mental well-being' to 141 items for 'PC characteristics of healthcare providers'.

TABLE 1 Characteristics of the participants in the focus groups and cognitive interviews

Characteristics	Focus groups	Cognitive interviews
Sample size	40	21
Age (in years)	$M = 54.2$ ($SD = 12.7$), range: 30–79	$M = 51.5$ ($SD = 15.0$), range: 22–80
Sex		
Female	31 (77.5%)	16 (76.2%)
Male	9 (22.5%)	5 (23.8%)
First language		
German	38 (95.0%)	20 (95.2%)
Other	2 (5.0%)	1 (4.8%)
Occupational status ^a		
Employed	9 (22.5%)	7 (33.3%)
Unemployed	6 (15.0%)	5 (23.8%)
Retired	19 (47.5%)	7 (33.3%)
Student/trainee	2 (5.0%)	3 (14.3%)
Marital status		
Unmarried	25 (62.5%)	10 (47.6%)
Married/partnered	6 (15.0%)	3 (14.3%)
Divorced	9 (22.5%)	8 (38.1%)
Formal education		
Low ^b	6 (15%)	1 (4.8%)
Intermediate ^c	10 (25.0%)	8 (38.1%)
High ^d	15 (37.5%)	5 (23.8%)
Very high ^e	9 (22.5%)	7 (33.3%)
Chronic disease group ^a		
Cancer	5 (12.5%)	2 (9.5%)
Cardiovascular disease	10 (25.0%)	7 (33.3%)
Mental disorder	27 (67.5%)	13 (61.9%)
Musculoskeletal disorder	22 (55.0%)	12 (57.1%)

Abbreviations: M, mean; SD, standard deviation.

^aMultiple answers possible.

^blow = no formal degree or graduation after less than 10 years at school.

^cIntermediate = graduation after 10 or 11 years at school.

^dHigh = graduation after more than 11 years at school.

^eVery high = college or university degree.

3.3 | Step 3: Item selection and testing of content validity

Deletion of repeating items as well as items that did not fit the context of use led to an item pool of 152 items ranging from three

TABLE 2 Distribution of professions in the relevance rating (each one rating one-third of the items)

	1st Third of items	2nd Third of items	3rd Third of items	Total
HCPs	3	3	4	11
Researchers	5	5	1	10
Patient representatives	3	4	3	10
Quality manager	0	1	0	1
Total	11	13	8	32

Abbreviation: HCPs, healthcare professionals.

items for the dimension 'support of mental well-being' to 15 items for the dimension 'good planning of care'.

We deleted items as follows: (1) 7 items had a CVI of 0.5 or lower, (2) 5 items had a mean feasibility to assess rating below 2 and (3) 15 items had a mean relevance rating below 3.

Based on the free-text comments in the relevance rating, we added four further items to the questionnaire.

This resulted in 129 items that were tested in cognitive interviews. Overall, item comprehensibility was high. Minor changes in 45 items were made after analysis of the first round. Those were changes of single words, addition of examples, or highlighting of single words. After the second round, no more adaptations were necessary.

Many participants considered the four response options derived from the HCAHPS too limiting and suggested adding more response categories (e.g., 'often' or 'seldom'). We tested the suggested response categories in the second round of cognitive interviews and found that participants had differing opinions on which response options we should add and perceived the order of the new response categories differently. Therefore, we decided to use a 6-point Likert scale rating agreement with statements instead (1 = completely disagree, 2 = strongly disagree, 3 = somewhat disagree, 4 = somewhat agree, 5 = strongly agree, 6 = completely agree). This response scale is used for example in the German SDM-Q-9 and SDM-Q-Doc,^{34,35} a widely used³⁶ questionnaire that was developed by team members and which was also transformed from a 4-point-scale to a 6-point-scale to reduce ceiling effects.³⁴ While we did not test them with our questionnaire, cognitive interviews conducted at that time by a colleague (H. C., cp. acknowledgements) showed that patients were well able to use its response scale.

Based on the feedback of the participants, we deleted four items that were clearly dichotomous (e.g., 'I received a discharge letter') and did therefore not fit the response scale. As a result, the final item pool comprises 125 items. The inpatient version of the questionnaire has 121 items and the outpatient version has 120 items (with 116 items being identical for both versions).

4 | DISCUSSION

4.1 | Summary of results

We reported the process of the item development and selection of the EPAT questionnaire. It will be the first measure to assess all dimensions of the integrative model of PC^{7,8} from the patients' perspective. Its current form has two versions, one for outpatient and one for inpatient healthcare while the items are the same for all four chronic disease groups considered here (cancer, cardiovascular disease, mental disorder, musculoskeletal disorder). All items in the final item pool showed good relevance and comprehensibility.

Our study offers a best-practice example that other researchers may use to develop patient-reported measures. In particular, we add to existing guidelines by describing a transparent process about how items were developed from qualitative data using a two-step process. First, we used qualitative analysis to link the data to the construct definition and complemented the definition with the data. Second, we used the data to generate items for each dimension of the construct based on the context of use.

4.2 | Strengths and limitations

A strength of this approach and the EPAT questionnaire is the involvement of patients in several steps: (1) in defining the construct to be measured, (2) during item generation using FGs and (3) during item selection in the relevance rating (as patient representatives) and in cognitive interviews. With these steps involving patients, we ensured that our item pool meets all three criteria given by the definition of content validity in the 'Consensus-based Standards for the selection of health Measurement INstruments' (COSMIN) criteria.²⁹ This is highly relevant as content validity can be considered one of the most central psychometric properties of measures.²⁴ To ensure high quality of patient-reported measures, the input from patients is crucial as they are the primary experts regarding their own health and care.²⁵ So far, many PREMs lack transparent reports of patient involvement. Two-thirds of all PREMs identified by Bull and colleagues¹³ did not provide enough information to rate content validity even though they already used a simplified version of the COSMIN criteria. Other reviews on measures assessing subscales of PC resulted in mostly poor ratings of content validity, again, often due to missing information.^{14,15,19}

One reason for such poor ratings might be that no guidelines exist for developing a PREM. This is also the reason why in this study we refer to references for patient-reported outcome measures (PROMs). As both PREMs and PROMs are patient-reported there are many similarities. Still, there is a big content-related difference between assessing experience during a treatment (PREMs) or outcomes after a treatment (PROMs). Guidelines that take into account those content-specific aspects for PREM development are much needed to foster the development of PREMs with high content validity.

A limitation of our study is that both patient samples are nonrepresentative regarding formal education. Further, the sample sizes in relevance rating and cognitive interviews are small. The COSMIN Study Design checklist recommends seven participants for qualitative studies while in our cognitive interviews, we had a minimum of three interviews per item. Having more participants was not feasible in this study due to the length of the item pool, resulting in a high number of cognitive interviews. Yet, we only stopped testing items if all three interviews showed good comprehensibility. All items that we adapted after the first round were tested in three more interviews and we stopped after that second round as all three interviews showed good comprehensibility.

Further limitations concern the literature search. As we lacked access to EMBASE, we searched only PubMed. COSMIN guidelines consider PubMed the minimum but still highly recommend including EMBASE.³⁷ To compensate for that, we included reference tracking of a systematic review on PREMs, which used several databases.¹² Another limitation is that references were screened by only one researcher. To reduce bias, this researcher included all measures she found during the screening. Whether these measures were relevant for our aim or not was decided during qualitative analysis where only relevant parts were coded and which was done by two researchers. For quantitative studies, the COSMIN checklist recommends a minimum of 50 participants while our items were rated by 8–13 experts. It has to be noted that this is above the minimum of 5–7 experts recommended by Boateng et al.²⁰ Yet, a larger and more representative sample including not only patient representatives but also patients without a professional background in healthcare might have been beneficial and would have allowed for more rigorous decisions based on the relevance rating. In addition, while patients were included as a source for data collection, they were not actively involved in the research process. For future questionnaire developments including future research on the EPAT questionnaire, we suggest including patient researchers in the entire research process to enhance relevance.³⁸ In particular, patients can be involved in setting priorities and defining measurement aims, designing guides for FGs and interviews, in the analysis of the qualitative data and in the item reduction process.

4.3 | Implications for future research

The resulting item pool of the EPAT questionnaire is planned to be psychometrically tested with $n = 2000$ patients from the four chronic disease groups in inpatient and outpatient settings.²³ We will examine item characteristics like floor and ceiling effects, the number of missing values and corrected item-total correlations as well as the factor structure to further reduce the item pool. We aim to select one item for each of the 16 dimensions to obtain a short version of the questionnaire that can be used easily in routine care. Further items will be available to assess each

dimension in more detail. This will allow users (e.g., researchers, HCPs, healthcare managers, etc.) to select parts of the questionnaire depending on their aim of measurement, their focus of interest and their available resources.

Eventually, the EPAT can be used to examine the effectiveness of interventions aiming to promote PC or certain dimensions of PC in Germany.

Furthermore, it can be used to explore relationships between the level of PC (e.g., of a healthcare centre or a single ward or HCP) and other outcomes like patient-reported outcomes (e.g., health-related quality of life).

The items of the EPAT were derived for four different chronic disease groups regardless of the different treatments, symptoms and so forth, and thus are roughly disease generic. The items should be tested with patients with further chronic diseases as well as acute diseases to investigate whether the EPAT can be used for them as well. So far, the EPAT will be available only in German. After the questionnaire is finalized we recommend it to be translated, culturally adapted and tested with patient groups in other countries. To our knowledge, there is no measure in any language that assesses PC directly from the patients' view in such detail while at the same time allowing users to adapt it flexibly to their own needs. As the model of PC that was used for development is based on international literature, we expect many items to be independent of the German context. Translating the EPAT questionnaire into other languages would allow for the comparison of PC in different countries and healthcare systems.

Moreover, the EPAT could be translated for migrant groups living and experiencing the healthcare system in Germany or adapted for other patient groups (e.g., patients with a disability). Hereby, special PCs needs or barriers for more vulnerable groups could be detected and addressed.

Finally, the EPAT could also be adapted to measure the perspective of HCPs on PC processes within their organization, at their workplace or in their patient contacts or on their perceived own competencies in PC. Therefore, the patients and the HCP perspective could be compared and PC needs on both sides could be detected. Possible results could be used to develop PC training and interventions to foster PC (e.g., to strengthen the perceived self-efficacy or competencies of HCP or to change work processes within organizations).

5 | CONCLUSIONS

We developed a preliminary version of the EPAT questionnaire, which will undergo psychometric testing. It can be used to assess PC from the patients' perspective in inpatient and outpatient healthcare. The thorough development process that included patients in several steps ensures good content validity of the EPAT questionnaire and can be used as a best-practice example to develop patient-reported measures from qualitative data.

AUTHOR CONTRIBUTIONS

Isabelle Scholl and Jödis Zill were the responsible principal investigators of the study. Isabelle Scholl, Levente Kriston and Martin Härter were involved in the planning and preparation of the study. Eva Christalle and Stefan Zeh recruited participants and collected data. Eva Christalle, Stefan Zeh and Jödis Zill analysed data. Eva Christalle, Stefan Zeh, Jödis Zill, Isabelle Scholl, Pola Hahlweg and Levente Kriston interpreted the results. Eva Christalle wrote the first draft of the manuscript. All authors critically revised the manuscript for important intellectual content. All authors gave final approval of the version to be published and agreed to be accountable for the work.

ACKNOWLEDGEMENTS

The authors would like to thank our student assistants Tanja Kloster and Sibel Gürü for transcribing the interviews and focus groups, and Hannah Cords for sharing her results of cognitive interviews on the Shared Decision Making Questionnaire 9. This study is funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung—BMBF) with grant number 01GY1614. The funder had no role in the decision to publish or preparation of the manuscript. Open Access funding enabled and organized by Projekt DEAL.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Deidentified data that support the findings of this study are available on reasonable request. Investigators who propose to use the data have to provide a methodologically sound proposal directed to the corresponding author. Signing a data use/sharing agreement will be necessary, and data security regulations both in Germany and in the country of the investigator who proposes to use the data must be complied with. Preparing the data set for use by other investigators requires substantial work and is thus linked to available or provided resources. The data sets used in this study are in the German language.

ETHICS STATEMENT

The study was carried out according to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good scientific practice were respected. The study had been approved by the Ethics Committee of the Medical Association Hamburg (study ID: PV5724). Study participation was voluntary and no foreseeable risks for participants resulted from the participation in this study. Participants were fully informed about the aims of the study, data collection and the use of collected data. Written informed consent was obtained before participation. Preserving principles of data sensitivity, data protection and confidentiality requirements were met.

ORCID

Eva Christalle  <https://orcid.org/0000-0001-7813-5800>
 Stefan Zeh  <https://orcid.org/0000-0002-1748-3155>
 Pola Hahlweg  <http://orcid.org/0000-0002-8768-8594>
 Levente Kriston  <http://orcid.org/0000-0003-0748-264X>
 Martin Härter  <https://orcid.org/0000-0001-7443-9890>
 Jödis Zill  <https://orcid.org/0000-0001-8995-5556>
 Isabelle Scholl  <https://orcid.org/0000-0002-7639-0880>

REFERENCES

1. Dwamena F, Holmes-Rovner M, Gauden CM, et al. Interventions for providers to promote a patient-centred approach in clinical consultations. *Cochrane Database Syst Rev*. 2012;(12).
2. McMillan SS, Kendall E, Sav A, et al. Patient-centered approaches to health care: a systematic review of randomized controlled trials. *Med Care Res Rev*. 2013;70(6):567-596.
3. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev*. 2013;70(4):351-379.
4. Berwick DM. A user's manual for the IOM's 'Quality Chasm' report. *Health Aff*. 2002;21(3):80-90.
5. World Health Organization (WHO). Innovative Care For Chronic Conditions: Building Blocks For Action. 2002. Accessed May 17, 2021. https://apps.who.int/iris/bitstream/handle/10665/42500/WHO_NMC_CCH_02.01.pdf
6. Deutscher Bundestag. Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten. *Bundesgesetzblatt*. 2013;(Teil I): 277-282.
7. Scholl I, Zill JM, Härter M, Dirmaier J. An integrative model of patient-centeredness—a systematic review and concept analysis. *PLoS One*. 2014;9(9):e107828.
8. Zeh S, Christalle E, Hahlweg P, Härter M, Scholl I. Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ Open*. 2019;9(12):e031741.
9. De Silva D. Helping measure person-centred care: a review of evidence about commonly used approaches and tools used to help measure person-centred care. 2014. Accessed May 17, 2021. <https://www.health.org.uk/publications/helping-measure-person-centred-care>
10. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Aff*. 2008;27(3):759-769.
11. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med*. 2014;12(6): 573-576.
12. Beattie M, Murphy DJ, Atherton I, Lauder W. Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Syst Rev*. 2015;4(1):1-21.
13. Bull C, Byrnes J, Hettiarachchi R, Downes M. A systematic review of the validity and reliability of patient-reported experience measures. *Health Serv Res*. 2019;54(5):1023-1035.
14. Barr PJ, Scholl I, Bravo P, Faber MJ, Elwyn G, McAllister M. Assessment of patient empowerment—a systematic review of measures. *PLoS One*. 2015;10(5):e0126553.
15. Christalle E, Zill JM, Frerichs W, et al. Assessment of patient information needs: a systematic review of measures. *PLoS One*. 2019;14(1):e0209165.
16. Gärtner FR, Bomhof-Roordink H, Smith IP, Scholl I, Stiggelbout AM, Pieterse AH. The quality of instruments to assess the process of shared decision making: a systematic review. *PLoS One*. 2018;13(2): e0191747.

17. Müller E, Zill JM, Dirmaier J, Härter M, Scholl I. Assessment of trust in physician: a systematic review of measures. *PLoS One*. 2014;9(9):e106844.
18. Sepucha KR, Scholl I. Measuring shared decision making: a review of constructs, measures, and opportunities for cardiovascular care. *Circ Cardiovasc Qual Outcomes*. 2014;7(4):620-626.
19. Zill JM, Christalle E, Müller E, Härter M, Dirmaier J, Scholl I. Measurement of physician-patient communication—a systematic review. *PLoS One*. 2014;9(12):e112637.
20. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health*. 2018;6:149.
21. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019. Accessed May 17, 2021. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf
22. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: a Practical Guide to their Development and Use*. Oxford University Press; 2015.
23. Christalle E, Zeh S, Hahlweg P, Kriston L, Härter M, Scholl I. Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open*. 2018;8(10):e025896.
24. Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for assessing the content validity of PROMs User manual version 1.0. 2021. Accessed May 19, 2021. <https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf>
25. Terwee CB, Prinsen C, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27(5):1159-1170.
26. VERBI Software. MAXQDA. VERBI Software, 2017. 2018.
27. Mayring P. Qualitative content analysis. *Forum: Qualitative Social Research*. 2000;1(2):Art. 20.
28. Mayring P. *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. 2014.
29. Prinsen C, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157.
30. LimeSurvey GmbH, Hamburg, Germany. Accessed April 11, 2022. <http://www.limesurvey.org>
31. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489-497.
32. Diamantopoulos A, Sarstedt M, Fuchs C, Wilczynski P, Kaiser S. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *J Acad Market Sci*. 2012;40(3):434-449.
33. Giordano LA, Elliott MN, Goldstein E, Lehrman WG, Spencer PA. Development, implementation, and public reporting of the HCAHPS survey. *Med Care Res Rev*. 2010;67(1):27-37.
34. Kriston L, Scholl I, Hölzel L, Simon D, Loh A, Härter M. The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. *Patient Educ Couns*. 2010;80(1):94-99.
35. Scholl I, Kriston L, Dirmaier J, Buchholz A, Härter M. Development and psychometric properties of the Shared Decision Making Questionnaire—physician version (SDM-Q-Doc). *Patient Educ Couns*. 2012;88(2):284-290.
36. Doherr H, Christalle E, Kriston L, Härter M, Scholl I. Use of the 9-item Shared Decision Making Questionnaire (SDM-Q-9 and SDM-Q-Doc) in intervention studies—a systematic review. *PLoS One*. 2017;12(3):e0173904.
37. Prinsen CA, Vohra S, Rose MR, et al. *Guideline for Selecting Outcome Measurement Instruments for Outcomes Included in a Core Outcome Set*. COMET COSMIN; 2016.
38. Greenhalgh T, Hinton L, Finlay T, et al. Frameworks for supporting patient and public involvement in research: systematic review and co-design pilot. *Health Expect*. 2019;22(4):785-801.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Christalle E, Zeh S, Hahlweg P, et al. Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—a best practice example for generating patient-reported measures from qualitative data. *Health Expect*. 2022;25: 1529-1538. doi:10.1111/hex.13494

9.4. Publication 4

Reference: Christalle, E., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J. M., Härter M., Bokemeyer C., Gallinat J., Gebhardt C., Magnussen C., Müller V., Schmalstieg-Bahr K., Strahl A., Kriston L., Scholl, I. (2024). Through the patients' eyes: psychometric evaluation of the 64-item version of the Experienced Patient-Centeredness Questionnaire (EPAT-64). *BMJ Quality & Safety*, 34(10), 643-656.



