Measuring capabilities in health and physical activity promotion: a systematic review.

Maike Till (maike.till@fau.de)
Friedrich-Alexander-Universitat Erlangen-Nurnberg Philosophische Fakultat und Fachbereich Theologie
https://orcid.org/0000-0002-9915-7512

Karim Abu-Omar
Friedrich-Alexander-Universitat Erlangen-Nurnberg Philosophische Fakultat und Fachbereich Theologie

Susanne Ferschl
Friedrich-Alexander-Universitat Erlangen-Nurnberg Philosophische Fakultat und Fachbereich Theologie

Anne Kerstin Reimers
Friedrich-Alexander-Universitat Erlangen-Nurnberg Philosophische Fakultat und Fachbereich Theologie

Peter Gelius
Friedrich-Alexander-Universitat Erlangen-Nurnberg Philosophische Fakultat und Fachbereich Theologie

Research article

Keywords: Capability Approach, Public Health, Measurement, Operationalization, Physical activity, Health Outcomes, Questionnaires

Posted Date: July 29th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-42211/v1

Version of Record: A version of this preprint was published on February 15th, 2021. See the published version at https://doi.org/10.1186/s12889-020-10151-3.

Abstract

Background: The capability approach by Amartya Sen and Martha Nussbaum has gained increasing attention in the field of public health. As it combines individual, social and structural factors and shifts the focus of attention from the actual behavior towards available options for health behaviors that people can actually choose from, it may help advance our understanding of complex health issues.

Objectives: The aim of this article is to identify and describe tools available to measure capabilities within the context of health, with a specific focus on capabilities for health-enhancing physical activity.

Method: We conducted a systematic literature review using 11 databases covering scientific journal articles published in English or German between the years 2000 and 2019 with a focus on capabilities for health or physical activity.

Results: We found a total of 44 articles meeting our inclusion criteria. Four articles measured capabilities using qualitative methods, one combined qualitative and quantitative methods, while the rest used quantitative methods. We identified a total 11 different capability questionnaires, all showing moderate to good validity/reliability. Only one questionnaire and one interview-based tool specifically dealt with capabilities for health enhancing physical activity.

Conclusion: Although we were able to identify measurement tools for capabilities in health, this review has shown that there is no generic tool available for the measurement across all population- and age-groups, and tools focusing on physical activity are scarce. However, our results can be used as guide for future projects that aim at measuring capabilities.

Background

Over the last years, the capability approach – originally developed by Amartya Sen [1] and Martha Nussbaum [24] in the field of welfare economics – has gained increasing attention in the field of health and has been used in multiple health promotion projects [2-5]. A recent review by Helter et al. [6] highlights the growing relevance of the capability approach in health promotion, particularly for the health economic evaluation of projects. The approach shifts the focus of attention from an individual's actual behavior (their "achieved functionings", e.g. healthy eating patterns) towards the real opportunities ("capabilities") available to individuals to choose from. It builds on the assumption that the mere presence of options and opportunities positively influences well-being, and that increasing disadvantaged people's capabilities can help reduce social inequality.

The shift from behavior towards opportunities can be particularly beneficial in the field of health promotion. For instance, physical activity (PA) has been generally proven to have a positive impact on people's health, e.g. in relation to obesity, non-communicable diseases (e.g. diabetes, high-blood pressure), cardio-respiratory health, cancer, mental health and all-cause mortality [7, 8]. Current efforts to promote PA, however, still tend to focus on "downstream" interventions (e.g. physical education in school or structured PA classes for older people) that promise to have immediate effects on the target group's health behavior

(e.g. by increasing the number of daily steps or reducing the time spent sitting) [9]. However, such interventions may be less sustainable than more “upstream” interventions whose effects cannot immediately be measured in terms of target group behavior change (e.g. interventions to increase individuals’ physical literacy or to initiate infrastructure change) [10].

The capability approach may help shift the focus of attention to these approaches by pointing to their benefits in terms of individual capabilities for PA.

In addition, it explicitly respects people’s freedom to decide against a healthy lifestyle rather than “forcing” them to behave in a healthy manner (e.g. mandatory through mandatory physical education in schools), thus potentially enhancing target group compliance.

In general, a person’s capabilities for health enhancing behavior can be assumed to be based on a set of capitals or resources [3] that are “translated” into capabilities through three sets of *conversion factors* [11]: (1) individual (e.g. physical condition, biological health or health literacy), (2) social (e.g. norms and values, social practices or political rules), and (3) environmental factors (e.g. climate, pollution, infrastructure). However, to operationalize a concept as complex as the capability approach [4] (or, to give another example, Antonovsky’s [12] “sense of coherence”) for actual measurement is challenging, as these concepts are theoretical in nature and were not designed with a mind to empirical application. Nonetheless, the increasing popularity of the capability approach in health and PA promotion obliges us to assess not only health status and indicators of behavior but also the opportunities that people have to engage in such behavior.

The aim of this paper is to support researchers and health promoters who intend to use the capability approach by (1) systematically identifying all currently available tools to measure capabilities for health, (2) compare the main features of these tools as well as their psychometric properties, and applicability to different areas, and (3) specifically identify whether capabilities for health-enhancing PA are measured by any of the identified tools.

Methods

Research for this paper was conducted in the context of Capital4Health, a research consortium funded by the German Federal Ministry of Education and Research [01EL1421 A-F] which aimed at promoting active lifestyles in four different settings across the life-course using the capability approach. A project (CAPCOM, [01EL1421A.]) tasked with fostering cooperation in the consortium conducted the systematic review at hand in order to strengthen its common methodological base. The presented work followed the Preferred Reporting Items for Systematic-Reviews and Meta-Analyses (PRISMA) guidelines [13].

An initial exploratory search for instruments to measure capabilities specifically for PA indicated that only a limited number of instruments were dedicated to this topic, we therefore decided to broaden the search to include capability measurement tools for health in general.