OPEN ACCESS

Through the patients' eyes: psychometric evaluation of the 64-item version of the Experienced Patient-Centeredness Questionnaire (EPAT-64)

Eva Christalle ,¹ Stefan Zeh,¹ Hannah Führes,¹ Alica Schellhorn,¹ Pola Hahlweg,¹ Jödis Maria Zill,¹ Martin Härter,¹ Carsten Bokemeyer,² Jürgen Gallinat,³ Christoffer Gebhardt,⁴ Christina Magnussen,^{5,6,7} Volkmar Müller,⁸ Katharina Schmalstieg-Bahr,⁹ André Strahl,¹⁰ Levente Kriston ,¹ Isabelle Scholl¹

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjqs-2024-017434>).

For numbered affiliations see end of article.

Correspondence to
Dr Isabelle Scholl;
i.scholl@uke.de

LK and IS are joint senior authors.

Received 19 April 2024
Accepted 31 July 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Christalle E, Zeh S, Führes H, et al. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2024-017434

ABSTRACT

Background Patient-reported experience measures (PREMs) are valuable tools to evaluate patient-centredness (PC) from the patients' perspective. Despite their utility, a comprehensive PREM addressing PC has been lacking. To bridge this gap, we developed the preliminary version of the Experienced Patient-Centeredness Questionnaire (EPAT), a disease-generic tool based on the integrative model of PC comprising 16 dimensions. It demonstrated content validity. This study aimed to test its psychometric properties and to develop a final 64-item version (EPAT-64).

Methods In this cross-sectional study, we included adult patients treated for cardiovascular diseases, cancer, musculoskeletal diseases and mental disorders in inpatient or outpatient settings in Germany. For each dimension of PC, we selected four items based on item characteristics such as item difficulty and corrected item-total correlation. We tested structural validity using confirmatory factor analysis, examined reliability by McDonald's Omega and tested construct validity by examining correlations with general health status and satisfaction with care.

Results Analysis of data from 2.024 patients showed excellent acceptance and acceptable item-total correlations for all EPAT-64 items, with few items demonstrating ceiling effects. The confirmatory factor analysis indicated the best fit for a bifactor model, where each item loaded on both a general factor and a dimension-specific factor. Omega showed high reliability for the general factor, while varying for specific dimensions. Construct validity was confirmed by absence of strong correlations with general health status and a strong correlation of the general factor with satisfaction with care.

Conclusions EPAT-64 demonstrated commendable psychometric properties. This tool allows comprehensive assessment of PC, offering flexibility to users who can measure each dimension with a four-item module or choose modules based on their needs. EPAT-64 serves multiple purposes, including quality improvement and

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Patient-reported experience measures (PREMs) can be used to assess patient-centredness (PC) from patients' perspective. The 'Experienced Patient-Centeredness Questionnaire' (EPAT) is the first PREM to comprehensively assess 16 dimensions of PC.

WHAT THIS STUDY ADDS

⇒ In this study, we tested the psychometric properties of all items developed for the EPAT and developed the 64-item version of the EPAT (EPAT-64), which demonstrated good psychometric properties.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The EPAT-64 can be used in research and routine care, for example, to evaluate interventions, provide feedback to healthcare professionals, support quality improvement, set benchmarks and consequently improve PC.

evaluation of interventions aiming to enhance PC. Its versatility empowers users in diverse healthcare settings.

BACKGROUND

Patient-centredness (PC) is a key element of high-quality healthcare^{1 2} and positively associated with several health-related outcomes.^{3–6}

To overcome ambiguities in the definition of PC, Scholl *et al* developed the integrative model of PC. The model includes 16 dimensions: (1) essential characteristics of the clinician, (2) clinician–patient relationship, (3) patient as a unique person, (4) biopsychosocial perspective, (5) clinician–patient communication, (6) integration of medical and non-medical care, (7) teamwork and team building, (8) access to care, (9) coordination and continuity of care, (10) patient safety, (11) patient information, (12) patient involvement in care, (13) involvement of family and friends, (14) patient empowerment, (15) physical support and (16) emotional support. This model is based on a thorough systematic review, integrating definitions from over 400 full texts and thus providing a comprehensive model.⁷ Furthermore, it has been validated in several empirical studies (eg, ref^{8,9}). Using this model grounded our study in a broad and nuanced understanding of the various existing definitions of PC.

Effective management of PC relies on sound measurement.^{10–11} Patient-reported experience measures (PREMs) capture patients' reports on whether or how often they have experienced specific processes or behaviours in their healthcare,¹² and positive patient experiences are associated with several positive outcomes.^{13–14} While many PREMs on PC have been developed,^{12–15} none assessed all 16 dimensions of PC as defined by Scholl and colleagues.^{7–16}

To bridge this gap, we developed the PREM 'Experienced Patient-Centeredness Questionnaire' (EPAT).¹⁷ We used a multistep process combining theory-driven approaches and the target group's perspective to ensure high content validity of EPAT. It is a disease-generic questionnaire in German language with an outpatient and an inpatient version. A reflective measurement model was used (ie, we assumed that specific aspects of patients' experiences are causally determined by underlying latent dimensions of PC). The preliminary EPAT version comprised 120 items for outpatients and 121 items for inpatients. Details of item development are reported elsewhere.¹⁷ However, a questionnaire of this length lacks feasibility and it was not yet psychometrically tested.

Thus, this study aimed to select items for the final 64-item version (EPAT-64) and test its psychometric properties.

METHODS

Study design

Development and psychometric testing of the EPAT was part of the ASPIRED study (Assessment of Patient-Centeredness Through Patient-Reported Experience Measures), a 6-year mixed-methods study aiming to assess patients' perspective on PC.^{17–18} Here, we report on the psychometric testing of EPAT with cross-sectional surveys. First, we assessed item characteristics of all items of the preliminary EPAT. Those results informed the item selection. For each of the 16

dimensions, a reduction to four items was decided, as the minimum number of items to test measurement models is 4.¹⁹ Hence, the final EPAT version has 64 items (EPAT-64). Consequently, we analysed structural validity, construct validity and reliability for the EPAT-64 items only in the same data set. As there are no guidelines for reporting psychometric evaluations for PREMs, we used the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guideline on reporting patient-reported outcome measures where applicable (see online supplemental appendix 1).²⁰

The EPAT questionnaire

The preliminary EPAT version and the final EPAT-64 comprise 16 domains^{7,8} and have an inpatient and an outpatient version each. All items are rated on a 6-point Likert scale of agreement (1=completely disagree, 2=strongly disagree, 3=somewhat disagree, 4=somewhat agree, 5=strongly agree, 6=completely agree). Every item offers the option 'does not concern me'. For most items, a high score implies high-experienced PC, while nine items are reversed in the preliminary version and only one item in the EPAT-64. EPAT was developed and tested in German. The final EPAT-64 was translated into English using a team-based translation process consisting of translation, review, adjudication, pretesting and documentation (TRAPD).²¹

Data collection and participants

We included participants from June 2020 to February 2022. As planned in the protocol,¹⁸ we used consecutive sampling in 10 inpatient and 14 outpatient healthcare institutions (eg, hospitals, primary care centres) in the metropolitan area of Hamburg, Germany. Outpatients received the questionnaire before their appointment and were asked to report on their experience during the last 4 weeks within the respective outpatient clinic. Inpatients received the questionnaire on day of discharge or the day before and were instructed to reply after discharge reporting on their whole stay. As participant inclusion was slowed down by the ongoing SARS-CoV-2 pandemic, we supplemented paper-and-pencil questionnaires with an online survey using LimeSurvey (LimeSurvey, Hamburg, Germany) in December 2020. We additionally invited participants via community-based strategies (ie, social media, postings in supermarkets, self-help groups and patient organisations from all parts of Germany). They were asked to refer to their last outpatient visit or inpatient stay within the past month.

Inclusion criteria were age of 18 years or older and being currently treated for at least one of the following four medical conditions: cardiovascular diseases, cancer, musculoskeletal diseases and mental disorders. Inclusion was based on self-report. As we aimed to create a generic measure, we covered the most prevalent treatment settings and medical conditions to

ensure representation of diverse healthcare experiences. We aimed to include 250 inpatients and 250 outpatients for each of the four medical conditions, resulting in a total target sample size of 2000.¹⁸ This allows for the use of maximum likelihood or weighted least squared estimators in the investigation of moderately complex models and for the detection of weak but meaningful correlations ($r=0.20$) with an alpha error probability of 0.05 and a power of 0.90.

In addition to EPAT, participants were asked to complete the first item of the German version of the 12-item Short Form Survey, assessing general health status.²² This item has an excellent item difficulty of 0.4 and high acceptability (less than 4% missing values). We assessed satisfaction with care with the German version of the 8-item Client Satisfaction Questionnaire,²³ a unidimensional questionnaire with high internal consistency (Cronbach's $\alpha=0.9$). We also administered the European Health Literacy Questionnaire (HLS-EU-Q16),²⁴ a short version of the HLS-EU-Q47 with low ceiling effects, strong correlation with the long version ($r=0.82$) and high internal consistency (Cronbach's $\alpha=0.90$).²⁵ The HLS-EU-Q16 can assess overall health literacy with a sum score of all items. Furthermore, we asked about demographic factors (eg, gender and age) and health status (eg, years since receiving the corresponding diagnosis). Finally, first in the online version and from June 2021 in the paper questionnaires, we included two control items ('To show that you are reading attentively, please check 'somewhat disagree'/'somewhat agree'). Data collection was anonymous. All participants gave written consent. Participants sent back paper questionnaires free of charge and had the possibility to receive an incentive of €10 on completion of the questionnaire.

Data analyses

First, we analysed all items of the preliminary EPAT version. We then selected items for EPAT-64 in a team discussion with five experts in the field of PC (IS, LK, MH, JMZ, EC). Three of them (IS, JMZ, EC) developed EPAT and were familiar with the qualitative data used for development.¹⁷ For each dimension, we chose four items resulting in 64 items covering 16 dimensions of PC. We considered the following psychometric properties during item selection: (1) percentage of missing values as proxy for acceptance,²⁶ (2) percentage of patients replying 'does not concern me' as proxy for relevance, (3) item difficulty calculated by standardising the mean to range from 0 to 1 (recommended range 0.2–0.8),²⁷ (4) within-dimension interitem correlations (recommended range >0.3),²⁸ (5) within-dimension corrected item–total correlation (recommended range >0.3)²⁸ and (6) content validity (eg, how important an item is to the definition of the given dimension).

For EPAT-64, we examined structural validity. We had a clear hypothesis concerning the number and

content of the dimensions to be measured, based on substantial previous work.^{7–9 17} Therefore, we used confirmatory factor analysis (CFA)^{29–31} using a robust version of maximum likelihood estimator to assess five different models: (1) unidimensional model, (2) correlated first-order dimension model, (3) hierarchical model, (4) bifactor model with uncorrelated dimensions and (5) bifactor model with correlated dimensions. See online supplemental appendix 2 for detailed descriptions and depictions of all models. For examination of model fit and selection, we followed recommendations by using a combination of model fit indices³²: root mean square error of approximation, standardised root mean square residual, Tucker-Lewis Index and Comparative Fit Index. For comparison of models, we used Akaike information criterion and Bayesian information criterion. Further, we report the χ^2 test statistic of model fit.³³ For the selected model, we report standardised factor loadings and average variance explained (AVE). We planned to test this model in the four medical conditions to examine measurement invariance between groups.²⁸ Further, we examined reliability by McDonald's Omega hierarchical which allows to estimate reliability in models with a bifactor structure.³⁴

Construct validity was assessed by correlating the model-based factor scores of each dimension with two other measures as described in the study protocol.¹⁸ Discriminant validity was tested by examining correlations with general health status.²² We hypothesised that the magnitude of these correlations should be below 0.3.³⁵ For convergent validity, we used correlations with satisfaction with care,²³ expecting correlations to exceed 0.5.³⁵

We used SPSS V.27.0 (IBM) to enter paper-and-pencil questionnaire data, to check their validity by a second person, to import data from LimeSurvey and to calculate sample characteristics. All other data cleaning and analyses were done with R V.4.3.2 (R Core Team, Vienna, Austria), specifically we used the lavaan package³⁶ for CFA, BifactorIndicesCalculator³⁷ for Omega hierarchical and psych package³⁸ for all other psychometric analyses. We excluded participants if they responded to less than 70% of EPAT items or if they failed both control items. Where applicable, we used full information maximum likelihood to deal with missing values,³⁹ treating 'does not concern me' replies as missing values. We report results for all medical conditions combined, but analysed data separately for inpatient and outpatient settings.

RESULTS

Sample

For the paper-and-pencil version, 4788 outpatients were invited to participate, of whom 2357 (49.2%) gave consent and 905 (18.9%) returned the questionnaire. Further, we invited 3046 inpatients, of whom 1799 (59.1%) gave consent and 704 (23.1%) returned

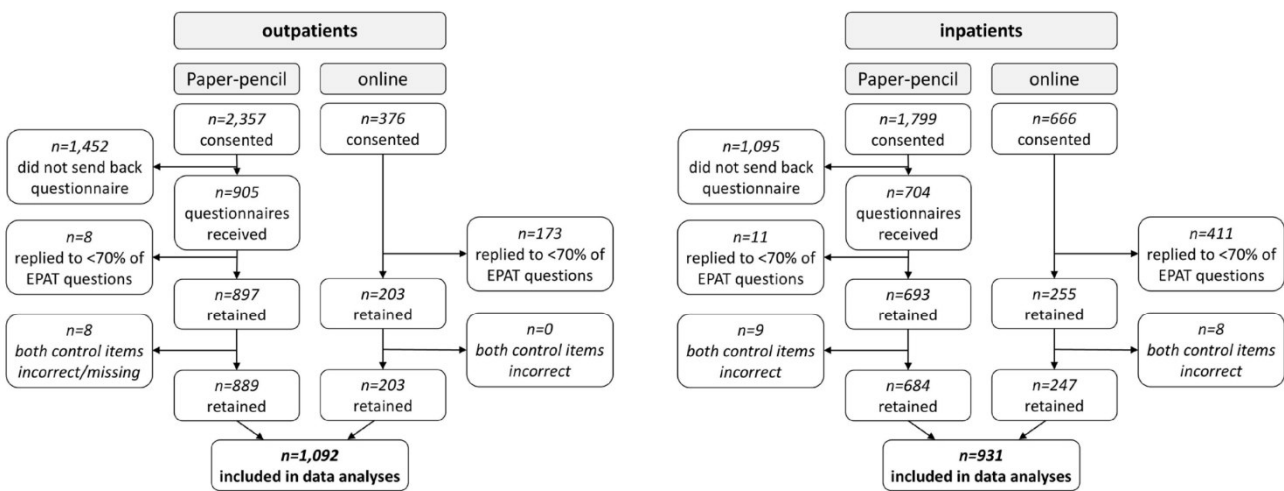


Figure 1 Participant flow chart. EPAT, Experienced Patient-Centeredness Questionnaire.

the questionnaire. For the online questionnaire, 1042 patients consented and 458 patients replied to at least 70% of the items (44.0%). We have no information on patients who denied consent or dropped out to perform dropout analyses. In total, questionnaires of 1092 outpatients and 931 inpatients were included in our analyses (see the flow chart in figure 1).

Of the included outpatients, 59.1% were female, 38.5% were male and 0.5% identified as diverse. Mean age was 52.8 years (SD=17.4, range 18–95). Mean time since initial diagnosis of the respective disease was 11.4 years (SD=11.0). Of the inpatients, 41.3% were female, 55.5% were male and 0.6% identified as diverse. Mean age was 55.5 years (SD=17.4, range 18–92). Mean time since initial diagnosis was 8.5 years (SD=10.0). See table 1 for further details on demographic and health-related variables. For subgroup sample characteristics regarding medical conditions refer to online supplemental appendix 3. For a comparison of paper-and-pencil and online questionnaires please see online supplemental appendix 4.

Item characteristics

Results on all items of the preliminary EPAT, which were the basis for item selection of EPAT-64, are reported in online supplemental appendices 5 and 6 for inpatients and outpatients, respectively.