Supported by a university librarian, three members of the CAPCOM research team developed a set of search strings consisting of variations of the terms “capability approach”, “measurement”, “health” and “physical activity” combined with Boolean operators. On 26th of June 2019, searches were conducted on the following databases: PsycINFO, Psychology and Behavioral Sciences Collection, SPORTDiscus, and PsycARTICLES via EBSCOhost, Applied Social Science Index & Abstracts, Sociological Abstracts, Social Services Abstracts, Worldwide Political Science Abstracts, International Bibliography of the Social Science, and the Sports Medicine & Education Index via ProQuest, and Pubmed.

Table 1 summarizes the inclusion/exclusion criteria applied to the results. Articles were included if they (a) were published between January 2000 and June 2019; (b) were written in English or German; (c) were scientific journal articles; (d) had a clear focus on the operationalization of the capability approach within the context of health or health-enhancing PA; and referred to any (e) population, (f) setting, or (g) country.

Two researchers independently screened all titles/abstracts based on the inclusion/exclusion criteria and discussed their results to resolve disagreement. Two researchers then independently screened the full texts of all remaining papers and discussed their results to reach consensus on the articles to be included for detailed analysis. In addition, the lead author carried out a supplementary hand search, the results of which were double-checked by another researcher. The included final search results were imported into Endnote X9 and analyzed regarding (i) the proposed types of measurement instruments for capabilities, (ii) the development process employed to develop these instruments, and (iii) the empirically tested validity, reliability, and responsiveness of the instruments among different target groups.

In the context of this paper, we rated instrument quality as follows: construct validity was categorized as “good” when correlations with other instruments had shown to be at least moderate and significant, or when chi-square analysis had shown to be significant at the 5% level. Discriminant validity was rated as “good” when the instrument showed a significant (at least $p < .01$) distinction between different areas. Internal consistency with $\alpha > .7$ was considered “good”, as well as test-retest reliability with a moderate Cohen’s kappa of at least (>.41) or an intraclass-correlation coefficient over .75.

Results

The search yielded in a total of N=9,444 hits matching the search terms across all eleven databases. After removing all duplicates, a total of 6,850 individual articles remained for screening. Researchers had substantial agreement on title/abstract screening (Cohen’s k=0.64), which yielded a total of 88 articles considered eligible for full-text screening. Researchers had moderate agreement in full-text screening (Cohen’s k=0.52), leading to the exclusion of another 49 articles. Five additional articles were identified in the hand search, resulting in a total of N=44 articles included in this review. A visual representation of the search is shown in figure 1 using the PRISMA-flowchart [13]

Types of measurement instruments

Table 2 provides an overview of the different measurement tools reported in the 49 identified articles. We found that instruments to assess capabilities fall into three major categories: (1) qualitative tools, e.g. using interviews or videography (n=4), (2) quantitative tools, e.g. questionnaires (n=39), and (3) mixed method approaches using a combination of interviews and questionnaires (n=1).

In the quantitative category, n=4 articles measured capabilities through analyzing secondary data (e.g. data from the British Panel Household Survey [14]), while n=35 papers covered a total of eleven individual questionnaires. Of these, four belong to the *ICECAP*-family (ICEpop CAPability index of the "Investigating Choice Experiments for the Preferences of Older People" (ICEpop) project) and use varying sets of items to cover specific target groups and outcome variables: the *ICECAP-O* for older adults [15] and *ICECAP-A* for adults [16] with 5 items each, the *ICECAP-SCM* measuring capabilities of people in need of supportive care [17] containing seven items, and the *ICECAP-FC* for adults measuring both functioning and capabilities [18] with ten items. Another set of questionnaires comes from the "Oxford Capability Questionnaire" family, including the original *OCAP* (Oxford capability Questionnaire) [19] with 64 items, the shortened *OCAP-18* [20] (18 items), and a version adapted to mental health, the *OxCAP-MH* (Oxford Capability Questionnaire for Mental Health) [21] (16 items). The most comprehensive questionnaires are the *CQ-CMH* (Capability Questionnaire for community mental health) [22] with 104 items and its adapted version, the *ACQ-CMH* (Achieved Capability Questionnaire for community mental health) [23] with 98 items. The systematic search further identified two questionnaires that did not belong to a larger "family" of tools, the *Capability Based Questionnaire for Patients with Chronic Pain* [24] (8 items) and the *Capability Assessment for Diet and Activity* (CADA) geared at adults suffering from obesity and diabetes [25]. All identified questionnaires use subjective/self-assessment measures to assess capabilities.

Main aims and methods employed

Table 3 reports on the main aims of the included articles as well as on the main methods used to develop the individual measurement tool and empirically test its measurement properties. Out of the 44 included articles, 8 described the development of a measurement instrument, 18 validated existing tools, 1 evaluated different instruments comparatively, and 7 reported results of actual measurements of health-related capabilities. Ten articles had a mixed focus on development/measurement (n=1), development/validation (n=6), or validation/comparison (n=3).

Among the qualitative tools, only Sauter et al. [26] provided details on the development process: Their interview guidelines were the result of literature screening and a conscious selection of specific items from the *OCAP* questionnaire [19]. The identified questionnaires were developed using different methodologies. For example, the OPCAP [19] is based on a set of largely theoretical criteria by Martha Nussbaum, who co-developed the original capability approach [27]. The ICECAP questionnaires [15-18], the *Capability Based Questionnaire for Patients with Chronic Pain* [24], and CADA [25] were compiled using expert opinion (e.g.

in the form of literature reviews) and target group interviews. The CQ-CMH [22] emanated from the analysis of focus group data, expert opinion, and an additional alignment with the Nussbaum criteria.

Articles reporting on the validation of questionnaires used different methodological approaches. Convergent and construct validity were mostly investigated by correlating results with those measured via other questionnaires (e.g. EQ-5D) [20, 21, 23, 28-37] or using Chi-Square analysis [15, 28, 38, 39]. Discriminant validity was ascertained by performing uni- or multivariate analysis [29-32, 40]. Some questionnaires were further been checked regarding their reliability using test-retest analysis [23, 33, 34, 41, 42], or regarding their responsiveness via anchor-based analysis [34, 37]. Moreover, the ICECAP-O and the Ox-CAP questionnaire were evaluated comparatively to the EQ-ED questionnaire by correlating their results [34, 43, 44].

Articles reporting on studies that directly measured capabilities without developing or validating any tools for future use were only found among the qualitative studies and secondary data analyses. Qualitative measurement was performed either by semi-structured interviews [45-47], observation [46] or video analysis [47], while secondary data was analyzed via methods such as regression [14] or equation modelling [48].

Measurement properties

The major measurement properties of the different tools are shown in Table 4. Sample sizes among the qualitative and mixed methods approaches varied between n=12 [55] and n=62 [56], while numbers were naturally much larger for the secondary data analyses (between n=2,814 [58] and n=25,180 [57]). Target groups varied widely, from adults in general [60], young adults [57] and older adults [54, 59] to adults with special conditions or characteristics [23, 53, 55, 56] and women [58].

Sample sizes for the quantitative tools varied substantially, ranging from n=20 [30] to n=2,501 [41]. For the ICECAP-O, six articles reported on the intended target group of adults over 65 [15, 28, 42, 43, 49, 50]. Later publications also applied it specifically to people with a medical condition [30, 51, 52], within a rehabilitation context [29, 38, 53], or for people over 60 with a hip fracture [44]. The ICECAP-A was developed to measure capabilities among adults over 18. Six of the identified articles used this target group [16, 35, 36, 39, 41, 54], while others validated it pointedly for adults with knee pain [37], opiate dependence [32], depression [40], or among women suffering irritative lower urinary tract syndrome [31].

More detailed psychometric properties were only reported for the quantitative measurement instruments. The most detailed results were available for questionnaires of the ICECAP-family. Both the ICECAP-O and the ICECAP-A were reported to have good construct [35, 37-39, 49], convergent, and discriminant validity [28-30, 44] when compared to the EQ-5D instrument to measure generic health status [55]. The ICECAP-O and ICECAP-A further showed good test-retest reliability [41, 42] and good internal consistency [30, 35, 36]. In addition, the ICECAP-A was also found to be significantly responsive among adults with knee pain [37] and women with irritative lower urinary tract syndrome [31]. No psychometric properties were reported for the ICECAP-SCM, ICECAP-FC questionnaires. In the OCAP family, no details were available for the original

questionnaire [19]. The OCAP-18 only yielded moderate construct validity when correlated with the EQ-5D-3L questionnaire [20]. The adaption of the OCAP for mental health showed good convergent validity, internal consistency and test-retest reliability [21, 33], which was confirmed for its German version [34]. With respect to the other questionnaires, Sacchetto et al. [23] reported good content and discriminant validity as well as internal consistency for the ACQ-CMH. The CADA questionnaire [25] reported good internal consistency for most questions, while the *Capability Measurement Tool for People with Chronic Pain* [24] did not report any psychometric properties.

Overall capabilities, capabilities for health, and capabilities for PA

While some of the questionnaires focus on the overall capabilities to pursue one's goals and being content with one's own life (e.g. the ICECAP questionnaires [15-17]), others are concerned with more specific aspects, such as enjoying recreational time, political views, making friends, or areas relevant to this study, e.g. bodily health and integrity (e.g. OCAP questionnaires [19-21]). Some questionnaires focus on specific subsets of health enhancing factors, such as the CADA [25], which is concerned with capabilities for healthy diet and PA but does not measure overall capabilities for health or well-being. A similar pattern can be for the qualitative tools: While Ndmoto et al. [46] focus on general capabilities for health, Abu-Zaineh et al. [48] explicitly deal with capabilities for health and self-management diabetes patients. Sauter et al. [26] is the only qualitative tool with a focus on capabilities for PA as an health-promoting factor.

Among the questionnaires, CADA [25] is the only one to directly measure capabilities for PA by specifically asking about resources (e.g. money to afford going to the gym) as well as environmental (e.g. indoor and outdoor PA spaces available), social (e.g. surrounding people are supportive of one's PA) and individual (e.g. mental and physical health influencing PA) factors of influence. The other questionnaires do not specifically ask for capabilities to pursue PA or sports but at least partially address areas that can be considered relevant for health-enhancing PA, such as physical suffering (ICECAP-SCM [17]), bodily health or enjoyment of recreational activities (OCAP [19] and OCAP-18 [20]). The qualitative tools do not explicitly address capabilities for PA. The only exception is Sauter et al. [26], which specifically asks for the individual (e.g. knowledge about PA), social (e.g. family and friends support) and environmental factors (e.g. offerings) that influence the opportunities of seniors in retirement homes to be physically active.

Discussion

The aim of this review has been to give an overview of the current state of research on available tools to measure capabilities for health based on the approach originally developed by Sen and Nussbaum, with a special focus on identifying those for health-enhancing PA. The systematic search was able to identify measurement tools for health and PA-related capabilities using qualitative, quantitative, and mixed methods between 2008 and 2019. It has explored the main features and psychometric properties of the identified tools, as well as their past application to different age and target groups.

Despite the number of papers identified, it is interesting to note that the number of distinct tools reported remains limited. For instance, there is a total of eleven questionnaire-based tools, most of which are

variations and adaptations of either the ICECAP or the OPAC questionnaire. Moreover, it is noteworthy that there is still no comprehensive measurement tool for health-related capabilities available that can be used among all target groups and all ages, especially when considering that the approach was first published in in 1985 [1], connected to well-being as early as 1993 [4], and has recently gained even more attention in the field of public health.

The analysis revealed a great degree of methodological variation regarding the development of the interview guidelines and questionnaires. Some studies approached the development from a more philosophical view and based their interview guideline [26] or questionnaire items [19, 22] on Martha Nussbaum's capability criteria [27]; others used an explorative approach, conducting focus-group [24, 25] or key-expert interviews [15, 16] to inductively develop their questionnaire. Another research group developed the questionnaire based solely on expert-group's opinion [22]. While our results allow no conclusions about which method is more appropriate or valid, those choosing a tool for a specific health promotion project should consider whether its development method and target group fit the intended application context.

The analyzed questionnaires that were empirically tested showed a moderate to good validity, reliability and responsiveness among different groups and compared to other questionnaires, mostly variations of the EQ-5D well-being questionnaire (i.e. EQ-5D-3L). This approach, however, poses an important theoretical issue, as it seems to imply that capability measures are better if they have a higher degree of correlation to measures of well-being. But according to Sen, well-being is a combination of "achieved functionings" [4], which are linked to but by no means perfectly correlated to a person's options (capabilities). If we take the capability approach seriously, we must necessarily expect a potentially considerable mismatch between the two measures, and using this kind of validation approach appears as generally problematic.