For outpatients, all 64 selected items showed high acceptability (highest percentage of missing responses was 2.7%). Regarding relevance, 47 items had less than 30% of participants answering ‘does not concern me’ and seven items had over 50%. While 46 items showed excellent item difficulty (between 0.2 and 0.8), 18 items showed ceiling effects (item difficulty >0.8). All but one item had an item–total correlation greater than 0.3.

For inpatients, again, all 64 items had high acceptability (highest percentage 2.7%). We included 55 items with a ‘does not concern me’ rate below 30% and two items above 50%. Item difficulty was excellent

for 42 items, 22 items showed ceiling effects. All items had item–total correlation greater than 0.3.

Item characteristics of all items included in the EPAT-64 are shown in table 2. For item wording, please refer to the full questionnaires at www.ukc.de/epat.

Item characteristics per medical condition are shown in online supplemental appendix 7 (outpatients) and online supplemental appendix 8 (inpatients).

Structural validity

Model fit indices for the five tested models are shown in table 3.

The best fit was found for the bifactor model with correlated dimensions. Here, all items loaded on their respective specific dimension and on a general dimension. Specific dimensions were intercorrelated, but their correlation with the general factor was restricted to be zero. Standardised factor loadings and AVE of this model are reported in table 4. Unstandardised factor loadings and error variances are shown in online supplemental appendix 2. Correlations between specific dimensions showed a wide variability ranging from –0.140 to 0.824 (see online supplemental appendix 9).

In the inpatient and outpatient samples, standardised loadings for the general factor were highest for the ‘Essential characteristics of the clinicians’ dimension and lowest for ‘Integration of medical and non-medical care’. On their specific factors, these two dimensions had the lowest and highest standardised loadings, respectively. The general factor AVE was highest for ‘Clinician–patient relationship’ for outpatients (41.0%) and ‘Coordination and continuity of care’ for inpatients (36.8%). Specific factors showed the highest AVE for ‘Essential characteristics of the clinicians’ in outpatients (87.9%) and for ‘Patient safety’ in inpatients (97.2%).

We further explored the general factor in the bifactor model by calculating correlations with other

Table 1 Sample characteristics

Characteristics	Outpatients	Inpatients
Total sample size, n	1092	931
In treatment for		
Cardiovascular disease, n (%)	277 (25.4)	286 (30.7)
Cancer, n (%)	273 (25.0)	351 (37.7)
Musculoskeletal disease, n (%)	217 (19.9)	92 (9.9)
Mental disorder, n (%)	273 (25.0)	202 (21.7)
No information, n (%)	52 (4.8)*	0 (0)
Age (in years), mean (SD)	53.1 (17.5)	56.0 (17.6)
No response, n (%)	23 (2.1)	19 (2.0)
Years since initial diagnosis, mean (SD)	11.4 (12.0)	8.5 (10.0)
No response, n (%)	124 (11.4)	118 (12.7)
Years as patient in this outpatient clinic, mean (SD)	4.53 (6.5)	—
No response, n (%)	129 (11.8)	—
Length of stay (in days), mean (SD)	—	18.1 (30.6)
No response, n (%)	—	74 (7.9)
Health literacy†, mean (SD)	50.7 (7.7)	50.6 (7.7)
No response, n (%)	439 (40.2)	304 (32.7)
Satisfaction‡, mean (SD)	27.5 (4.6)	28.0 (4.9)
No response, n (%)	15 (1.4)	13 (1.4)
Health status§, mean (SD)	3.4 (2.8)	3.4 (4.1)
No response, n (%)	44 (4.0)	36 (3.9)
Comorbidity (Do you have any further diseases?)		
Yes, n (%)	551 (50.5)	467 (50.2)
No, n (%)	450 (41.2)	393 (42.2)
No response, n (%)	91 (8.3)	71 (7.6)
Gender, n (%)		
Female	646 (59.2)	384 (41.2)
Male	420 (38.5)	517 (55.5)
Diverse	5 (0.5)	6 (0.6)
No response	21 (1.9)	24 (2.6)
Marital status, n (%)		
Unmarried and unpartnered	343 (31.4)	237 (25.5)
Married or partnered	552 (50.5)	524 (56.3)
Divorced	107 (9.8)	80 (8.6)
Widowed	61 (5.6)	54 (5.8)
No response	29 (2.7)	36 (3.9)
Formal education, n (%)		
Low¶	10 (0.9)	20 (2.1)
Intermediate**	404 (37.0)	383 (41.2)
High††	267 (24.5)	204 (21.9)
Very high‡‡	376 (34.4)	279 (30.0)
No response	23 (2.1)	27 (2.9)
Occupational status§§, n (%)		
Employed	447 (40.9)	357 (38.3)
Unemployed	66 (6.0)	64 (6.9)
Student/trainee	91 (8.3)	50 (5.4)
Parental leave/stay at home	57 (5.2)	35 (3.8)
Retired	394 (36.1)	392 (42.1)
No response	26 (2.4)	17 (1.8)
Health insurance§§, n (%)		
Statutory	910 (83.3)	721 (77.4)
Private	164 (15.0)	214 (23.0)
No response	32 (2.9)	34 (3.7)

Continued

Table 1 Continued

Characteristics	Outpatients	Inpatients
Migration background (Were you or your parents born in another country than Germany?)		
Yes, n (%)	167 (15.3)	150 (16.1)
No, n (%)	898 (82.2)	758 (81.4)
No response, n (%)	27 (2.5)	23 (2.5)
*These patients were recruited in a general practice. They assured us that they were treated for one of the four health conditions but failed to indicate which one in the questionnaire.		
†Health literacy measured by the total sum score of the European Health Literacy Questionnaire (HLS-EU-Q16), ⁵¹ range 16–64, high value=high health literacy.		
‡Satisfaction with care measured by German version of the 8-item Client Satisfaction Questionnaire (ZUF-8), ²³ range 8–32, high value=high satisfaction.		
§General health status measured by the first item of 12-item Short Form Survey (SF-12), ²² range 0–5, high value=low health status.		
¶Low=no formal degree or graduation after less than 10 years at school.		
**Intermediate=graduation after 9 or 10 years at school.		
††High=graduation after more than 10 years at school.		
‡‡Very high=college or university degree.		
§§Multiple answers possible.		
SD, standard deviation.		

measures. We found a strong correlation between the model-based scores of the general factor and satisfaction with care (outpatients: 0.73, inpatients: 0.83). Correlations of the general factor with health literacy, general health status, comorbidity, age and time since initial diagnosis were all below 0.3.

We were not able to examine measurement invariance across medical conditions. The combination of small sample sizes within condition subgroups, high rates of missing values due to ‘does not concern me’ and the complexity of the tested model led to estimation problems (eg, non-convergent models, not positive definite variance-covariance matrix of estimated parameters).

Reliability

McDonald’s Omega hierarchical was high for the general factor (outpatients: 0.868, inpatients: 0.908) and varied widely for the specific dimensions (outpatients: 0.214–0.872, inpatients: 0.128–0.847). To support interpretability of sum scores for specific dimensions, we examined their correlations with the model-based factor scores (online supplemental appendix 10). Results suggest that these sum scores measure a mixture of the general factor and the specific dimension.

Construct validity

Using model-based factor scores, discriminant validity was confirmed for all specific dimensions and the general factor (correlations with general health status <0.3). Convergent validity was confirmed for the general factor, which was strongly correlated with satisfaction with care, while the specific dimensions showed no strong correlations. Correlations for all dimensions and the general factor are shown in online supplemental appendix 11.

DISCUSSION

Summary of the findings

EPAT-64 is a PREM to assess 16 dimensions of PC from patients’ perspective. All items demonstrated high acceptability. Most items had low rates of patients answering ‘does not concern me’. Regarding item difficulty, most items were within the recommended range, but about 30% of the items showed ceiling effects. Item–total correlations were good for all but one item. CFA showed the best fit for a bifactor model. Discriminant validity was confirmed for all dimensions and the general factor. Convergent validity was confirmed for the general factor only. McDonald’s Omega hierarchical suggests good reliability of the general factor, while there was wide variability for the specific dimensions.

Strengths and limitations

EPAT-64 is the first PREM comprehensively assessing PC.^{15 16} One of its strengths is the thorough development and item selection process, which prioritised content validity.^{7 8 17} Some items of EPAT-64 showed suboptimal item characteristics, in particular ceiling effects, which may limit the ability to discriminate between high and low PC. These items were included nevertheless as we selected items based on several quantitative criteria and content. Further, we decided to choose four items per dimension, as we aimed for parsimony while requiring at least four items for statistical model identifiability.¹⁹ For dimensions that were covered by only four items in the preliminary EPAT,¹⁷ we had no choice between items. Deleting dimensions was not an option, as all dimensions are relevant to patients.⁸ We argue that selecting items solely based on quantitative data may lead to more reliable questionnaires, but risks compromising content validity.¹⁹ In particular in outpatients, some items had high rates of ‘does not concern me’ answers, especially those beginning with ‘if’ or ‘when’ conditions. For instance,

Table 2 Item characteristics of the EPAT-64													
Dimension	Item	Outpatient				Inpatient							
		M	SD	'Does not concern me' (%)	No reply (%)	Item difficulty	Item total	M	SD	'Does not concern me' (%)	No reply (%)	Item difficulty	Item total
Essential characteristics of the clinician	Item 1	5.1	1.2	4.2	1.2	0.82	0.73	5.1	1.2	3.9	0.8	0.82	0.73
	Item 2	5.5	0.9	0.7	0.7	0.91	0.63	5.5	0.9	0.2	0.9	0.90	0.65
	Item 3	5.2	1.1	4.5	0.8	0.85	0.71	5.3	1.0	2.9	1.1	0.85	0.65
	Item 4	4.8	1.5	32.8	1.1	0.77	0.73	4.7	1.5	30.3	1.0	0.75	0.68
Clinician–patient relationship	Item 1	5.2	1.1	1.8	1.2	0.85	0.71	5.3	1.1	1.2	0.4	0.86	0.70
	Item 2	4.9	1.4	12.8	1.0	0.78	0.73	4.9	1.3	12.9	0.8	0.79	0.73
	Item 3	5.1	1.2	3.8	0.5	0.82	0.57	4.8	1.6	12.6	0.6	0.75	0.53
	Item 4	5.0	1.3	23.0	0.7	0.79	0.69	5.2	1.0	1.3	0.5	0.84	0.62
Patient as a unique person	Item 1	4.8	1.4	8.3	0.8	0.75	0.78	4.8	1.3	5.5	0.9	0.76	0.78
	Item 2	5.1	1.2	2.0	0.5	0.82	0.73	5.1	1.2	1.6	0.6	0.82	0.71
	Item 3	4.1	1.8	12.5	0.8	0.62	0.72	4.1	1.7	12.8	0.9	0.62	0.68
	Item 4	3.7	1.8	17.7	0.7	0.55	0.71	3.7	1.7	14.5	0.9	0.54	0.64
Biopsychosocial perspective	Item 1	3.9	1.8	16.1	1.2	0.57	0.66	3.8	1.8	14.2	0.8	0.56	0.63
	Item 2	4.0	1.8	10.1	1.6	0.61	0.66	3.8	1.9	10.2	1.0	0.57	0.54
	Item 3	4.9	1.4	7.9	1.5	0.77	0.63	4.9	1.4	4.9	0.6	0.78	0.60
	Item 4	3.4	1.9	20.2	1.1	0.48	0.68	3.4	1.8	13.0	0.9	0.48	0.60
Clinician–patient communication	Item 1	5.3	1.1	2.6	1.0	0.86	0.69	5.2	1.1	2.8	0.5	0.85	0.71
	Item 2	5.4	1.0	0.8	0.8	0.87	0.48	5.2	1.1	0.5	0.5	0.84	0.55
	Item 3	5.5	0.9	0.4	0.5	0.90	0.64	5.5	0.9	0.2	0.4	0.90	0.64
	Item 4	5.0	1.3	5.2	0.9	0.80	0.71	5.0	1.2	2.3	0.9	0.80	0.71
Integration of medical and non-medical care	Item 1	3.1	2.0	30.4	1.1	0.42	0.54	3.4	2.0	25.8	1.5	0.49	0.57
	Item 2	4.5	1.8	57.5	1.7	0.71	0.52	4.3	1.8	56.7	2.6	0.66	0.63
	Item 3	3.1	1.9	43.3	1.7	0.42	0.62	3.3	1.9	40.9	2.4	0.46	0.61
	Item 4	3.3	2.1	46.1	1.9	0.45	0.60	3.6	2.0	42.3	2.7	0.52	0.62
Teamwork and team building	Item 1	5.1	1.0	7.1	0.8	0.83	0.41	5.1	1.1	1.2	1.0	0.82	0.64
	Item 2	5.0	1.2	12.4	1.4	0.80	0.50	5.1	1.2	1.4	1.2	0.82	0.65
	Item 3	4.6	1.4	22.1	2.1	0.72	0.59	5.0	1.2	3.0	1.6	0.81	0.64
	Item 4*	5.1	1.4	26.1	0.9	0.82	0.27	4.8	1.6	7.5	1.2	0.75	0.35
Access to care	Item 1	4.7	1.4	23.4	1.3	0.73	0.50	4.8	1.2	6.6	1.0	0.76	0.65
	Item 2	5.1	1.2	2.1	1.4	0.83	0.37	5.4	1.1	18.2	0.9	0.88	0.39
	Item 3	5.0	1.4	9.2	1.3	0.79	0.30	5.2	1.3	30.4	1.9	0.83	0.44
	Item 4	5.1	1.3	8.6	0.9	0.83	0.33	5.5	0.9	20.0	1.0	0.89	0.52

Continued

Table 2 Continued

Dimension	Outpatient				Inpatient								
	Item	M	SD	'Does not concern me' (%)	No reply (%)	Item difficulty	Item total	M	SD	'Does not concern me' (%)	No reply (%)	Item difficulty	Item total
Coordination and continuity of care	Item 1	5.1	1.4	13.1	1.4	0.82	0.52	5.0	1.4	7.2	1.0	0.80	0.56
	Item 2	3.9	1.8	11.0	0.9	0.58	0.44	4.2	1.6	9.0	1.4	0.65	0.61
	Item 3	5.3	1.1	0.5	0.9	0.85	0.70	5.1	1.1	0.5	0.5	0.83	0.71
	Item 4	3.5	2.0	27.2	1.5	0.50	0.52	4.8	1.4	6.7	1.7	0.76	0.69
Patient safety	Item 1	3.9	1.8	23.4	0.8	0.58	0.65	4.1	1.7	14.2	1.2	0.62	0.64
	Item 2	5.1	1.2	16.6	2.7	0.82	0.59	5.3	1.0	1.0	2.3	0.87	0.60
	Item 3	5.0	1.5	41.5	1.0	0.80	0.60	5.2	1.3	13.1	1.1	0.84	0.56
	Item 4	2.3	1.7	30.6	0.9	0.26	0.45	2.9	1.8	17.1	1.3	0.38	0.51
Patient information	Item 1	4.6	1.6	10.9	0.7	0.71	0.67	4.6	1.5	6.4	1.3	0.72	0.65
	Item 2	3.9	1.8	16.1	1.3	0.57	0.62	4.1	1.8	6.6	0.8	0.62	0.59
	Item 3	4.8	1.5	19.2	1.1	0.75	0.65	4.8	1.5	4.6	1.2	0.75	0.62
	Item 4	4.0	1.8	14.5	1.0	0.61	0.67	4.2	1.8	8.9	1.4	0.63	0.64
Patient involvement in care	Item 1	5.0	1.2	7.6	1.3	0.79	0.70	4.8	1.3	4.4	1.8	0.76	0.69
	Item 2	4.3	1.7	27.6	0.9	0.65	0.64	4.4	1.7	16.4	0.8	0.69	0.65
	Item 3	4.6	1.6	25.2	1.3	0.72	0.65	4.3	1.7	18.6	1.5	0.66	0.64
	Item 4	4.5	1.6	24.8	1.1	0.71	0.73	4.3	1.6	18.5	1.7	0.67	0.67
Involvement of family and friends	Item 1	3.1	2.0	42.5	1.3	0.41	0.64	3.3	1.9	29.6	1.0	0.46	0.62
	Item 2	3.0	2.0	60.6	1.5	0.41	0.66	3.1	1.9	50.2	1.4	0.43	0.58
	Item 3	3.8	2.1	57.5	1.5	0.57	0.63	4.1	1.9	40.3	1.4	0.62	0.61
	Item 4	4.0	2.1	56.4	1.6	0.60	0.63	4.0	1.9	43.0	1.9	0.59	0.63
Patient empowerment	Item 1	4.0	1.8	26.0	0.8	0.61	0.63	4.0	1.7	25.8	0.5	0.61	0.52
	Item 2	4.6	1.5	9.7	1.5	0.72	0.73	4.8	1.4	4.8	1.1	0.75	0.67
	Item 3	3.1	1.8	21.9	2.0	0.43	0.63	3.3	1.7	16.6	1.5	0.46	0.62
	Item 4	3.6	1.8	35.7	0.8	0.51	0.65	3.7	1.8	29.1	1.0	0.55	0.55
Physical support	Item 1	4.8	1.4	54.7	1.0	0.76	0.61	5.4	1.0	19.4	0.8	0.87	0.56
	Item 2	4.7	1.4	56.6	1.2	0.75	0.66	5.2	1.1	23.8	0.5	0.85	0.55
	Item 3	5.3	1.0	40.3	0.8	0.87	0.45	5.5	0.9	5.8	0.3	0.89	0.59
	Item 4	3.2	1.9	67.5	0.8	0.44	0.56	4.2	1.8	48.2	0.6	0.64	0.60
Emotional support	Item 1	4.4	1.7	19.2	0.8	0.68	0.76	4.4	1.6	15.9	1.5	0.69	0.72
	Item 2	4.4	1.7	16.8	0.7	0.69	0.70	4.3	1.7	18.3	1.4	0.66	0.63
	Item 3	3.8	1.9	17.9	1.3	0.57	0.71	3.7	1.8	19.5	1.7	0.54	0.57
	Item 4	3.5	2.1	31.5	0.9	0.50	0.58	3.9	2.0	23.4	1.0	0.58	0.57

* Item was reversed prior to analyses.
EPAT-64, 64-item version of the Experienced Patient-Centeredness Questionnaire; M, mean; SD, standard deviation.

Table 3 Model fit indices for confirmatory factor analysis on EPAT-64

			χ^2								
			χ^2	df	P value	RMSEA	SRMR	TLI	CFI	AIC	BIC
Recommendation					>0.05	≤0.06	<0.08	≥0.90	≥0.90	Model with smallest value	
Outpatient version	1	Unidimensional	14 957	1952	<0.001	0.078	0.087	0.659	0.670	166 616	167 575
	2	Correlated first order	7098	1832	<0.001	0.051	0.072	0.853	0.866	158 996	160 555
	3	Hierarchical	8556	1936	<0.001	0.056	0.084	0.825	0.832	160 247	161 286
	4	Bifactor without dimension correlations	No results. Model estimation did not converge.								
	5	Bifactor with dimension correlations	4871	1768	<0.001	0.040	0.055	0.910	0.921	156 898	158 776
Inpatient version	1	Unidimensional	15 201	1952	<0.001	0.085	0.101	0.631	0.643	151 891	152 819
	2	Correlated first order	7276	1832	<0.001	0.056	0.083	0.838	0.853	144 206	145 714
	3	Hierarchical	8839	1936	<0.001	0.062	0.102	0.806	0.814	145 561	146 567
	4	Bifactor without dimension correlations	No results. Model estimation did not converge.								
	5	Bifactor with dimension correlations	4721	1768	<0.001	0.042	0.051	0.909	0.920	141 779	143 597
AIC, Akaike information criterion; BIC, Bayesian information criterion; CFI, Comparative Fit Index; EPAT-64, 64-item version of the Experienced Patient-Centeredness Questionnaire; RMSEA, root mean square error of approximation; SRMR, standardised root mean square residual; TLI, Tucker-Lewis Index.											

AIC, Akaike information criterion; BIC, Bayesian information criterion; CFI, Comparative Fit Index; EPAT-64, 64-item version of the Experienced Patient-Centeredness Questionnaire; RMSEA, root mean square error of approximation; SRMR, standardised root mean square residual; TLI, Tucker-Lewis Index.

54.7% of the outpatient sample replied this way to the item ‘When I had pain, I was helped quickly’. However, this item is essential for the definition of the ‘Physical support’ dimension⁸ and highly relevant for everyone with pain. Overall, some items being less relevant to certain patient groups influences EPAT-64’s generic usability.

CFA showed best fit for a bifactor model with correlated dimensions. This model postulates that all items load on a general factor and on their specific dimension, which is independent of the general factor. How well items’ variance was explained by their specific and general factors and specific factors’ reliability varied. Thus, not all items are well represented by the bifactor model. We decided against further modifying this model, as factor structure decisions should combine data and theory-driven approaches to avoid overfitting and ensure meaningful interpretation.^{7–9 17} Overall, the bifactor model showed good results for global model fit indices, reflecting our underlying model of PC, but the local fit for specific items needs further investigation. An advantage of the bifactor model is that it enables users to partition scores into an estimation of the general factor and of the specific dimension.⁴⁰ The general factor likely captures more subjective aspects of PC, as it strongly correlates with satisfaction with care (see convergent validity). The specific dimensions of PC seem to explain variance not captured by the general factor, confirming the underlying conceptual model. We assume that these dimensions capture more objective aspects of PC (ie, whether something happened rather than how well it went). No strong correlations between their factor scores and satisfaction also indicate this (see convergent validity). This differentiation is important in light of definitions of PREMs¹² which stress patient experience and satisfaction being

two distinct constructs.⁴¹ Satisfaction is subjective and influenced by individual expectations, patient experience is more objective and focuses on whether certain processes occurred.¹² While EPAT-64 was designed as a PREM, the general factor reflects satisfaction. Since experiences and satisfaction are ‘complexly related’⁴¹ and all self-reports introduce some subjectivity,⁴² PREMs likely correlate with satisfaction. The EPAT-64 bifactor model may allow statistical differentiation with the general factor likely relating to satisfaction and the specific dimensions likely capturing experiences. Thus, we recommend users to apply the whole EPAT-64. This way, they can model the bifactor model in their data and calculate both general and dimension scores. Yet, in routine care, there might be sparse resources for administering a 64-item questionnaire or insufficient sample size for estimating bifactor models. Hence, choosing the most relevant dimensions for a context and calculating sum scores per dimension might be more feasible. This flexibility allows adapting the questionnaire to specific questions, areas of application and resources. Missing values should best be dealt with using multiple imputation or full information maximum likelihood.⁴³ In routine care, when sample size and resources are limited, one missing value per person and dimension can be replaced by the average of the other three items of that person.⁴³ For persons with more missings, no sum score for the respective dimension should be calculated. When interpreting those sum scores, one should consider that this score combines subjective and objective PC assessments. Further, the specific dimensions are differently reliable and correlations between them make interpretation of sum scores more challenging. Online supplemental appendix 10 helps guide sum score interpretation, but more research is needed.

Table 4 Standardised factor loadings, average variance explained and reliability coefficients of the correlated bifactor model												
Dimension	Standardised loadings on general factor				AVE (%)	Standardised loadings on specific dimension				Reliability		
	Item 1	Item 2	Item 3	Item 4		Item 1	Item 2	Item 3	Item 4	AVE (%)	Omega_H	
Outpatient sample	Essential characteristics of the clinicians	0.759	0.795	0.767	0.575	7.8	0.440	0.216	0.387	0.571	87.9	0.214
	Clinician–patient relationship	0.754	0.682	0.565	0.627	41.0	0.474	0.595	0.225	0.345	26.3	0.242
	Patient as a unique person	0.704	0.795	0.366	0.288	8.1	0.498	0.332	0.782	0.833	85.1	0.515
	Biopsychosocial perspective	0.495	0.290	0.352	0.243	15.8	0.488	0.805	0.717	0.814	40.9	0.705
	Clinician–patient communication	0.760	0.457	0.739	0.608	22.9	0.371	0.424	0.384	0.498	57.6	0.256
	Integration of medical and non-medical care	0.143	0.218	0.188	0.165	6.0	0.836	0.731	0.829	0.854	41.8	0.859
	Teamwork and team building	0.347	0.417	0.493	0.390	13.2	0.576	0.689	0.518	0.884	47.0	0.415
	Access to care	0.475	0.323	0.235	0.283	11.4	0.502	0.712	0.638	0.535	41.8	0.599
	Coordination and continuity of care	0.430	0.198	0.739	0.291	14.6	0.345	0.516	0.318	0.505	23.5	0.355
	Patient safety	0.458	0.359	0.586	0.114	4.6	0.453	0.683	0.298	0.685	80.5	0.506
	Patient information	0.489	0.374	0.534	0.426	23.6	0.620	0.649	0.564	0.665	33.3	0.559
	Patient involvement in care	0.732	0.447	0.506	0.529	12.2	0.356	0.566	0.651	0.734	74.7	0.461
	Involvement of family and friends	0.187	0.201	0.233	0.255	10.1	0.882	0.895	0.825	0.791	45.7	0.872
	Patient empowerment	0.232	0.570	0.227	0.224	7.7	0.764	0.515	0.732	0.790	65.9	0.718
Physical support	0.567	0.586	0.457	0.190	18.7	0.706	0.725	0.209	0.443	34.4	0.461	
Emotional support	0.516	0.464	0.359	0.196	19.9	0.755	0.766	0.857	0.680	46.1	0.736	
General factor											0.868	
											Continued	

Continued

Table 4 Continued

Dimension	Standardised loadings on general factor				Standardised loadings on specific dimension				Reliability			
	Item 1	Item 2	Item 3	Item 4	AVE (%)	Item 1	Item 2	Item 3	Item 4	AVE (%)	Omega_H	
Inpatient sample	Essential characteristics of the clinicians	0.783	0.830	0.758	0.575	7.6	0.312	0.155	0.230	0.517	88.0	0.128
	Clinician–patient relationship	0.785	0.674	0.323	0.657	31.1	0.338	0.539	0.612	0.263	38.6	0.291
	Patient as a unique person	0.720	0.775	0.340	0.250	4.7	0.444	0.238	0.779	0.869	90.6	0.484
	Biopsychosocial perspective	0.257	0.180	0.445	0.263	11.8	0.813	0.776	0.476	0.826	39.0	0.757
	Clinician–patient communication	0.744	0.559	0.727	0.798	26.2	0.334	0.483	0.352	0.395	56.0	0.215
	Integration of medical and non-medical care	0.233	0.286	0.265	0.252	13.7	0.833	0.835	0.824	0.846	45.1	0.847
	Teamwork and team building	0.718	0.744	0.643	0.448	31.4	0.465	0.425	0.553	0.918	35.3	0.213
	Access to care	0.692	0.541	0.528	0.655	28.9	0.114	0.538	0.711	0.821	44.1	0.217
	Coordination and continuity of care	0.521	0.459	0.784	0.659	36.8	0.288	0.518	0.233	0.355	18.8	0.199
	Patient safety	0.392	0.734	0.558	0.256	1.0	0.695	0.123	0.249	0.726	97.2	0.365
	Patient information	0.489	0.378	0.592	0.538	29.9	0.622	0.673	0.443	0.575	26.6	0.493
	Patient involvement in care	0.695	0.534	0.532	0.479	14.7	0.378	0.564	0.676	0.738	71.5	0.474
	Involvement of family and friends	0.273	0.265	0.415	0.378	21.4	0.850	0.862	0.689	0.755	38.0	0.780
	Patient empowerment	0.212	0.572	0.330	0.223	6.0	0.742	0.398	0.687	0.866	75.8	0.668
Physical support	0.662	0.670	0.677	0.415	27.6	0.622	0.632	0.169	0.230	31.5	0.268	
Emotional support	0.493	0.327	0.186	0.272	17.0	0.735	0.846	0.878	0.619	39.4	0.768	
General factor											0.908	
AVE, average variance explained (average percentage of the variance in the indicators which is explained by the respective latent factors).												

Original research

As a limitation, we were not able to examine measurement invariance across different medical conditions. To make transferability across medical conditions likely, we developed items generically and tested them in four medical conditions that affect a large proportion of the population. Yet, due to the complex measurement model and the large number of ‘does not concern me’ replies in some items, we were not able to confirm a common factor structure across conditions. Further, the high dropout rate after consent is a limitation. As self-selection bias is possible, acceptance and feasibility of EPAT-64 data collection need further testing. The ongoing SARS-CoV-2 pandemic might have influenced our results, in particular, dimensions like ‘Involvement of family and friends’ or ‘Access to care’. Furthermore, pandemic access restrictions led us to adding an online survey. While this broadened the sample and might have improved generalisability, there might be differences between online and paper-and-pencil subgroups. Paper-and-pencil participants were treated in or near Hamburg and received the questionnaire on the day of their treatment; online participants came from across Germany and might have responded much later. Further investigation of influences of data collection modality on EPAT-64 is needed. Notably, since we aimed to design a PREM, we defined PC based on interactions between patients and healthcare professionals. Other models view PC more broadly and include aspects that occur ‘further away’ from the care system (eg, physical activities and self-management practices).

Implications

There is a wide range of possible applications of EPAT-64. Measuring PC is essential for research on the current state of PC and on fostering PC. Internationally, PREMs are used to provide feedback to healthcare providers, inform quality improvement, public reporting, benchmarking and value-based purchasing.^{44–46} To improve patient experiences, data from PREMs need to be used in a coordinated approach.⁴⁷ Emerging evidence suggests that being informed about patient experiences is associated with improved communication,⁴⁸ and meetings to discuss PREM results are associated with improved patient experience.⁴⁵

Psychometric testing in independent samples is indispensable to ensure that the psychometric properties found here hold in other populations and contexts. Particular attention should be paid to the replication of the bifactor structure with correlated dimensions and measurement invariance. In addition to validity and reliability, aspects of fairness and responsiveness should be investigated.⁴⁹ Regarding testing in other patient groups, the EPAT-64 is currently adapted to assess PC in the context of healthcare and support services for women with unintended pregnancy.⁵⁰ Further adaptations to other healthcare contexts, such

as rehabilitation facilities or non-academic hospitals, could be considered.

Internationally, no comparable PREM that comprehensively assesses all dimensions of PC exists.¹⁵ Therefore, cross-cultural adaptations and translations of the EPAT-64 could provide an opportunity to promote PC around the world. In addition, translations would help reach people in Germany not fluent in German. An English EPAT-64 translation that used the TRAPD process²¹ can be downloaded from www.uke.de/epaten. However, further tests on comprehensibility and psychometric properties are required before it can be used.

CONCLUSION

The EPAT-64, a novel questionnaire assessing 16 dimensions of PC from the patient’s perspective, emerges from a meticulous development process that blends theory-driven with data-driven decisions. Psychometric testing within a substantial patient cohort underscores its robust psychometric properties. Nevertheless, further research is imperative to refine the interpretation of scores. The adaptability of this tool to specific purposes positions it as a valuable instrument for both assessment and cultivation of PC.

Author affiliations

¹Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

²Department of Internal Medicine II (Oncology/Hematology/BMT/Pneumology), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

³Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁴Department of Dermatology and Venereology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁵Department of Cardiology, University Heart & Vascular Center Hamburg, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁶German Center for Cardiovascular Research (DZHK), partner site Hamburg/Kiel/Luebeck, Germany

⁷Center for Population Health Innovation (POINT), University Heart and Vascular Center Hamburg, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁸Department of Gynecology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁹Department of General Practice and Primary Care, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹⁰Department of Orthopedics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Acknowledgements We thank our student assistant Sibel Gürü and our research interns Lucas Bolz, Antonia Buchal, Sophia Ehrentraut, Carolin Guddat, Zoe Henning, Constanze Mahn, Yannick Mosemann, Philip Niemann, Svea-Marie Nohrden, Paul Packheiser, Ronja Platzeck, Mareike Remmers, Ramona Rusch, Laura Stefan, Melvin Schneider and Lisa Strelow for carrying out data collection. Further, we thank all our collaboration partners who gave us access to their patients for the data collection: Klinik für Psychosomatische Medizin und Psychotherapie (UKE) III, Medizinische Klinik und Poliklinik (UKE), Strahlentherapie und Radioonkologie (UKE), Martini-Klinik, Evangelisches Krankenhaus Alsterdorf, Verhaltenstherapie Falkenried, Bundeswehrkrankenhaus Hamburg, Asklepios Westklinikum Hamburg, Psychiatrische Ambulanz Dr Guth and Fachzentrum Altona für Psychiatrie und Psychotherapie.

Contributors IS was the responsible principal investigator of the study and is the guarantor. IS, LK and MH were involved in planning and preparation of the study. SZ, AS, HF and EC recruited the participants and collected the data. EC analysed the data with supervision from LK. EC, SZ, JMZ, IS, MH, PH and LK interpreted the results and discussed the item selection. EC wrote the first draft of the manuscript. CB, JG, CG, CM, VM, KS-B and AS actively supported the data collection. All authors critically revised the manuscript for important intellectual content. All authors gave final approval of the version to be published and agreed to be accountable for the work.

Funding This study was funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung – BMBF) with the grant number 01GY1614. The funder had no role in decision to publish or preparation of the manuscript.