Another issue is that the number of items used to measure capabilities also varied considerably between questionnaires, i.e. between four items (ICECAP-O/ICECAP-A) and 104 items (OCAP). This raises the question whether all identified tools – even though they may have been validated – allow for measuring with the same accuracy. More research is required to investigate this, but in any case, health promoters interested in measuring capabilities will have to consider whether it will be feasible to administer the tool of their choice in practice, esp. regarding those with a large number of items.

Most questionnaires were developed for a specific population group, e.g. adults (ICECAP-A [16], OCAP/OCAP-18 [19, 20], CADA [25]), older adults (ICECAP-O [15]) or people suffering from mental illnesses (OxCAP-MH [21]; CQ-CMH [22]). However, even the general population questionnaires were often validated using samples of vulnerable population groups (e.g. adults with dementia [30], diabetes and obesity [25], or post hospitalization [29]). This may have implications for both the applicability of the results to the general populace and the question whether the tools adequately capture social inequalities, the reduction of which is at the heart of the capability approach [56].

Our research was guided by the intention to identify suitable tools for measuring capabilities for PA across the life-course. However, only two of the identified measurement instruments explicitly address PA, i.e. the

CADA questionnaire [25] and the interview-based tool by Sauter et al. [26]. However, CADA is not geared exclusively at PA but combines it with capabilities for healthy diet. In addition, it was developed for populations suffering from obesity rather than general populations. Similarly, Sauter et al.'s tool has a specific focus on senior citizens. In other questionnaires, only individual items might be considered relevant for PA, e.g. questions on bodily health [19-21]. Therefore, they cannot be applied to draw precise conclusions on PA capabilities of people. All in all, our study shows that more research is needed to develop appropriate (PA) capability instruments. First, these should focus on measuring PA and all its facets, including the individual (e.g. PA-related competence), social (e.g. social support for PA), and environmental (e.g. PA infrastructures and offers) conversion factors. Second, future measures should ideally be applicable to a maximum range of different settings, populations, and age-groups, thus allowing for standardized and comparable assessments of PA intervention effectiveness.

Despite our best of efforts, this study has some limitations which need to be borne in mind when interpreting its results and drawing conclusions. First, due to the heterogeneity of the tools identified, comparing individual instruments with each other was difficult, and it was therefore not possible to recommend a single tool that, in general, could be considered to be particularly appropriate. For the same reason, a more systematic quality assessment of the primary studies, as required by the PRISMA checklist, was not possible. All in all, however, we are confident that this review provides a good initial overview in an innovative and increasingly relevant area of research. Having been conducted on a large number of databases and employing an additional hand search, it presents details on different types of instruments that may guide the selection of appropriate tools for specific purposes in future research projects.

Conclusion

This systematic review has shown that there is currently no measurement tool available that can be adopted across various target groups and settings to assess capabilities for health and health-related PA. Available tools vary substantially regarding their underlying assumptions, focus on capabilities, properties (e.g. language, number of items), development processes, measurement approaches, and addressees. Most of the quantitative tools have been empirically shown to be valid, reliable and responsive, but the methods employed for validation invite skepticism as to whether all instruments truly measure capabilities and/or do so in a meaningful way. At this point in time, it is not possible to recommend a single tool for general use, and health promoters may want to choose carefully or even consider adapting a tool to their specific needs.

Our findings thus seem to echo Sen's own concerns about the empirical difficulties of operationalizing the capability approach [1, 56], as well as those of other researches who have demurred that the multidimensional, context-dependent, and normative nature of the approach can pose problems for operationalization [57-59].

These difficulties notwithstanding, the Capital4Health consortium, under whose auspices this review was conducted, is planning to contribute to the further development of capability measurement in two important ways: First, two of the consortium projects – one working in childcare centres (QueB, [60]) and

another in the community setting (Action for Men, [61]) – will apply the qualitative tool developed by Sauter et al. [26] in order to test its applicability in different environments and its potential for the comparative evaluation of capabilities. Second, in an attempt to strike a compromise between detailed but setting-exclusive tools and overly generic instruments, the consortium is currently employing a participatory approach to co-develop a framework for conceptualizing and measuring capabilities for PA across the life-course. This framework [62] is intended to define a number of principles that will ensure a greater amount of comparison between age groups and settings while still allowing for the use of adapted instruments in different contexts. Such research may eventually result in a set of improved measurement tools, thus helping to further establish the capability approach more firmly in health promotion and PA intervention research.

References

1. Sen A: **Commodities and Capabilities**: North-Holland; 1985.
2. Ruger JP: **Health capability: conceptualization and operationalization**. *American Journal of Public Health* 2010, **100**(1):41-49.
3. Abel T, Frohlich KL: **Capitals and capabilities: linking structure and agency to reduce health inequalities**. *Social Science & Medicine* 2012, **74**(2):236-244.
4. Sen A: **Capability and Well-Being**. In: *The Quality of Life.* edn. Edited by Nussbaum, Sen. Oxford: Clarendon Press; 1993.
5. Robeyns I: **The Capability Approach in Practice***. *Journal of Political Philosophy* 2006, **14**(3):351-376.
6. Helter TM, Coast J, Laszewska A, Stamm T, Simon J: **Capability instruments in economic evaluations of health-related interventions: a comparative review of the literature**. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2020, **29**(6):1433-1464.
7. Warburton DE, Nicol CW, Bredin SS: **Health benefits of physical activity: the evidence**. *Canadian Medical Association Journal* 2006, **174**(6):801-809.
8. Reiner M, Niermann C, Jekauc D, Woll A: **Long-term health benefits of physical activity–a systematic review of longitudinal studies**. *BMC public health* 2013, **13**:813.
9. Brownson RC, Seiler R, Eyler AA: **Measuring the impact of public health policy**. *Preventing chronic disease* 2010, **7**(4):A77.
10. McKinlay JB: **Paradigmatic obstacles to improving the health of populations–implications for health policy**. *Salud publica de Mexico* 1998, **40**(4):369-379.
11. Frahsa A, Abel T, Gelius P, Rütten A, consortium ObotCH: **The capability approach as a bridging framework across health promotion settings: Theoretical and empirical considerations**. *Health Promotion International* under Review.
12. Antonovsky A: **Unraveling the mystery of health how people manage stress and stay well**; 1987.

13. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement**. *PLoS medicine* 2009, **6**(7):e1000097.
14. Anand P, Hunter G, Smith R: **Capabilities and Well-Being: Evidence Based on the Sen–Nussbaum Approach to Welfare**. *Social Indicators Research* 2005, **74**(1):9-55.
15. Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, Peters TJ: **Valuing the ICECAP capability index for older people**. *Social Science & Medicine* 2008, **67**(5):874-882.
16. Al-Janabi H, Flynn TN, Coast J: **Development of a self-report measure of capability wellbeing for adults: the ICECAP-A**. *Quality of Life Research* 2012, **21**(1):167-176.
17. Sutton EJ, Coast J: **Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods**. *Palliative Medicine* 2014, **28**(2):151-157.
18. Al-Janabi H: **Do capability and functioning differ? A study of UK survey responses**. *Health Economics* 2018, **27**(3):465-479.
19. Anand P, Hunter G, Carter I, Dowding K, Guala F, Van Hees M: **The Development of Capability Indicators**. *Journal of Human Development and Capabilities* 2009, **10**(1):125-152.
20. Lorgelly PK, Lorimer K, Fenwick EAL, Briggs AH, Anand P: **Operationalising the capability approach as an outcome measure in public health: The development of the OCAP-18**. *Social Science & Medicine* 2015, **142**:68.
21. Simon J, Anand P, Gray A, Rugkåsa J, Yeeles K, Burns T: **Operationalising the capability approach for outcome measurement in mental health research**. *Social Science & Medicine* 2013, **98**:187.
22. Sacchetto B, Aguiar R, Vargas-Moniz MJ, Jorge-Monteiro MF, Neves MJ, Cruz MA, Coimbra JA, Ornelas J: **The Capabilities Questionnaire for the Community Mental Health Context (CQ-CMH): A measure inspired by the capabilities approach and constructed through consumer–researcher collaboration**. *Psychiatric Rehabilitation Journal* 2016, **39**(1):55-61.
23. Sacchetto B, Ornelas J, Calheiros MM, Shinn M: **Adaptation of Nussbaum's Capabilities Framework to Community Mental Health: A Consumer-Based Capabilities Measure**. *American Journal of Community Psychology* 2018, **61**(1/2):32-46.
24. Kinghorn P, Robinson A, Smith RD: **Developing a Capability-Based Questionnaire for Assessing Well-Being in Patients with Chronic Pain**. *Social Indicators Research* 2015, **120**(3):897-916.
25. Ferrer RL, Cruz I, Burge S, Bayles B, Castilla MI: **Measuring capability for healthy diet and physical activity**. *The Annals of Family Medicine* 2014, **12**(1):46-56.
26. Sauter A, Curbach J, Rueter J, Lindacher V, Loss J: **German senior citizens' capabilities for physical activity: a qualitative study**. *Health Promot International* 2019, **34**(6):1117-1129.
27. Nussbaum M: **Women's Capabilities and Social Justice**. *Journal of Human Development* 2000, **1**(2):219-247.
28. Makai P, Brouwer WBF, Koopmanschap MA, Nieboer AP: **Capabilities and quality of life in Dutch psycho-geriatric nursing homes: An exploratory study using a proxy version of the ICECAP-O**. *Quality of Life Research* 2012, **21**(5):801-812.

29. Makai P, Koopmanschap MA, Brouwer WB, Nieboer AA: **A validation of the ICECAP-O in a population of post-hospitalized older people in the Netherlands**. *Health and quality of life outcomes* 2013, **11**:57.
30. Sarabia-Cobo CM, Paras-Bravo P, Amo-Setien FJ, Alconero-Camarero AR, Saenz-Jalon M, Torres-Manrique B, Sarabia-Lavin R, Fernandez-Rodriguez A, Silio-Garcia T, Fernandez-Pena R *et al*: **Validation of the Spanish Version of the ICECAP-O for Nursing Home Residents with Dementia**. *PloS one* 2017, **12**(1):e0169354.
31. Goranitis I, Coast J, Al-Janabi H, Latthe P, Roberts TE: **The validity and responsiveness of the ICECAP-A capability-well-being measure in women with irritative lower urinary tract symptoms**. *Quality of Life Research* 2016, **25**(8):2063-2075.
32. Goranitis I, Coast J, Day E, Copello A, Freemantle N, Seddon J, Bennett C, Frew E: **Measuring Health and Broader Well-Being Benefits in the Context of Opiate Dependence: The Psychometric Performance of the ICECAP-A and the EQ-5D-5L**. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2016, **19**(6):820-828.
33. Vergunst F, Jenkinson C, Burns T, Anand P, Gray A, Rugkåsa J, Simon J: **Psychometric validation of a multi-dimensional capability instrument for outcome measurement in mental health research (OxCAP-MH)**. *Health and Quality of Life Outcomes* 2017, **15**.
34. Laszewska A, Schwab M, Leutner E, Oberrauter M, Spiel G, Simon J: **Measuring broader wellbeing in mental health services: validity of the German language OxCAP-MH capability instrument**. *Quality of Life Research* 2019, **28**(8):2311-2323.
35. Tang C, Xiong Y, Wu H, Xu J: **Adaptation and assessments of the Chinese version of the ICECAP-A measurement**. *Health and quality of life outcomes* 2018, **16**(1):45.
36. Linton M-J, Mitchell PM, Al-Janabi H, Schlander M, Richardson J, Iezzi A, Ubels J, Coast J: **Comparing the german translation of the icecap-a capability wellbeing measure to the original english version: Psychometric properties across healthy samples and seven health condition groups**. *Applied Research in Quality of Life* 2018.
37. Keeley T, Al-Janabi H, Nicholls E, Foster NE, Jowett S, Coast J: **A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention**. *Quality of Life Research* 2015, **24**(10):2319-2331.
38. Couzner L, Ratcliffe J, Crotty M: **The relationship between quality of life, health and care transition: an empirical comparison in an older post-acute population**. *Health and quality of life outcomes* 2012, **10**:69.
39. Al-Janabi H, Peters TJ, Brazier J, Bryan S, Flynn TN, Clemens S, Moody A, Coast J: **An investigation of the construct validity of the ICECAP-A capability measure**. *Quality of Life Research* 2013, **22**(7):1831-1840.
40. Mitchell PM, Al-Janabi H, Byford S, Kuyken W, Richardson J, Iezzi A, Coast J: **Assessing the validity of the ICECAP-A capability measure for adults with depression**. *BMC psychiatry* 2017, **17**(1):46.
41. Al-Janabi H, Van Exel J, Brouwer W, Trotter C, Glennie L, Hannigan L, Coast J: **Measuring Health Spillovers for Economic Evaluation: A Case Study in Meningitis**. *Health Economics* 2016, **25**(12):1529-