Competing interests PH is a board member of the International Shared Decision Making Society, a charitable scientific society. IS received honoraria for presentations and speeches on patient-centred care from the following commercial entities: onkowissen.de GmbH and ClinSol GmbH & Co KG.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Ethics Committee of the Medical Association Hamburg (study ID: PV5724). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Eva Christalle <http://orcid.org/0000-0001-7813-5800>

Levente Kriston <http://orcid.org/0000-0003-0748-264X>

REFERENCES

- 1 Institute of Medicine. *Crossing the quality chasm: a new health system for the 21st century*. Washington, DC: The National Academies Press, 2001.
- 2 Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med* 2014;12:573–6.
- 3 Bertakis KD, Azari R. Patient-centered care is associated with decreased health care utilization. *J Am Board Fam Med* 2011;24:229–39.
- 4 Dwamena F, Holmes-Rovner M, Gaulden CM, *et al*. Interventions for providers to promote a patient-centred approach in clinical consultations. *Cochrane Database Syst Rev* 2012;12:CD003267.
- 5 McMillan SS, Kendall E, Sav A, *et al*. Patient-centered approaches to health care: a systematic review of randomized controlled trials. *Med Care Res Rev* 2013;70:567–96.
- 6 Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev* 2013;70:351–79.
- 7 Scholl I, Zill JM, Härter M, *et al*. An integrative model of patient-centeredness - a systematic review and concept analysis. *PLoS One* 2014;9:e107828.
- 8 Zeh S, Christalle E, Hahlweg P, *et al*. Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ Open* 2019;9:e031741.
- 9 Zill JM, Scholl I, Härter M, *et al*. Which Dimensions of Patient-Centeredness Matter? - Results of a Web-Based Expert Delphi Survey. *PLoS ONE* 2015;10:e0141978.
- 10 Epstein RM, Street RL. The values and value of patient-centered care. *Ann Fam Med* 2011;9:100–3.
- 11 Browne K, Roseman D, Shaller D, *et al*. Analysis & commentary. Measuring patient experience as a strategy for improving primary care. *Health Aff (Millwood)* 2010;29:921–5.
- 12 Beattie M, Murphy DJ, Atherton I, *et al*. Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Syst Rev* 2015;4:97:97.
- 13 Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open* 2013;3:e001570.
- 14 Anhang Price R, Elliott MN, Zaslavsky AM, *et al*. Examining the role of patient experience surveys in measuring health care quality. *Med Care Res Rev* 2014;71:522–54.
- 15 Bull C, Byrnes J, Hettiarachchi R, *et al*. A systematic review of the validity and reliability of patient-reported experience measures. *Health Serv Res* 2019;54:1023–35.
- 16 Mihaljevic AL, Doerr-Harim C, Kalkum E, *et al*. Measuring patient centeredness with German language Patient-Reported Experience Measures (PREM)-A systematic review and qualitative analysis according to COSMIN. *PLoS One* 2022;17:e0264045.
- 17 Christalle E, Zeh S, Hahlweg P, *et al*. Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)-A best practice example for generating patient-reported measures from qualitative data. *Health Expect* 2022;25:1529–38.
- 18 Christalle E, Zeh S, Hahlweg P, *et al*. Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open* 2018;8:e025896.
- 19 Diamantopoulos A, Sarstedt M, Fuchs C, *et al*. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *J Acad Mark Sci* 2012;40:434–49.
- 20 Gagnier JJ, Lai J, Mokkink LB, *et al*. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res* 2021;30:2197–218.
- 21 Harkness J, Pennell BE, Schoua-Glusberg A. Survey questionnaire translation and assessment. In: *Methods for testing and evaluating survey questionnaires*. 2004: 453–73.
- 22 Wirtz MA, Morfeld M, Glaesmer H, *et al*. Konfirmatorische Prüfung der Skalenstruktur des SF-12 version 2.0 in einer

- deutschen bevölkerungs-repräsentativen Stichprobe. *Diagn* 2018;64:84–96.
- 23 Kriz D, Nübling R, Steffanowski A, *et al.* Patientenzufriedenheit in der stationären Rehabilitation: Psychometrische Reanalyse des ZUF-8 auf der Basis multizentrischer Stichproben verschiedener Indikation. *Z Med Psychol* 2008;17:67–79.
 - 24 Röthlin F, Pelikan JM, Ganahl K. Die Gesundheitskompetenz der 15-jährigen Jugendlichen in Österreich. Abschlussbericht der österreichischen Gesundheitskompetenz Jugendstudie im Auftrag des Hauptverbands der österreichischen Sozialversicherungsträger (HVSFV). 2013.
 - 25 Jordan S, Hoebel J. Gesundheitskompetenz von Erwachsenen in Deutschland. *Bundesgesundheitsbl* 2015;58:942–50.
 - 26 Sepucha KR, Matlock DD, Wills CE, *et al.* “It’s Valid and Reliable” Is Not Enough: Critical Appraisal of Reporting of Measures in Trials Evaluating Patient Decision Aids. *Med Decis Making* 2014;34:560–6.
 - 27 Döring N, Bortz J. *Forschungsmethoden und evaluation*. Wiesbaden: Springerverlag, 2006.
 - 28 Boateng GO, Neilands TB, Frongillo EA, *et al.* Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health* 2018;6:149.
 - 29 Flora DB, Flake JK. The purpose and practice of exploratory and confirmatory factor analysis in psychological research: decisions for scale development and validation. *Can J Behav Sci / Rev Can Des Sci Du Comport* 2017;49:78–88.
 - 30 Suhr DD. Exploratory or confirmatory factor analysis? 2006.
 - 31 Bandalos DL, Finney SJ. Factor analysis: exploratory and confirmatory. In: *The reviewer’s guide to quantitative methods in the social sciences*. Routledge, 2018: 98–122.
 - 32 Irwing P, Booth T, Hughes DJ. *The Wiley handbook of psychometric testing: a multidisciplinary reference on survey, scale and test development*. John Wiley & Sons, 2018.
 - 33 Cochran WG. The χ^2 test of goodness of fit. In: *The annals of mathematical statistics*. 1952: 315–45.
 - 34 Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychol Methods* 2016;21:137–50.
 - 35 Olstedal S, Garratt A, Bjertnaes Ø, *et al.* The NORPEQ patient experiences questionnaire: data quality, internal consistency and validity following a Norwegian inpatient survey. *Scand J Public Health* 2007;35:540–7.
 - 36 Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw* 2012;48:1–36.
 - 37 Dueber D. Package ‘Bifactorindicescalculator’. 2017. Available: <https://CRAN.R-project.org/package=BifactorIndicesCalculator> [Accessed 29 Jan 2024].
 - 38 Revelle W. *Psych: procedures for psychological, psychometric, and personality research. R Package Version 2.2.5*. Evanston: Northwestern University, 2022. Available: <https://CRAN.R-project.org/package=psych>
 - 39 Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Eq Model: A Multidiscipl J* 2001;8:430–57.
 - 40 Chen FF, Hayes A, Carver CS, *et al.* Modeling general and specific variance in multifaceted constructs: a comparison of the bifactor model to other approaches. *J Pers* 2012;80:219–51.
 - 41 Bull C. Patient satisfaction and patient experience are not interchangeable concepts. *Int J Qual Health Care* 2021;33:mzab023.
 - 42 Larson E, Sharma J, Bohren MA, *et al.* When the patient is the expert: measuring patient experience and satisfaction with care. *Bull World Health Organ* 2019;97:563–9.
 - 43 Newman DA. Missing data: five practical guidelines. *Org Res Methods* 2014;17:372–411.
 - 44 Bull C, Callander EJ. Current PROM and PREM use in health system performance measurement: still a way to go. *Pat Exp J* 2022;9:12–8.
 - 45 Reeves R, West E, Barron D. Facilitated patient experience feedback can improve nursing care: a pilot study for a phase III cluster randomised controlled trial. *BMC Health Serv Res* 2013;13:259.
 - 46 Hibbard JH, Stockard J, Tusler M. Hospital performance reports: impact on quality, market share, and reputation. *Health Aff (Millwood)* 2005;24:1150–60.
 - 47 Coulter A, Locock L, Ziebland S, *et al.* Collecting data on patient experience is not enough: they must be used to improve care. *BMJ* 2014;348:g2225.
 - 48 Murante AM, Vainieri M, Rojas D, *et al.* Does feedback influence patient - professional communication? Empirical evidence from Italy. *Health Policy* 2014;116:273–80.
 - 49 Eignor DR. *The standards for educational and psychological testing*. 2013.
 - 50 Zill JM, Lindig A, Reck LM, *et al.* Assessment of person-centeredness in healthcare and social support services for women with unintended pregnancy (CarePreg): protocol for a mixed-method study. *BMJ Open* 2022;12:e066939.
 - 51 Jordan S, Hoebel J. [Health literacy of adults in Germany: Findings from the German Health Update (GEDA) study]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2015;58:942–50.

9.5. Publication 5

Publication 5: Christalle, E., Kriston, L., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J.M., Härter, M. & Scholl, I. (2025). Development and Psychometric Testing of EPAT-16: A Short and Valid Measure for Patient-Centeredness From the Patient's Perspective. *Health Expectations*, 28(3), e70296.

ORIGINAL ARTICLE OPEN ACCESS

Development and Psychometric Testing of EPAT-16: A Short and Valid Measure for Patient-Centeredness From the Patient's Perspective

Eva Christalle  | Levente Kriston  | Stefan Zeh  | Hannah Führes  | Alica Schellhorn | Pola Hahlweg  | Jödis Maria Zill  | Martin Härter  | Isabelle Scholl 

Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Germany

Correspondence: Isabelle Scholl (i.scholl@uke.de)

Received: 24 January 2025 | **Revised:** 7 April 2025 | **Accepted:** 5 May 2025

Funding: This study was supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung – BMBF) (no. 01GY1614).

ABSTRACT

Background: We previously developed the EPAT-64, a patient-reported experience measure designed to assess patient-centeredness (PC) across 16 dimensions of the validated integrative model of PC. While its modular structure makes it highly adaptable to examine certain dimensions of PC, generating an overall PC score requires assessing all 64 items. This is often challenging in routine settings with limited resources. Therefore, we developed and psychometrically tested a 16-item short form.

Methods: We conducted a cross-sectional study involving adult inpatients and outpatients receiving treatment for cardiovascular diseases, cancer, musculoskeletal disorders and mental health conditions in Germany. To ensure comprehensive content coverage, we selected one item per dimension based on content relevance as well as item characteristics such as item difficulty and item-total correlation. We assessed the structural validity of a unidimensional model using confirmatory factor analysis (CFA), measured reliability with McDonald's Omega and evaluated construct validity by investigating the inter-correlation of the EPAT-16 sum score with measures of general health status and satisfaction with care.

Results: All items of the final EPAT-16 showed high acceptability and good item-total correlations, with approximately two-thirds demonstrating appropriate item difficulty. CFA showed a slight misfit of the unidimensional model, but a high average variance explained. McDonald's Omega showed high reliability. All hypotheses about construct validity were confirmed.

Conclusions: The EPAT-16 demonstrated good psychometric properties, making it a feasible tool for assessing overall PC when resource constraints preclude using the full EPAT-64. In particular, it can be used in routine care for feedback and quality improvement, as well as in research to assess relationships with other relevant variables. Since its items were designed generically, it can be used for different medical conditions and settings, for example for public reporting. Future research should evaluate the EPAT-16 in diverse, independent patient samples to confirm that its positive characteristics are consistent across different populations and cultures.

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; CFA, confirmatory factor analysis; CFI, Comparative Fit Index; df, degrees of freedom; FIML, full information maximum likelihood; HLS-EU-Q16, European Health Literacy Questionnaire; M, mean; MLR, robust maximum likelihood estimator; PC, patient-centeredness; PREMs, patient-reported experience measures; PROMs, patient-reported outcome measures; RMSEA, root mean square error of approximation; SD, standard deviation; SRMR, standardized root mean square residual; TLI, Tucker-Lewis Index; ZUF-8, 8-item client satisfaction questionnaire.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Health Expectations* published by John Wiley & Sons Ltd.

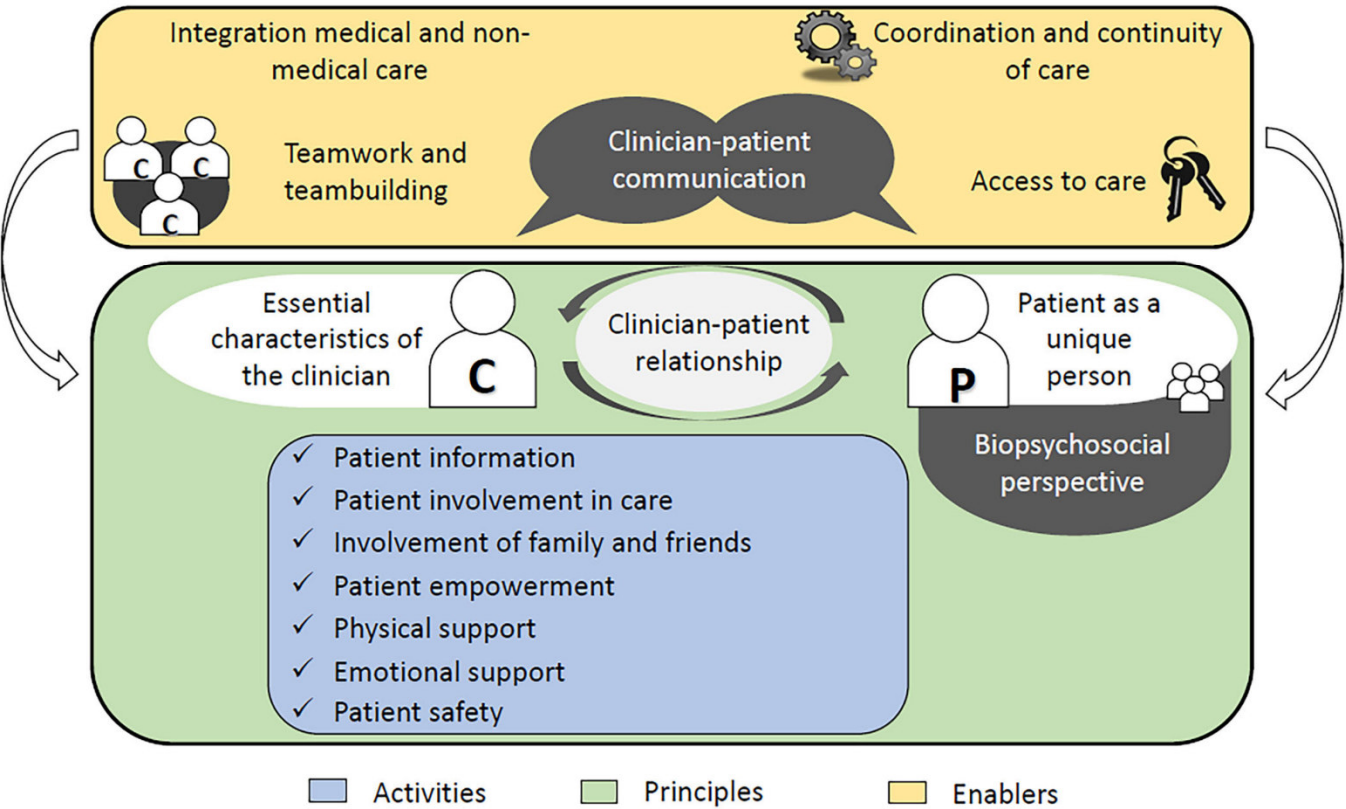
Patient or Public Contribution: Patients did not participate as active members of the research team. However, for the research proposal, several patient organizations provided positive feedback on the study aims and their relevance, signed a collaboration agreement and later supported recruitment by distributing study information.

1 | Introduction

Robust measures for assessing the degree of patient-centeredness (PC) are essential for promoting and improving PC in healthcare settings [1, 2]. A commonly used approach to evaluate PC from the patient's perspective is through patient-reported experience measures (PREMs), which focus on what or how patients experienced specific processes or behaviours in their healthcare [3, 4]. We developed the EPAT-64 questionnaire [5], designed to assess 16 dimensions of PC, based on the integrative model of PC (see Figure 1) [6, 7].

The development of the EPAT-64 followed a rigorous process, combining both theory- and data-driven approaches, and including multiple steps to involve patients in the development [8]. It has demonstrated robust psychometric properties in a sample of 2024 patients with chronic diseases [5]. The EPAT-64 is available in both outpatient and inpatient versions, with each of the 16 dimensions represented by a module consisting of four items. Users can choose to use all 64 items or select specific modules that are most relevant to their objectives. While this modular approach

ensures high adaptability, calculating an overall PC score requires administering all 64 items, which can be challenging. In particular, for routine care—given constraints such as limited time, personnel and financial resources—there is a need for a shorter version of the EPAT-64. Short scales can provide more efficient assessments in terms of time and cost [9]. Additionally, shorter instruments reduce the burden on respondents [10], an important consideration for patients who may be cognitively impaired or fatigued due to treatments or their medical conditions. While longer measures offer greater content coverage, they can lead to higher dropout rates, thus compromising the representativeness of the sample [11]. Reducing the number of items usually comes at the cost of reducing reliability, thus making short forms less applicable for individual-level decision making. However, PC is usually assessed for a whole unit of care with feedback from several patients, which increases the reliability of the average score. Further, short forms are useful tools for multivariate studies, which examine a high number of variables and thus need to assess each variable with concise questionnaires to reduce the burden on participants. Furthermore, short forms are in general useful tools for research questions focused on correlational relationships [11].



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). To view a copy of the license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/>
Authors: Isabelle Scholl, Martin Härter, Jörg Dirmaier, Stefan Zeh, Eva Christalle, Pola Hahlweg, & Jödis Zill (University Medical Center, Hamburg, Germany)

FIGURE 1 | Integrative model of PC [6, 7].

The 16 dimensions of PC of the integrative model are considered to be interrelated [6]. Psychometric analyses of the EPAT-64 revealed that the measure can be effectively represented by a bifactor model, which includes a general factor loading on all items as well as 16 specific factors representing each dimension [5]. In this model, the general factor showed a strong correlation with treatment satisfaction, suggesting it may reflect the more subjective aspects of patient experiences. While the general factor was independent of the specific dimensions, the dimensions themselves were interrelated. Consequently, both theoretical and empirical evidence support the feasibility of deriving an overall PC score. However, it is important to note that such a score would encompass a combination of both the general factor and the specific factors. To preserve the breadth of the construct PC, each of the 16 dimensions should be represented in the short form [12].

In this paper, we describe the development process of the EPAT-16, a shortened version of the EPAT-64, which evaluates each of the 16 dimensions using a single item. We present the psychometric properties of the EPAT-16 based on the sample used for item selection. Our goal was to design a tool to assess overall PC that can be used for all patients in routine care (e.g. as tool for continuous feedback or quality improvement), in research (e.g. to examine the relationship of a high number of constructs) and to compare different healthcare institutions regarding their overall level of PC.

2 | Methods

2.1 | Study Design

The EPAT-64 and its short form, the EPAT-16, were developed and tested as part of the ASPIRED study (Assessment of Patient-Centeredness Through Patient-Reported Experience Measures) [13], a 6-year mixed-methods research project on the measurement of patients' perspectives on PC. Here, we report on the item selection process and psychometric evaluation of the EPAT-16, using the same data set that was used for the development and evaluation of the EPAT-64 [5]. Since no specific guidelines exist for reporting psychometric evaluations of PREMs, we followed the COSMIN guidelines for reporting patient-reported outcome measures (PROMs) where applicable (see Appendix 1).

2.2 | The EPAT Questionnaire

For psychometric testing, we conducted a cross-sectional survey that included all preliminary EPAT items. We developed these items based on the integrative model of PC [6, 7] shown in Figure 1, informed by a literature review, key informant interviews with experts and focus groups with patients with chronic diseases [8]. The preliminary questionnaire assessed 16 dimensions of PC, comprising 121 items for inpatients and 120 items for outpatients. Our goal was to create a generic instrument that would be applicable across different diseases. Participants rated the EPAT items on a 6-point Likert scale of agreement (1 = completely disagree, 2 = strongly disagree,

3 = somewhat disagree, 4 = somewhat agree, 5 = strongly agree, 6 = completely agree), with an additional 'does not concern me' option for each item. Higher scores indicated a greater degree of PC, except for nine items, which we reverse-coded before analysis. We originally developed and tested the EPAT in German. For this paper, we translated the EPAT-16 into English using the TRAPD process, which involves a team-based approach consisting of translation, review, adjudication, pre-testing and documentation [14].

2.3 | Data Collection and Participants

We used the same sample for the EPAT-16 item selection that we had used for the EPAT-64. We collected data through both paper-pencil questionnaires and an online survey conducted via LimeSurvey (LimeSurvey GmbH, Hamburg, Germany). For the paper-pencil data collection, we used consecutive sampling at 10 inpatient and 14 outpatient healthcare institutions (e.g. hospitals, primary care centres) in the metropolitan area of Hamburg, Germany, from June 2020 to February 2022. Outpatients received the questionnaire before their appointment and were asked to report on their experiences in the past 4 weeks. Inpatients received the questionnaire shortly before discharge and were asked to report on their entire inpatient stay. For the online survey, we recruited participants using community-based strategies, including social media and patient organizations from across Germany. Online participants were asked to refer to their last outpatient visit or inpatient stay within the past month. We included adults (18 years or older) with at least one of the following four medical conditions: cardiovascular diseases, cancer, musculoskeletal diseases and mental disorders. Inclusion was based on self-report. As outlined in our study protocol, we aimed for a sample size of 250 participants in each of the two settings (outpatient and inpatient) for each of the four medical conditions, resulting in a target sample size of 2000 participants [13].

The full questionnaire for data collection included the preliminary version of the EPAT, the first item of the German version of the 12-item Short Form Survey [15], the German version of the 8-item Client Satisfaction Questionnaire (ZUF-8) [16], the European Health Literacy Questionnaire (HLS-EU-Q16) [17] in German as well as questions on demographic characteristics and health status. We also added two control items (e.g. 'To show that you are reading attentively, please check "somewhat disagree"/"somewhat agree"'). Data collection was anonymous. Participants gave explicit informed consent (either written before receiving the paper-pencil questionnaire or by active click at the start of the online survey) and could opt to receive a €10 incentive for completing the questionnaire.

2.4 | Data Analyses

2.4.1 | Item Characteristics and Item Selection Strategy

We selected the items for the EPAT-16 during the same team discussion session in which we chose the items for the EPAT-64 [5]. Given that the original instrument assesses 16 dimensions, it has a

rather broad scope. Even though the EPAT-16 omits those dimensions and instead aims to measure an overall degree of PC, it was important that all dimensions were represented in the final short form to ensure content coverage [12]. Further, it is recommended that item selection for short forms should be guided by both content of the items and item characteristics [9]. Therefore, our selection strategy was to choose one item per dimension, focusing on how well the content represented the dimension's core definition, but also considering the following item characteristics: (1) acceptance indicated by the percentage of missing responses [18], (2) relevance indicated by the percentage of 'does not concern me' replies, (3) item-difficulty indicated by the mean standardized to range from 0 to 1 (recommended range 0.2–0.8) [19], (4) corrected item-total correlation within each dimension (recommended range > 0.3) [20]. Additionally, in contrast to EPAT-64, we decided to create one version of EPAT-16 for both inpatient and outpatient settings. Hence, we excluded any items that applied to only one setting. This allowed us to create a short form that could be used across different healthcare contexts, providing comparability.

2.4.2 | Psychometric Analyses of Final EPAT-16

After item selection, we examined further psychometric properties of the EPAT-16 within the same data set. First, we examined structural validity using confirmatory factor analysis (CFA) [21–23] testing a unidimensional model. We used a robust maximum likelihood estimator (MLR). As recommended [24], we used several indices to assess model fit (see Table 2). We report standardized factor loadings and average variance explained by calculating the average percentage of the variance in the items which is explained by the common factor. Further, based on the factor analysis, we calculated McDonald's Omega to test reliability [25]. We examined measure invariance between medical conditions by imposing the same factor structure (configural invariance), same factor loadings (weak invariance) and same factor loadings and intercepts (strong invariance) for each of the four groups [26]. We compared the fit of these models using likelihood ratio tests.

Next, we examined construct validity. To assess convergent validity, we hypothesized that the EPAT-16 would correlate with the German ZUF-8, a unidimensional measure of patient satisfaction with high internal consistency (Cronbach's $\alpha = 0.9$) [16]. We expected the correlation to exceed 0.5 [27]. For divergent validity, we used the first item of the German SF-12, which assesses general health status, has an item difficulty of 0.4 and high acceptability [15]. Given that health status is theoretically unrelated to PC, we hypothesized a correlation magnitude below 0.3 [27]. Finally, we examined the distribution of the EPAT-16 sum score using its mean, standard deviation and percentage frequencies.

We used SPSS Version 27.0 (IBM Corp., Armonk, NY, USA) to enter data, calculate sample characteristics as well as sum scores and item total correlations of EPAT-16 and to test construct validity. All other data cleaning and analyses were performed using R Version 4.3.2 (R Core Team, Vienna, Austria). During data cleaning, we excluded patients who responded to fewer than 70% of the EPAT items (where 'does not concern me' was counted as a response) or who failed both control items. We treated 'does not concern me' responses as missing values in all

further analyses. For factor analyses, we used full information maximum likelihood (FIML) [28] to deal with missing values. For analysing construct validity, item-total-correlations and sum scores of EPAT-16, we used multiple imputation [29] with fully conditional specification [30] and predictive mean matching (40 imputations and maximum 100 iterations) to handle missing data [29]. We analysed data of inpatient and outpatient settings separately, with results reported for all medical conditions combined.

3 | Results

3.1 | Sample

Details of the sample were already reported in the primary publication on EPAT-64 [5]. The participant flow chart and the table with all sample characteristics are replicated in Appendix 2. The final analyses included $n = 1092$ outpatient and $n = 931$ inpatient questionnaires. Of the outpatients, 59.1% were female, 38.5% male and 0.5% identified as diverse. Their mean age was 52.8 years (standard deviation [SD] = 17.4, range 18–95). Regarding medical conditions, 25.4% were treated for cardiovascular diseases, 25.0% for cancer, 19.9% for musculoskeletal diseases and 25.0% for mental disorders. In the inpatient sample, 41.3% were female, 55.5% male and 0.6% identified as diverse. The mean age was 55.5 years (SD = 17.4, range 18–92). Of the inpatients, 30.7% were treated for cardiovascular diseases, 37.7% for cancer, 9.9% for musculoskeletal diseases and 21.7% for mental disorders.

3.2 | Item Characteristics

The items selected for EPAT-16 are shown in Table 1 with their corresponding dimension of the EPAT-64, their English translation and their item characteristics. None of the final EPAT-16 items had to be reversed.

Regarding relevance, in the outpatient sample three items had 'does not concern me' response rates above 30% (highest rate was for the 'physical support'-item with 54.7%). In the inpatient sample all items had 'does not concern me' response rates below 30% (highest rate was for the item 'involvement of family and friends' with 29.6%). Acceptance was good for all items (highest rate of missing responses was 1.4% for outpatients and 1.5% for inpatients). Regarding item difficulty, 12 and 11 items in the outpatient and inpatient sample, respectively, were within the recommended range. The remaining items showed ceiling effects. Item total correlations ranged from 0.403 to 0.748 in the outpatient sample and from 0.489 to 0.742 in the inpatient sample. Item characteristics per medical condition are shown in Appendix 3.

3.3 | Psychometric Analyses of Final EPAT-16

Standardized factor loadings are shown in Table 1. They ranged from 0.407 to 0.829 in the outpatient sample and from 0.444 to 0.803 in the inpatient sample. Except for SRMR, all model fit

TABLE 1 | Item characteristics of the EPAT-16.

Dimension of EPAT-64	Item	Outpatient sample						Inpatient sample					
		Does not concern me			Item difficulty			Does not concern me			Item difficulty		
		M	SD	No reply	Item difficulty	Total Corr	Std. factor loadings	M	SD	No reply	Item difficulty	Total Corr	Std. factor loadings
Essential characteristics of the clinician	The healthcare professionals were sensitive (e.g. they addressed my feelings, showed understanding, or empathized with my situation).	5.1	1.2	1.2%	0.82	0.714	0.816	5.1	1.2	0.8%	0.82	0.701	0.798
Clinician–patient relationship	I trusted my healthcare professionals.	5.2	1.1	1.8%	0.85	0.694	0.792	5.3	1.1	0.4%	0.86	0.676	0.778
Patient as a unique person	My wishes, needs and expectations were asked and taken into account in the treatment.	4.8	1.4	8.3%	0.75	0.748	0.829	4.8	1.3	5.5%	0.76	0.742	0.803
Biopsychosocial perspective	My entire personal life was taken into account during the treatment (e.g. job, family and friends, partnership and sexuality, culture and religion, age, or financial circumstances).	3.9	1.8	16.1%	0.57	0.646	0.617	3.8	1.8	14.2%	0.56	0.593	0.538
Clinician–patient communication	I was given enough time to describe my concerns and my situation (e.g. medical history or current symptoms)	5.3	1.1	2.6%	0.86	0.673	0.760	5.2	1.1	2.8%	0.85	0.690	0.775
Integration of medical and nonmedical care	I was asked if I use or would like to use additional services (e.g. support groups, counselling, health courses, complementary and alternative medicine, or spiritual support/pastoral care).	3.1	2.0	30.4%	0.42	0.493	0.439	3.4	2.0	25.8%	0.49	0.531	0.464
Teamwork and teambuilding	The processes within the team were well organized.	5.1	1.0	7.1%	0.83	0.403	0.407	5.1	1.1	1.2%	0.82	0.605	0.695

(Continues)

TABLE 1 | (Continued)

Dimension of EPAT-64	Item	Outpatient sample					Inpatient sample				
		Does not concern me			Item difficulty		Does not concern me			Item difficulty	
		M	SD	No reply	Item difficulty	Item Total Corr	M	SD	No reply	Item difficulty	Item Total Corr
Access to care	If I wanted to speak to a physician, they were easily accessible.	4.7	1.4	1.3%	0.73	0.496	4.8	1.2	1.0%	0.76	0.611
Coordination and continuity of care	It was discussed with me whether follow-up appointments would be useful (e.g. for aftercare or further treatment).	5.1	1.4	1.4%	0.82	0.479	5.0	1.4	1.0%	0.80	0.547
Patient safety	I was encouraged to speak up if I noticed inconsistencies in my treatment.	3.9	1.8	23.4%	0.58	0.614	4.1	1.7	14.2%	0.62	0.610
Patient information	I received information about my condition from my healthcare professionals (e.g. causes, symptoms, effects or course).	4.6	1.6	10.9%	0.71	0.630	4.6	1.5	6.4%	0.72	0.606
Patient involvement in care	I was an equal partner with my healthcare professionals (e.g. in making decisions or sharing information).	5.0	1.2	7.6%	0.79	0.665	4.8	1.3	4.4%	0.76	0.643
Involvement of family and friends	I was informed about the options for involving my family members in the treatment (e.g. accompanying to appointments, participating in conversations, or assisting with medication intake).	3.1	2.0	42.5%	0.41	0.553	3.3	1.9	29.6%	0.46	0.578
Patient empowerment	I was encouraged to improve my health by changing my behaviour (e.g. through diet, exercise, reducing tobacco or alcohol).	4.0	1.8	26.0%	0.61	0.592	4.0	1.7	25.8%	0.61	0.492

(Continues)

TABLE 1 | (Continued)

Dimension of EPAT-64	Item	Outpatient sample					Inpatient sample								
		M	SD	Does not concern me	No reply	Item difficulty	Item Total Corr	Std. factor loadings	M	SD	Does not concern me	No reply	Item difficulty	Item Total Corr	Std. factor loadings
Physical support	When I had pain, I was helped quickly.	4.8	1.4	54.7%	1.0%	0.76	0.551	0.624	5.4	1.0	19.4%	0.8%	0.87	0.492	0.599
Emotional support	The healthcare professionals addressed my fears and concerns (e.g. by showing understanding and providing encouragement).	4.4	1.7	19.2%	0.8%	0.68	0.731	0.771	4.4	1.6	15.9%	1.5%	0.69	0.688	0.688

Note: M = mean; SD = standard deviation; Item Total Corr = corrected item total correlation with the other EPAT-16 items (recommended range > 0.3); Std. factor loadings = standardized factor loadings in the unidimensional model; item difficulty: recommended range 0.2–0.8.

indices slightly differed from the recommended range (see Table 2). We further explored this misfit by calculating residual correlations between items and modification indices (see Appendix 4). The highest residual correlation was found between the items for ‘Integration of medical and nonmedical care’ and ‘Involvement of family and friends’ (0.249 for outpatients and 0.278 for inpatients). The average variance explained of the unidimensional model was 94.7% for the outpatient sample and 94.5% for the inpatient sample. McDonald’s Omega indicated a high reliability of the unidimensional model (0.908 for the outpatient version and 0.906 for the inpatient version). Regarding measurement invariance, the model fit indices exhibited only slight changes for configural invariance and weak invariance when compared to the CFA of the full data set. For strong invariance, all model fit indices worsened noticeably. Yet, likelihood ratio tests for model comparisons (configural invariance vs. weak invariance, and weak invariance vs. strong invariance) showed a significant decline of fit ($p < 0.001$) for all comparisons in both samples.

Regarding construct validity, all hypotheses were confirmed. In the outpatient sample, the EPAT-16 sum score had a pooled correlation of 0.757 with satisfaction with care and of −0.113 with general health status. In the inpatient sample, those correlations were 0.794 and −0.050, respectively.

EPAT-16 sum scores had mean = 72.6 (SD = 15.7) in the outpatient sample and mean = 73.6 (SD = 15.2) in the inpatient sample (maximum possible range is 16 [lowest degree of PC] to 96 [highest degree of PC]). Distribution of EPAT-16 sum scores is shown in Figure 2. Mean, SD and percentiles for the EPAT-16 sum scores for the overall samples and per medical condition can be found in Appendix 5.

4 | Discussion

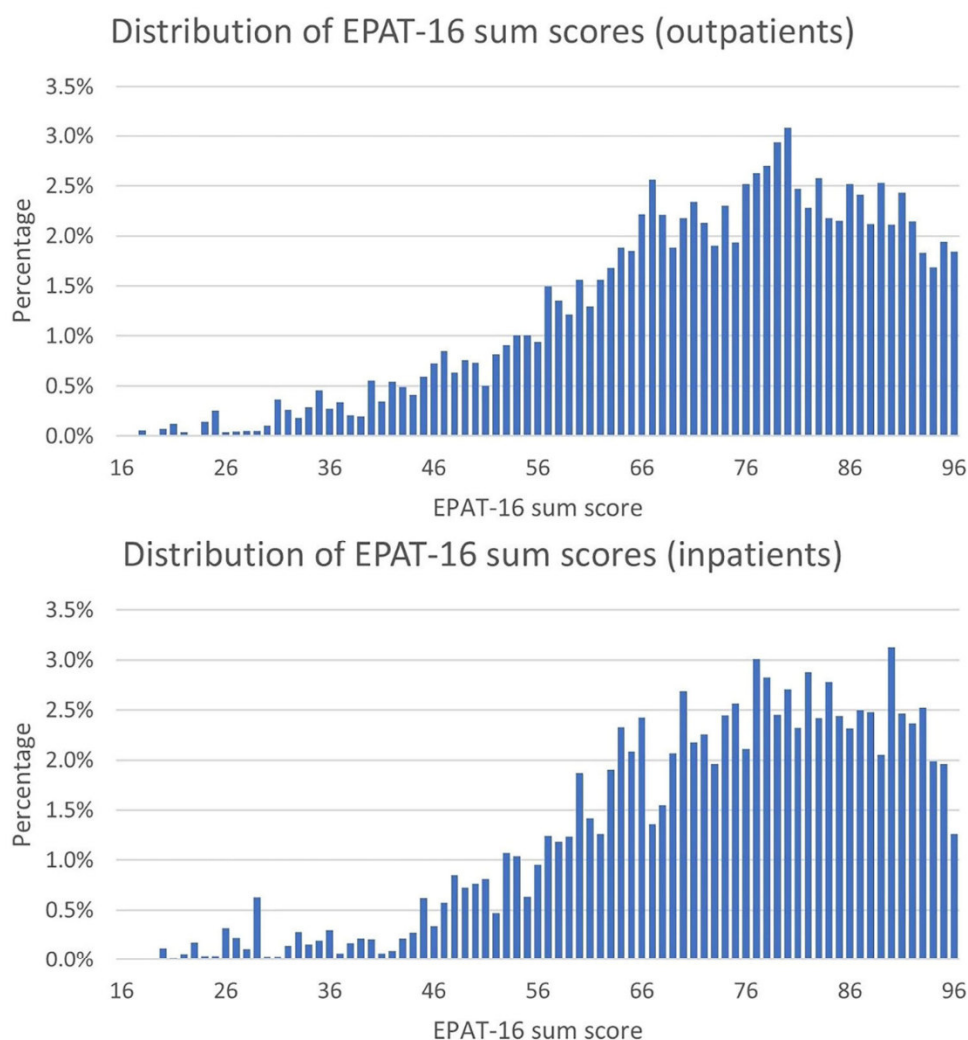
4.1 | Summary

We reported on the development and item selection for the EPAT-16, a short form designed to assess the overall degree of PC. We also evaluated the psychometric properties of the final EPAT-16 within the item selection sample. All items of the EPAT-16 demonstrated high acceptability and good corrected item-total correlations. While all items showed high relevance in the inpatient sample, three items were marked as ‘does not concern me’ by more than 30% of the outpatients. Approximately three-quarters of the items displayed appropriate difficulty levels, while the remaining items exhibited ceiling effects. The sum score showed a good degree of variance, although most scores were concentrated in the upper half of the possible range. Regarding structural validity, most model fit indices for the unidimensional model were slightly outside the recommended range. However, the model explained a substantial portion of the item variance, and McDonald’s Omega indicated high reliability. Even though the model fit indices changed only slightly when imposing the same factor loadings, based on the likelihood ratio tests the EPAT-16 did not achieve weak measurement invariance. Finally, all hypotheses regarding construct validity were confirmed.