1544.

42. Hörder H, Gustafsson S, Rydberg T, Skoog I, Waern M: **A Cross-Cultural Adaptation of the ICECAP-O: Test-Retest Reliability and Item Relevance in Swedish 70-Year-Olds**. *Societies* 2016, **6**(4):30-n/a.
43. Franklin M, Payne K, Elliott RA: **Quantifying the Relationship between Capability and Health in Older People: Can't Map, Won't Map**. *Medical Decision Making* 2018, **38**(1):79-94.
44. Milte R, Crotty M, Miller MD, Whitehead C, Ratcliffe J: **Quality of life in older adults following a hip fracture: an empirical comparison of the ICECAP-O and the EQ-5D-3 L instruments**. *Health and quality of life outcomes* 2018, **16**(1):173.
45. Weaver RR, Lemonde M, Payman N, Goodman WM: **Health capabilities and diabetes self-management: the impact of economic, social, and cultural resources**. *Social science & medicine* 2014, **102**:58-68.
46. Ndomoto L, Hibble A, Obuzor G, Nthusi N, Quine A, Chahal P, Barasa SO, Nyanja N, Tulinius C: **Understanding the fundamental elements of global health: Using the sen capability approach as the theoretical framework for a health needs assessment in deprived communities**. *Education for Health (Abingdon)* 2018, **31**(1):43-47.
47. Petros R, Solomon P, Linz SJ, DeCesaris M, Hanrahan NP: **Capabilities approach: Contextualizing participants' perspectives on systems barriers to recovery**. *International Journal of Mental Health* 2016, **45**(4):262-278.
48. Abu-Zaineh M, Woode ME, Giacaman R: **Youth wellbeing through the lens of the Senian capability approach: insights from the occupied Palestinian territory: a cross-sectional study**. *Lancet* 2018, **391**:S53-S53.
49. Coast J, Smith RD, Lorgelly P: **Welfarism, extra-welfarism and capability: the spread of ideas in health economics**. *Social Science & Medicine* 2008, **67**(7):1190-1198.
50. Flynn TN, Chan P, Coast J, Peters TJ: **Assessing quality of life among British older people using the ICEPOP CAPability (ICECAP-O) measure**. *Applied health economics and health policy* 2011, **9**(5):317-329.
51. Horwood J, Sutton E, Coast J: **Evaluating the face validity of the ICECAP-O capabilities measure: A 'think aloud' study with hip and knee arthroplasty patients**. *Applied Research in Quality of Life* 2014, **9**(3):667-682.
52. Davis JC, Best JR, Dian L, Khan KM, Hsu CL, Chan W, Cheung W, Liu-Ambrose T: **Are the EQ-5D-3L and the ICECAP-O responsive among older adults with impaired mobility? Evidence from the Vancouver Falls Prevention Cohort Study**. *Quality of Life Research* 2017, **26**(3):737-747.
53. Davis JC, Liu-Ambrose T, Richardson CG, Bryan S: **A comparison of the ICECAP-O with EQ-5D in a falls prevention clinical setting: are they complements or substitutes?** *Quality of Life Research* 2013, **22**(5):969-977.
54. Al-Janabi H, Keeley T, Mitchell P, Coast J: **Can capabilities be self-reported? A think aloud study**. *Social Science & Medicine* 2013, **87**:116-122.

55. EuroQoL-Group: **EuroQol - a new facility for the measurement of health-related quality of life**. *Health Policy* 1990, **16**(3):199-208.
56. Sen A: **Inequality Reexamined**: Russell Sage Foundation; 1992.
57. Sugden R: **Welfare, Resources, and Capabilities: A Review of Inequality Reexamined by Amartya Sen**. *Journal of Economic Literature* 1993, **31**(4):1947-1962.
58. Ysander B: **Robert Erikson: descriptions of inequality**. In: *Quality of life* edn. Edited by Nussbaum M, Sen A. Oxford Clarendon Press; 1993.
59. Srinivasan TN: **Human Development: A New Paradigm or Reinvention of the Wheel?** *The American economic review* 1994, **84**(2):238-243.
60. Müller C, Foitzik E, Hassel H: **Bewegte Kitas durch Organisationsentwicklung**. *Prävention und Gesundheitsförderung* 2019, **15**(1):50-55.
61. Strobl H, Brew-Sam N, Curbach J, Metz B, Tittlbach S, Loss J: **ACTION for Men: Study Protocol of a Community Capacity Building Intervention to Develop and Implement Gender-Sensitive Physical Activity Programs for Men 50 Plus**. *Frontiers in Public Health* 2020, **8**:4.
62. Till M, P. PG, Abu-Omar K, Abel T: **Using the Capability Approach in health promotion projects: a Framework for implementation.** in preparation.
63. Bucki B, Spitz E, Etienne AM, Le Bihan E, Baumann M: **Health capability of family caregivers: how different factors interrelate and their respective contributions using a Bayesian approach**. *BMC public health* 2016, **16**:364.
64. Douptcheva N, Hill AG: **Functionings and capabilities as tools for explaining differences in self-assessed health: the case of women's health in Accra, Ghana**. *African population studies / Etude de la population Africaine* 2014, **28**(2):744-757.
65. Tellez J, Krishnakumar J, Bungener M, Le Galès C: **Capability deprivation of people with Alzheimer's disease: An empirical analysis using a national survey**. *Social science & medicine* 2016, **151**:56.
66. Anand P, Santos C, Smith R: **The measurement of capabilities**. *Open Discussion Paper in Economics Open University* 2007, **67**:49-89.
67. Grewal I, Lewis J, Flynn T, Brown J, Bond J, Coast J: **Developing attributes for a generic quality of life measure for older people: preferences or capabilities?** *Social Science & Medicine* 2006, **62**(8):1891-1901.
68. Makai P, Brouwer WBF, Koopmanschap MA, Nieboer AP: **Capabilities and quality of life in Dutch psycho-geriatric nursing homes: An exploratory study using a proxy version of the ICECAP-O**. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation* 2012, **21**(5):801-812.
69. Al-Janabi H, Flynn TN, Peters TJ, Bryan S, Coast J: **Test–retest reliability of capability measurement in the UK general population**. *Health Economics* 2015, **24**(5):625-630.
70. Simon J, Łaszewska A, Leutner E, Spiel G, Churchman D, Mayer S: **Cultural and linguistic transferability of the multi-dimensional OxCAP-MH capability instrument for outcome measurement in mental health: The German language version**. *BMC psychiatry* 2018, **18**.

Abbreviations

ADL	activities of daily living
ADRQL	Alzheimer's disease related Quality of life
BPRS	Brief Psychiatric Rating Scale
BPRS	Brief Psychiatric Rating Scale
C	Comparison
CQ-CMH	Capability Questionnaire for Community Mental Health
CTM-3	3-Item Care Transition Measure
D	Development
DASS-D	Depression Anxiety Stress Scales
EFA	exploratory factor analysis
FG	Focus group
GAF	Global Assessment of Functioning
ICC	Intra-class correlation coefficient
ICECAP-A	ICEpop CAPability measure for adults
ICECAP-FC	ICEpop CAPability and Functioning measure
ICECAP-O	ICEpop CAPability measure for older people
ICECAP-SCM	ICEpop CAPability measure for supportive care
K6	Kessler Psychological Distress Scale
M	Mean
M	Measurement
OCAP	Oxford Capability Questionnaire
OLS	ordinary least square
OXCAP-MH	Oxford Capability Questionnaire for Mental Health

PA	Physical Activity
PHQ-8	Patient Health Questionnaire depression scale
RAS	Recovery Assessment Scale
RT	Researcher Team
SEM	Structural equation modeling
SIX	Objective Social Outcomes Index
V	Validation
WB	Well-Being

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and materials: Not applicable

Competing interests: The authors declare that they have no competing interests

Funding: The research network Capital4Health [grant number: 01EL1421A-F] and the associated subproject CAPCOM [grant number: 01EL1421A] are funded by the German Federal Ministry for Education and Research (BMBF)

Authors' contributions: Conceptualization, MT., KAO., and PG.; methodology, MT., KAO, PG, AKR.; Screening, MT, SF; supervision, PG, KAO; full-text analysis, MT; writing—original draft preparation MT, PG; All authors have read and agreed to the published version of the manuscript.

Acknowledgements: We thank Leonie Birkholz for her support during the literature search.

Tables

Table 1: Inclusion and exclusion criteria

Criterion	Inclusion	Exclusion
Time	January 2000 – June 2019	Studies before 2000 and after June 2019
Language	English, German	Any other language
Type of publication	Journal Articles	Scientific papers published outside a journal
Focus of study	• Operationalization of the Capability Approach in terms of Amartya Sen/ Martha Nussbaum • Capability approach used in the context of health/heath enhancing physical activity	• Outside context of health • Pure article on theory without operationalization of the capability approach
Study population	Any Population	Nil
Setting	Any Setting	Nil
Country	Any country	Nil

Table 2: Distribution of measurement tools

			No.		
Capability Measurement Tools	**Qualitative tools**		1	Interviews [26, 45, 46]	
			2	Videography [47]	
	Mixed-Method		3	Questionnaire and Interviews [63]	
	Quantitative tools		4	Questionnaire combinations (secondary data) [14, 48, 64, 65]	
		Questionnaires	*5*	*ICECAP*	*ICECAP-O* [15]*
			6		*ICECAP-A* [16]*
			7		*ICECAP-SCM* [17]*
			8		*ICECAP-FC* [18]*
			9	*OCAP*	*OCAP* [19]*
			10		*OCAP-18* [20]*
			11		*OXCAP-MH* [21]*
			12	*CQ-CMH*	*CQ-CMH* [22]*
			13		*ACQ-CMH* [23]*
			14	*Capability-based questionnaire –well-being in patients with chronic pain* [24]*	
			15	*CADA* [25]*	

*Note: This table only indicates the articles reporting on the development of the respective tool. ICECAP-O= ICEpop CAPability measure for older people; ICECAP-A= ICEpop CAPability measure for adults; ICECAP-FC= ICEpop CAPability and Functioning measure; ICECAP-SCM= ICEpop CAPability measure for supportive care; OCAP= Oxford Capability Questionnaire; OXCAP-MH= Oxford Capability Questionnaire for Mental Health; CQ-CMH= Capability Questionnaire for Community Mental Health

Table 3: Description of included studies and tools

Qualitative tools

No.	Tool	No. of items	Author (year) country	Study Aim	Focus of tool	Language of tool	Target population	Method
1.	Interview	8	Weaver et al, (2014)[45], Canada	M	Health and diabetes self-management	English	Adults with diabetes	Measurement via semi-structured interviews; Analysis via two researchers
		n.a.	Ndomoto et al. (2018) [46], UK	M	Health	English	Adults living in rural Kenya and urban deprived UK	Measurement via key informant interviews; *FG* and participant observation.
		10	Sauter et al. (2018)[26], Germany	D/M	Health enhancing PA	German/ English	Older adults living in senior residences	Development of interview-guide by *RT* based on Anand's capability questionnaire [66] and literature on older adults physical activity; Measurement via semi-structured interviews
2.	Videography	n.a.	Petros et al. (2016)[47], USA	M	Mental health recovery	English	Adults with mental illness	Four-week measurement via videography on the topic *Tell us about your recovery*; No *RT* present during recording; Transcription and analysis of data by *RT*

Mixed method

No.	Instrument	No. of items	Author (year) country	Study Aim	Focus of Tool	Language of tool	Target population	Method
3.	Questionnaire and Interview	20	Bucki et al. (2016)[63], Luxembourg	C	Health	Luxembourgish, Portuguese, French, German	Adult care givers	Measurement of relations between health capability factors of care givers using questionnaire-based (HCFC-8) interviews. Statistical analysis using Monte Carlo Markov Chain algorithms.

Quantitative tools

No.	Instrument	No. of	Author (year) country	Study Aim	Focus of Tool	Language of tool	Target population	Method

		items						
4.	Questionnaires used in secondary data	n.a.	Abu-Zaineh & Woode (2018) [48], France	M	Health and self-management	English	Adults living in Palestine	Measurement of capabilities (health awareness, knowledge and living conditions) via Exploratory Structural Equation Modelling using data from the Palestinian Family Survey.
		n.a.	Anand et al. (2005)[14], UK	M	General WB	English	Adults living in British households	Measurement of capabilities and well-being by regression using data of the British Household Panel Survey
		n.a.	Douptcheva et al. (2014) [64], UK	M	Health	English	Women living in Accra	Measurement of capabilities and functionings to identify factors that influence our health using data from the Women's Health Study of Accra – Wave II.
		1,760	Tellez et al. (2016)[65], France	M	WB	French	Older adults	Measurement of capabilities (freedom to perform self-care activities, freedom to participate in life of the household) by use of a latent variable modelling framework analyzing the 2008 Disability and Health Household Survey of France.
5.	ICECAP/ ICECAP-O	5	Coast et al. (2008)[15], UK	D	General WB	English	Adults ≥65	Lay terms defined by *RT* based on in-depth interviews [67]. Iterative semi-structured interviews to ensure understandable language. Valuation via survey interviews.
		5	Coast et al.	V	General WB	English	Adults	Validation via

	(2008)[49], UK				≥65	Chi-square analysis against socio-demographic information, health, nature of locality and environment, social support, participation, and comparison of data to priori set *RT*-expectations
5	Flynn et al. (2011)[50], UK	V	General WB	English	Adults ≥65	Construct validity measurement of tariff scores (Comparison with qualitative interviews of attribute development [67] and subjective wellbeing literature)
5	Couzner et al. (2012)[38], Australia	V	General WB	English	Adults ≥65	Measurement of relationship of ICECAP-O to EQ-5D and CTM-3 through Spearman's rho, *t*-tests and chi-square tests.
5	Makai et al. (2012)[68], Netherlands	D/V	General WB	Dutch	Adults ≥65	Forward-backward-translation into Dutch by two independent translators; Measurement of concurrent (correlations of the nursing and family version with EQ-5D, EQ-VAS, Cantril's ladder, overall life satisfaction) and discriminant validity (chi-square and Mann-Whitney *U* tests*)*
5	Davis et al. (2013)[53], Canada	V	General WB	English	Adults ≥65	Comparison against the EQ-5D using EFA
5	Makai et al. (2013)[29], Netherlands	V	General WB	Dutch	Adults ≥65	Measurement of convergent (correlation with EQ-5D, IADL, GDS-15, SPF-IL and Cantril's

							ladder) and discriminant validity (*t* test, one-way ANOVA and stepwise regression analyses)
		5	Horwood et al. (2014)[51], UK	V	General WB	English	Adults ≥65	Face-validity measurement via “think aloud” study analysis and frequency of participant’s problems
		5	Hörder et al. (2016)[42], Sweden	V	General WB	Swedish	Adults ≥65	Test-retest reliability (1-2 weeks apart) and item relevance measure (participants rated items from 0-100)
		5	Davis et al. (2017)[52], Canada	V	General WB	English	Adults ≥65	Measurement of responsiveness (regression on age, sex, and faller status)
		5	Sarabia-Cobo et al. (2017)[30], Spain	V	General WB	Spanish	Adults ≥65	Measurement of construct (factor analysis) and convergent validity (correlation with dimensions of the EQ-5D+C, ADRQL, ADL), and reliability (internal consistency-Cronbach Alpha)
		5	Franklin et al. (2018)[43], UK	C/V	General WB	English	Adults ≥65	Comparison of (1) tariff scores using OLS and CLAD regression models and (2) domain scores using MNL regression against the EQ-5D-3L
		5	Milte et al. (2018)[44], Australia	C/V	General WB	English	Adults ≥65	Comparison against the EQ-5D-3L using Spearman correlation coefficient and multiple linear regression
6.	ICECAP-A	5	Al-Janabi et al. (2012) [16], UK	D	General WB	English	Adults ≥18	Identification of important components of life

						through in-depth interviews; Iterative semi-structured interviews to refine attributes to a self-completion measure with one item per attribute
5	Al-Janabi et al. (2013) [54], UK	V	General WB	English	Adults ≥18	Think-aloud and semi-structured interviews to assess the feasibility of a self-reporting capability measurement
5	Al-Janabi et al. (2013) [39], UK	V	General WB	English	Adults ≥18	Measurement of construct validity (univariate analysis and correlations based on hypotheses made in advance)
5	Al-Janabi et al. (2016) [69], UK	V	General WB	English	Adults ≥18	Measurement of test-retest reliability (ICC-baseline and 2-week capability index scores)
5	Keeley et al. (2015)[37], UK	V	General WB	English	Adults ≥18	Measurement of responsiveness (anchor-based analysis; anchors: EQ-5D-3L, GAD-7, PHQ-8)
5	Goranitis et al (2016)[31], UK	V	General WB	English	Adults ≥18	Measure of acceptability, construct validity (convergent: Pearson's correlation with EQ-5D-3L and ICIQ-OAB, Spearman's correlation coefficient across dimension scores, and index and dimension scores; discriminant: one-way ANOVA and Kruskal-Wallis *H* test)
5	Goranitis et al (2016)[32], UK	V	General WB	English	Adults