TABLE 2 | Model fit, average variance explained, reliability and measure invariance of the unidimensional model.

	Chi square			RMSEA	SRMR	TLI	CFI	AIC	BIC
	Chi ²	df	<i>p</i> value						
Recommendation			> 0.05	≤ 0.06	< 0.08	≥ 0.90	≥ 0.90	Smallest value	
Outpatient sample	843	104	< 0.001	0.081	0.066	0.877	0.894	43,910	44,149
Configural invariance	1170	416	< 0.001	0.097	0.080	0.856	0.875	41,279	42,229
Weak invariance	1280	461	< 0.001	0.095	0.100	0.861	0.866	41,299	42,026
Strong invariance	1816	506	< 0.001	0.114	0.114	0.799	0.788	41,744	42,249
Inpatient sample	856	104	< 0.001	0.088	0.073	0.865	0.883	39,789	40,022
Configural invariance	1024	416	< 0.001	0.086	0.069	0.881	0.897	38,345	39,272
Weak invariance	1122	461	< 0.001	0.086	0.096	0.881	0.886	38,353	39,063
Strong invariance	1796	506	< 0.001	0.115	0.122	0.790	0.778	38,937	39,429

Note: Configural invariance = same factor structure imposed on all four groups; weak invariance = same factor loadings imposed on all four groups; strong invariance = same factor loadings and intercepts imposed on all four groups; df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; SRMR = standardized root mean square residual; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; AIC = Akaike information criterion; BIC = Bayesian information criterion.

**FIGURE 2** | Percentage frequencies of EPAT-16 sum scores.

The EPAT-16 offers an efficient and concise method for assessing PC while ensuring representation of all 16 dimensions in the integrative model of PC [6, 7]. This was achieved by an item selection strategy that, next to item characteristics, took

into account to which extent an item reflected the dimension definition. While the EPAT-64 allows to explore detailed questions about specific dimensions of PC, the EPAT-16 serves as a potentially valuable tool when the focus is on an overall PC

score and resources are constrained. Reducing the number of items to a quarter makes the EPAT-16 more feasible in routine care, where time and personnel are limited, and patients might have low cognitive abilities or energy due to their medical conditions and treatment. The EPAT-16 can be utilized in routine care for continuous feedback on the overall degree of PC or for quality improvement. It can be used in large-scale multivariate studies assessing the relationship of a large number of constructs and their relationship to PC. It also enables large-scale collection of patient feedback across healthcare institutions, facilitating comparisons and reporting of PC levels.

Although other short questionnaires assessing PC exist [31, 32], EPAT stands out due to its rigorous item development and selection process, which ensures a strong theoretical foundation and a comprehensive evaluation of PC [33], even with a reduced number of items.

4.2 | Limitations and Future Research

The results presented in this study are limited by the fact that all psychometric analyses were conducted in the same sample as had been used for item selection [12]. It is also the same sample which we used to develop and evaluate the long version, EPAT-64. To confirm that the psychometric properties hold true when participants respond to only the 16 items, the EPAT-16 should be administered to a new, independent sample in future studies. Future research should specifically ask patients to complete both the EPAT-16 and EPAT-64 separately, allowing for an accurate calculation of the correlation between the two questionnaires [12]. We did not report this correlation here, as using the same sample would have resulted in overestimation due to using responses on the EPAT-16 items on both sides of the calculation [12].

Although the items were selected based on both content and psychometric characteristics—a key strength of the EPAT-16—this process could have been more transparent by incorporating formal content rating and more representative by involving a larger group of experts in the selection process [12].

Furthermore, most model fit indices were slightly outside the recommended range, suggesting that exploring alternatives to the unidimensional model may be worthwhile. The residual correlations and modification indices presented in Appendix 4, in conjunction with theoretical considerations, can guide the development of such alternative models. An example for how such a model might be developed can be found in Appendix 4. Wording of the EPAT-16 items was designed to be applicable for all patients regardless of their medical condition. Yet, measurement invariance between medical conditions was not supported in the data analysed here. Hence, while it can be applied generically to different healthcare institutions, comparability of results between different medical conditions is limited. Appendix 5 provides results of EPAT-16 sum scores per setting and medical condition to allow users to compare results of similar samples. Future research should assess the psychometric properties of the EPAT-16, including its factor structure and how it changes, in patient groups with further medical conditions that were not included here. In general, before using the

EPAT-16 in different patient populations, its psychometric properties first should be tested in comparable samples. Adaptations of the EPAT-16 to specific populations might be required, in particular where items show high rates of ‘does not concern me’ responses, indicating limited relevance for the respective population. Furthermore, ceiling effects of some items reduce the ability of the EPAT to differentiate between healthcare facilities with high PC ratings. This might be addressed by changing the response scale (e.g. frequency-based scales rather than agreement-based scales) or incorporating additional discriminating items to enhance sensitivity.

In addition, the strong association of the sum score with the results from the Client Satisfaction Questionnaire deserves further attention, as the construct measured by the EPAT is likely to be a complex mixture of subjective and objective experiences [5]. This should be further explored in future studies. Finally, validating the English version of the EPAT-16, along with translating and testing it in other languages, would expand access to more patients and facilitate research on PC across different cultures, countries and healthcare systems.

5 | Conclusions

The EPAT-16 provides a brief tool for assessing PC while covering all 16 dimensions from the integrative model of PC. Its concise format makes it well-suited for routine care settings, where time, resources and patient capacity may be limited. It can be used for quality improvement, studies examining relationships of a large number of variates, and public reporting of PC across different healthcare institutions. Future research should focus on validating the EPAT-16 in independent samples, exploring its factor structure across various patient populations and expanding its use across different languages and healthcare systems. The EPAT-16 is licensed under a Creative Commons license and can be downloaded in German and English at www.uke.de/epat.

Author Contributions

Eva Christalle: investigation, formal analysis, methodology, resources, data curation, writing – original draft, writing – review and editing, visualization, project administration. **Levente Kriston:** conceptualization, methodology, formal analysis, investigation, writing – review and editing, supervision. **Stefan Zeh:** investigation, resources, writing – review and editing, project administration. **Hannah Führes:** project administration, resources, writing – review and editing. **Alica Schellhorn:** project administration, resources, writing – review and editing. **Pola Hahlweg:** conceptualization, investigation, writing – review and editing. **Jördis Maria Zill:** investigation, writing – review and editing. **Martin Härter:** conceptualization, investigation, resources, project administration. **Isabelle Scholl:** conceptualization, methodology, investigation, writing – review and editing, supervision, project administration, funding acquisition.

Acknowledgements

During the writing of the manuscript, Eva Christalle used ChatGPT (OpenAI, Version 2) for language editing to improve readability and conciseness, since she is not a native English speaker. All content was created by the authors, and only prompts to improve the language without altering the content were used. After using ChatGPT, all

authors reviewed and edited the content as needed and take full responsibility for the content of the publication. Open Access funding enabled and organized by Projekt DEAL.

Ethics Statement

The study was carried out according to the latest version of the Helsinki Declaration of the World Medical Association. Principles of good scientific practice were respected. The study had been approved by the Ethics Committee of the Medical Association Hamburg (Study ID: PV5724). Study participation was voluntary, and no foreseeable risks for participants resulted from the participation in this study. Participants were fully informed about the aims of the study, data collection and the use of collected data. Written informed consent was obtained before participation. Preserving principles of data sensitivity, data protection and confidentiality requirements were met.

Conflicts of Interest

Isabelle Scholl received honoraria for presentations and speeches on patient-centred care from the following commercial entities: onkowsissen.de GmbH, ClinSol GmbH & Co. KG. The authors declare no conflicts of interest.

Data Availability Statement

Data are available upon reasonable request.

References

1. K. Browne, D. Roseman, D. Shaller, and S. Edgman-Levitan, "Analysis & Commentary Measuring Patient Experience as a Strategy for Improving Primary Care," *Health Affairs* 29, no. 5 (2010): 921–925.
2. R. M. Epstein and R. L. Street, "The Values and Value of Patient-Centered Care," *Annals of Family Medicine* 9 (2011): 100–103.
3. M. Beattie, D. J. Murphy, I. Atherton, and W. Lauder, "Instruments to Measure Patient Experience of Healthcare Quality in Hospitals: A Systematic Review," *Systematic Reviews* 4, no. 1 (2015): 97.
4. C. Bull, J. Byrnes, R. Hettiarachchi, and M. Downes, "A Systematic Review of the Validity and Reliability of Patient-Reported Experience Measures," *Health Services Research* 54, no. 5 (2019): 1023–1035.
5. E. Christalle, S. Zeh, H. Führes, et al., "Through the Patients' Eyes: Psychometric Evaluation of the 64-item Version of the Experienced Patient-Centeredness Questionnaire (EPAT-64)," *BMJ Quality & Safety* (2024): bmjqs-2024-017434, ahead of print, October 16, 2024, <https://doi.org/10.1136/bmjqs-2024-017434>.
6. I. Scholl, J. M. Zill, M. Härter, and J. Dirmaier, "An Integrative Model of Patient-Centeredness – A Systematic Review and Concept Analysis," *PLoS One* 9, no. 9 (2014): e107828.
7. S. Zeh, E. Christalle, P. Hahlweg, M. Härter, and I. Scholl, "Assessing the Relevance and Implementation of Patient-Centredness From the Patients' Perspective in Germany: Results of a Delphi Study," *BMJ Open* 9, no. 12 (2019): e031741.
8. E. Christalle, S. Zeh, P. Hahlweg, et al., "Development and Content Validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—A Best Practice Example for Generating Patient-Reported Measures From Qualitative Data," *Health Expectations* 25 (2022): 1529–1538.
9. S. Schipolowski, U. Schroeders, and O. Wilhelm, "Pitfalls and Challenges in Constructing Short Forms of Cognitive Ability Measures," *Journal of Individual Differences* 35 (2014): 190–200.
10. C. J. Sandy, S. D. Gosling, and T. Koelkebeck, "Psychometric Comparison of Automated Versus Rational Methods of Scale Abbreviation," *Journal of Individual Differences* 35 (2014): 221–235.
11. M. Ziegler, C. J. Kemper, and P. Krüyen, *Short Scales – Five Misunderstandings and Ways to Overcome Them* (Hogrefe Publishing, 2014).
12. G. T. Smith, D. M. McCarthy, and K. G. Anderson, "On the Sins of Short-Form Development," *Psychological Assessment* 12, no. 1 (2000): 102–111.
13. E. Christalle, S. Zeh, P. Hahlweg, L. Kriston, M. Härter, and I. Scholl, "Assessment of Patient Centredness Through Patient-Reported Experience Measures (ASPIRED): Protocol of a Mixed-Methods Study," *BMJ Open* 8, no. 10 (2018): e025896.
14. J. Harkness, B. E. Pennell, and A. Schoua-Glusberg, "Survey Questionnaire Translation and Assessment," in *Methods for Testing and Evaluating Survey Questionnaires* (2004), 453–473, <https://doi.org/10.1002/0471654728.ch22>.
15. M. A. Wirtz, M. Morfeld, H. Glaesmer, and E. Brähler, "Konfirmatorische Prüfung der Skalenstruktur des SF-12 Version 2.0 in einer deutschen bevölkerungs-repräsentativen Stichprobe," *Diagnostica* 64, no. 2 (2018): 84–96.
16. D. Kriz, R. Nübling, A. Steffanowski, et al., "Patientenzufriedenheit in der stationären Rehabilitation: Psychometrische Reanalyse des ZUF-8 auf der Basis multizentrischer Stichproben verschiedener Indikation," *Zeitschrift für medizinische Psychologie* 17, no. 2–3 (2008): 67–79.
17. F. Röthlin, J. M. Pelikan, and K. Ganahl "Die Gesundheitskompetenz der 15-jährigen Jugendlichen in Österreich," Abschlussbericht der österreichischen Gesundheitskompetenz Jugendstudie im Auftrag des Hauptverbands der österreichischen Sozialversicherungsträger (HVS) 2013.
18. K. R. Sepucha, D. D. Matlock, C. E. Wills, et al., "It's Valid and Reliable' Is Not Enough: Critical Appraisal of Reporting of Measures in Trials Evaluating Patient Decision Aids," *Medical Decision Making* 34, no. 5 (2014): 560–566.
19. N. Döring and J. Bortz "Forschungsmethoden und Evaluation," (Springerverlag, 2006).
20. G. O. Boateng, T. B. Neilands, E. A. Frongillo, H. R. Melgar-Quinonez, and S. L. Young, "Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer," *Frontiers in Public Health* 6 (2018): 149.
21. D. B. Flora and J. K. Flake, "The Purpose and Practice of Exploratory and Confirmatory Factor Analysis in Psychological Research: Decisions for Scale Development and Validation," *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* 49, no. 2 (2017): 78–88.
22. D. D. Suhr, Exploratory or Confirmatory Factor Analysis?, 2006.
23. D. L. Bandalos and S. J. Finney, "Factor analysis: Exploratory and confirmatory," *Reviewer's Guide to Quantitative Methods in the Social Sciences* (Routledge, 2018), 98–122.
24. P. Irwing, T. Booth, and D. J. Hughes, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development* (John Wiley & Sons, 2018).
25. D. B. Flora, "Your Coefficient Alpha is Probably Wrong, but Which Coefficient Omega is Right? A Tutorial on Using R to Obtain Better Reliability Estimates," *Advances in Methods and Practices in Psychological Science* 3, no. 4 (2020): 484–501, <https://doi.org/10.1177/2515245920951747>.
26. The Lavaan Project. Measurement Invariance Testing, accessed 7 December 2024, <https://lavaan.ugent.be/tutorial/groups.html#measurement-invariance-testing>.
27. S. Olteidal, A. Garratt, Ø. Bjertnæs, M. Bjørnsdottir, M. Freil, and M. Sachs, "The NORPEQ Patient Experiences Questionnaire: Data Quality, Internal Consistency and Validity Following a Norwegian Inpatient Survey," *Scandinavian Journal of Public Health* 35, no. 5 (2007): 540–547.
28. C. K. Enders, "A Primer on Maximum Likelihood Algorithms Available for Use With Missing Data," *Structural Equation Modeling: A Multidisciplinary Journal* 8, no. 1 (2001): 128–141.

29. D. A. Newman, "Missing Data: Five Practical Guidelines," *Organizational Research Methods* 17, no. 4 (2014): 372–411.
30. S. Van Buuren, J. P. L. Brand, C. G. M. Groothuis-Oudshoorn, and D. B. Rubin, "Fully Conditional Specification in Multivariate Imputation," *Journal of Statistical Computation and Simulation* 76, no. 12 (2006): 1049–1064.
31. C. Gehrlach, T. Altenhöner, and D. Schwappach "Der Patients' Experience Questionnaire," Patientenerfahrungen Vergleichbar Machen Bertelsmann Stiftung, Gütersloh 2009.
32. C. Jenkinson, "The Picker Patient Experience Questionnaire: Development and Validation Using Data From in-Patient Surveys in Five Countries," *International Journal for Quality in Health Care* 14, no. 5 (2002): 353–358.
33. A. L. Mihaljevic, C. Doerr-Harim, E. Kalkum, and G. Strunk, "Measuring Patient Centeredness With German Language Patient-Reported Experience Measures (PREM)—A Systematic Review and Qualitative Analysis According to COSMIN," *PLoS One* 17, no. 11 (2022): e0264045, <https://doi.org/10.1371/journal.pone.0264045>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

10. OWN CONTRIBUTIONS TO THE PUBLICATIONS

Publication 1: Christalle, E.*, Zeh, S.*, Hahlweg, P., Kriston, L., Härter, M., & Scholl, I. (2018).

Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ open*, 8(10), e025896.

(*Shared first authorship)

Own contributions: Based on the grant proposal of the ASPIRED study, EC planned the development process and psychometric evaluation in detail, including research to identify hypotheses for testing construct validity. Furthermore, EC wrote the manuscript as shared first author, focusing particularly on the sections about item development and psychometric evaluation.

Publication 2: Zeh, S., **Christalle, E.**, Hahlweg, P., Härter, M., & Scholl, I. (2019). Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ open*, 9(12), e031741.

Own contributions: EC made substantial contributions to the study design and conducted the data collection. She prepared the data, analyzed the data together with the first author and finalized the resulting model with the first and the senior author. Further, she provided critical feedback to the manuscript.

Publication 3: Christalle, E., Zeh, S., Hahlweg, P., Kriston, L., Härter, M., Zill, J., & Scholl, I. (2022). Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)—A best practice example for generating patient-reported measures from qualitative data. *Health Expectations*, 25(4), 1529-1538.

Own contributions: EC conducted the literature search, contributed substantially to preparation of data collections and analyzed all data independently as one of two raters and compiled the item pool together with a co-author. EC drafted the manuscript and revised it several times after receiving feedback from co-authors and peer reviewers.

Publication 4: Christalle, E., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J. M., Härter M., Bokemeyer C., Gallinat J., Gebhardt C., Magnussen C., Müller V., Schmalstieg-Bahr K., Strahl A., Kriston L., Scholl, I. (2024). Through the patients' eyes: psychometric evaluation of the 64-item version of the Experienced Patient-Centeredness Questionnaire (EPAT-64). *BMJ Qual Saf.* 2024 Oct 16:bmjqs-2024-017434.

Own contributions: EC drafted the analysis plan for the psychometric evaluation and revised it after team discussion with the senior authors and a co-author. She autonomously cleaned and analyzed the data with regular feedback from the two senior authors, prepared the item selection session and moderated it. EC drafted the manuscript and revised it several times after receiving feedback from co-authors and peer reviewers.

Publication 5: Christalle, E., Kriston, L., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J.M., Härter, M. & Scholl, I. (2025). Development and Psychometric Testing of EPAT-16: A Short and Valid Measure for Patient-Centeredness From the Patient's Perspective. *Health Expectations*, 28(3), e70296.

Own contributions: EC drafted the analysis plan for the psychometric evaluation and revised it after team discussion with the senior author and two co-authors. She autonomously cleaned and analyzed the data with regular feedback from a co-author and the senior author prepared the item selection session and moderated it. EC drafted the manuscript and revised it several times after receiving feedback from co-authors and peer reviewers.

11. DANKSAGUNG

Als erstes möchte ich meiner Betreuerin und langjährigen Forschungsgruppenleiterin Isabelle Scholl danken. Mit deiner Anleitung, deiner Unterstützung und deinem Verständnis konnte ich über all die Jahre in meiner Rolle wachsen und mein Potential als Wissenschaftlerin entfalten. Ich bin unglaublich stolz auf das, was ich mit deinem Mentoring erreicht habe! Ich danke Levente Kriston für viele, stets bereichernde Diskussionen, für dein frühes Vertrauen in mein Potential und deine Unterstützung als Teil meines Thesis-Komitees. Ich danke Herrn Prof. Dr. Scherer für Ihre Rückmeldungen und Ihre Unterstützung im Rahmen des Thesis-Komitees. Ich danke meinen Kolleg:innen am Institut für Medizinische Psychologie und ganz besonders in der PACE AG. Ihr habt über Jahre dafür gesorgt, dass ich mich im Institut wohlfühlt habe und ich mich stets auf den Austausch mit euch gefreut habe. Ganz besonders möchte ich außerdem meinem „Doktorbruder“ Stefan Zeh danken, dafür dass du mit mir durch all die Höhen und Tiefen unserer gemeinsamen Karriere gegangen bist, und immer wertschätzend und aufmerksam zugehört hast.

Ich danke meiner Familie. Als allererstes möchte ich meinem absolut wunderbaren Ehemann Thorsten danken, dafür dass du mich immer unterstützt das zu tun, was mir wichtig ist, einfach nur weil es mir wichtig ist, dass du mir dabei immer den Rücken freigehalten hast, damit ich meine Projekte veröffentlichen kann, für deine unglaubliche Geduld und bedingungslose Liebe. Ich danke meinem geliebten Sohn Leon. Es ist eine reine Freude, wie du selbst als kleiner Wissenschaftler voller Neugier und Begeisterung die Welt entdeckst. Danke, dass du mir zeigst, was wirklich wichtig ist, wie wunderbar das Leben ist und wie groß Liebe sein kann. Ich danke meinen Eltern, Biggi und Joppich für all eure Liebe. Ihr seid das Fundament, auf dem ich mein Leben aufbauen konnte – und ich bin unendlich dankbar für dieses Leben! Ich danke meiner Schwester Lena, dafür, dass du dieser wunderbare, einfühlsame Mensch bist, immer interessiert und bereit wirklich zuzuhören. Ich danke meiner Seelenverwandten Nastja für eine Freundschaft über Grenzen hinaus. Danke, für all deine Unterstützung, deine Geduld, deine Ermutigungen und all die wunderbaren gemeinsamen Erinnerungen. Ich wünsche dir ganz viel Kraft für deine Dissertation! Ich danke Maren dafür, dass du immer ein offenes Ohr für mich hast, selbst wenn wie uns in stressigen Phasen nicht persönlich treffen. Danke für endlose Stunden geduldigen Zuhörens und einfühlsamer Sprachnachrichten. I would like to thank my dear friends, Ashton and Rose. Thank you, Ashton, for providing the soundtrack that kept me sane no matter what, and for involving me in all those exciting endeavors that truly fueled me during my final PhD year. Thank you, Rose, for being so inspirational and for encouraging me to pursue dreams I never even thought possible. Gracias, Angie, por una noche genial y los tacos más ricos que he probado en mi vida. ¡Que disfrutes leyendo mi tesis!

12. CURRICULUM VITAE

Persönliche Daten

Name	Eva Christalle
Geburtsdatum / -ort	27.12.1988 in Berlin
Staatsangehörigkeit	deutsch
Familienstand	verheiratet, ein Kind

Bildung

07.2024	B. Sc. of Mathematics and Statistics (Open University, UK, Gesamtnote Pass Grade 2)
04.2016	M. Sc. Psychologie (Universität Hamburg, Gesamtnote 1,6)
01.2013	B. Sc. Psychologie (Universität Bremen, Gesamtnote 1,3)
06.2008	Abitur (Heinrich-Hertz-Gymnasium Berlin, Gesamtnote 1,9)

Beruflicher Werdegang (Auszug)

Seit 2016	Wissenschaftliche Mitarbeiterin (Institut und Poliklinik für Medizinische Psychologie, UKE)
2013 - 2016	Studentische Hilfskraft (Institut und Poliklinik für Medizinische Psychologie, UKE)
2012 - 2013	Werkstudentin (Deutsches Zentrum für Luft- und Raumfahrt (DLR), Hamburg)
2010 - 2012	Statistik-Tutorin (Universität Bremen, Bremen)

Bisherige Forschungsprojekte am Institut für Medizinische Psychologie

Seit 10.2025	Physical and psychological symptoms in patients with cancer: meta-analytic and longitudinal network analysis (OncoSympNet) (Hauptaufgabe: Statistische Analysen und methodische Beratung)
Seit 07.2024	Wirksamkeitsstudie digitale Intervention „MyMinddistrict“ (Hauptaufgabe: Statistische Analysen und methodische Beratung)
10.2023 - 06.2024	Kommunikation in der patient:innenorientierten Pflege (KOMPAT) (Hauptaufgabe: Datenanalyse der RCT-Ergebnisse)
Seit 07.2023	Patient Safety in German Radiation Oncology (PaSaGeRO) (Hauptaufgabe: Methodische Planung, Studiendurchführung im Team, alle Datenanalysen)
10.2017 - 09.2023	Assessment of patient centredness through patient-reported experience measures (ASPIRED) (Promotionsprojekt) (Hauptaufgabe: Durchführung aller Studienschritte im Team, Fokus auf psychometrische Evaluation)
10.2016 - 03.2017	Expertise zur Versorgung traumatisierter Kinder & Jugendlicher in Deutschland (Hauptaufgabe: Analyse von Befragungsdaten und Verfassen des Ergebnisberichtes)
04.2016 - 09.2017	Psychological distress of patients with psoriasis (Hauptaufgabe: Systematisches Review und Meta-Analyse)
04.2016 - 09.2017	Wirksamkeitsstudie digitale Intervention „Vorvida“ zur Reduktion von Alkoholkonsum (Hauptaufgabe: Analyse und Publikation der RCT-Ergebnisse)

Peer-reviewed Publikationen

- **Christalle, E.**, Kriston, L., Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J. M., Härter, M., & Scholl, I. (2025). Development and psychometric testing of EPAT-16: A short and valid measure for patient-centeredness from the patient's perspective. *Health Expect.* <https://doi.org/10.1111/hex.70296>
- Baehr, A., Grohmann, M., **Christalle, E.**, Schwenzer, F., & Scholl, I. (2025). Aiming for patient safety indicators in radiation oncology - Results from a systematic literature review as part of the PaSaGeRO study. *Radiother Oncol*, 203, 110657. <https://doi.org/10.1016/j.radonc.2024.110657>
- **Christalle, E.**, Zeh, S., Führes, H., Schellhorn, A., Hahlweg, P., Zill, J. M., Härter, M., Bokemeyer, C., Gallinat, J., Gebhardt, C., Magnussen, C., Müller, V., Schmalstieg-Bahr, K., Strahl, A., Kriston, L., & Scholl, I. (2024). Through the patients' eyes: psychometric evaluation of the 64-item version of the Experienced Patient-Centeredness Questionnaire (EPAT-64). *BMJ Qual Saf.* <https://doi.org/10.1136/bmjqs-2024-017434>
- Baehr, A., **Christalle, E.**, Grohmann, M., & Scholl, I. (2024). Development and psychometric validation of a patient safety assessment tool in German Radiation Oncology: the PaSaGeRO Study protocol. *BMJ Open*, 14(8), e086214. <https://doi.org/10.1136/bmjopen-2024-086214>
- **Christalle, E.**, Zeh, S., Hahlweg, P., Kriston, L., Härter, M., Zill, J., & Scholl, I. (2022). Development and content validity of the Experienced Patient-Centeredness Questionnaire (EPAT)-A best practice example for generating patient-reported measures from qualitative data. *Health Expect.*, 25(4), 1529-1538. <https://doi.org/10.1111/hex.13494>
- Zeh, S., **Christalle, E.**, Zill, J. M., Härter, M., Block, A., & Scholl, I. (2021). What do patients expect? Assessing patient-centredness from the patients' perspective: an interview study. *BMJ Open*, 11(7), e047810. <https://doi.org/10.1136/bmjopen-2020-047810>
- Lindig, A., Hahlweg, P., **Christalle, E.**, Giguere, A., Härter, M., von dem Knesebeck, O., & Scholl, I. (2021). Translation and psychometric evaluation of the German version of the IcanSDM measure - a cross-sectional study among healthcare professionals. *BMC Health Serv Res*, 21(1), 541. <https://doi.org/10.1186/s12913-021-06430-3>
- Nordholt, P. U., **Christalle, E.**, Zill, J. M., & Dirmaier, J. (2020). Engagement With a Web-Based Intervention to Reduce Harmful Drinking: Secondary Analysis of a Randomized Controlled Trial. *J Med Internet Res*, 22(11), e18826. <https://doi.org/10.2196/18826>
- Lindig, A., Hahlweg, P., **Christalle, E.**, & Scholl, I. (2020). Translation and psychometric evaluation of the German version of the Organisational Readiness for Implementing Change measure (ORIC): a cross-sectional study. *BMJ Open*, 10(6), e034380. <https://doi.org/10.1136/bmjopen-2019-034380>
- Zill, J. M., **Christalle, E.**, Tillenburg, N., Mrowietz, U., Augustin, M., Härter, M., & Dirmaier, J. (2019). Effects of psychosocial interventions on patient-reported outcomes in patients with psoriasis: a systematic review and meta-analysis. *Br J Dermatol*, 181(5), 939-945. <https://doi.org/10.1111/bjd.17272>
- Zill, J. M., **Christalle, E.**, Meyer, B., Härter, M., & Dirmaier, J. (2019). The Effectiveness of an Internet Intervention Aimed at Reducing Alcohol Consumption in Adults. *Dtsch Arztebl Int*, 116(8), 127-133. <https://doi.org/10.3238/arztebl.2019.0127>
- Zeh, S., **Christalle, E.**, Hahlweg, P., Härter, M., & Scholl, I. (2019). Assessing the relevance and implementation of patient-centredness from the patients' perspective in Germany: results of a Delphi study. *BMJ Open*, 9(12), e031741. <https://doi.org/10.1136/bmjopen-2019-031741>
- **Christalle, E.**, Zill, J. M., Frerichs, W., Härter, M., Nestoriuc, Y., Dirmaier, J., & Scholl, I. (2019). Assessment of patient information needs: A systematic review of measures. *PLoS One*, 14(1), e0209165. <https://doi.org/10.1371/journal.pone.0209165>
- Zill, J. M., Dirmaier, J., Augustin, M., Dwinger, S., **Christalle, E.**, Härter, M., & Mrowietz, U. (2018). Psychosocial Distress of Patients with Psoriasis: Protocol for an Assessment of Care Needs and the Development of a Supportive Intervention. *JMIR Res Protoc*, 7(2), e22. <https://doi.org/10.2196/resprot.8490>

- **Christalle, E.,** Zeh, S., Hahlweg, P., Kriston, L., Härter, M., & Scholl, I. (2018). Assessment of patient centredness through patient-reported experience measures (ASPIRED): protocol of a mixed-methods study. *BMJ Open*, 8(10), e025896. <https://doi.org/10.1136/bmjopen-2018-025896>
- Kivelitz, L., Kriston, L., **Christalle, E.,** Schulz, H., Watzke, B., Härter, M., Götzmann, L., Bailer, H., Zahn, S., & Melchior, H. (2017). Effectiveness of telephone-based aftercare case management for adult patients with unipolar depression compared to usual care: A randomized controlled trial. *PLoS One*, 12(10), e0186967. <https://doi.org/10.1371/journal.pone.0186967>
- Doherr, H., **Christalle, E.,** Kriston, L., Härter, M., & Scholl, I. (2017). Use of the 9-item Shared Decision Making Questionnaire (SDM-Q-9 and SDM-Q-Doc) in intervention studies-A systematic review. *PLoS One*, 12(3), e0173904. <https://doi.org/10.1371/journal.pone.0173904>
- Cömert, M., Zill, J. M., **Christalle, E.,** Dirmaier, J., Härter, M., & Scholl, I. (2016). Assessing Communication Skills of Medical Students in Objective Structured Clinical Examinations (OSCE)--A Systematic Review of Rating Scales. *PLoS One*, 11(3), e0152717. <https://doi.org/10.1371/journal.pone.0152717>
- Zill, J. M., **Christalle, E.,** Müller, E., Härter, M., Dirmaier, J., & Scholl, I. (2014). Measurement of physician-patient communication--a systematic review. *PLoS One*, 9(12), e112637. <https://doi.org/10.1371/journal.pone.0112637>

Beiträge auf Konferenzen (Auszug)

- 2025 Poster: „Patient:innensicherheit in der Strahlentherapie messen - Entwicklung eines Fragebogens im mehrstufigen Verfahren“
Christalle E, Grohmann M, Schwenzer F, Baehr A, Scholl I
24. Deutscher Kongress für Versorgungsforschung (DKVF), Hamburg
- 2021 Vortrag: “Assessment of patient-centeredness through patient-reported experience measures (ASPIRED)”
Christalle E, Hahlweg P, Scholl I, Zeh S, Zill J
22. Jahrestagung EbM-Netzwerk, online
- 2019 Vortrag: “Psychometric testing of a core set of patient-reported experience measures (PREMs) to assess patient-centeredness from the patients’ perspective (research in progress)”
Christalle E, Zeh S, Scholl I
International Conference on Communication in Healthcare (ICCH), San Diego, USA
- 2019 Poster: “Development of a core set of patient-reported experience measures (PREMs) to assess patient-centeredness from the patients’ perspective”
Christalle E, Zeh S, Hahlweg P, Scholl I
International Conference on Communication in Healthcare (ICCH), San Diego, USA
- 2018 Vortrag: “Assessment of patient information needs: A systematic review of measures”
Christalle E, Zill J, Härter M, Dirmaier J, Scholl I
International Conference on Communication in Healthcare (ICCH), Porto, Portugal
- 2018 Poster: “Assessment of Patient-Centeredness through Patient-Reported Experience Measures – A Study Protocol”
Christalle E, Zeh S, Hahlweg P, Scholl I
International Conference on Communication in Healthcare (ICCH), Porto, Portugal
- 2018 Vortrag: “Effekte psychosozialer Interventionen für Psoriasis-Patienten auf Lebensqualität und psychologische Outcomes - eine Meta-Analyse”
Christalle E, Zill J, Härter M, Mrowietz U, Augustin M, Dirmaier J
19. Jahrestagung EbM-Netzwerk, Graz, Österreich
- 2017 Vortrag: “Erfassung von Informationsbedürfnissen von Patienten – eine systematische Übersichtsarbeit zu Messinstrumenten”
Christalle E, Zill J, Härter M, Dirmaier J, Scholl I
18. Jahrestagung EbM-Netzwerk, Hamburg, Deutschland

13. EIDESSTATTLICHE ERKLÄRUNG

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Während des Schreibens des Manuskripts habe ich ChatGPT (OpenAI, GPT-4o) zur sprachlichen Überarbeitung verwendet, um die Lesbarkeit und Prägnanz zu verbessern. Sämtliche Inhalte wurden von mir selbst erstellt, und ich habe ausschließlich Prompts verwendet, um die Sprache zu verbessern, ohne den Inhalt zu verändern. Nach der Nutzung von ChatGPT habe ich die Vorschläge sorgfältig überprüft und bei Bedarf bearbeitet und übernehme die volle Verantwortung für den Inhalt der Dissertation. Es wurden keine Inhalte durch Darüber hinaus habe ich die „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“ vom September 2023 berücksichtigt.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: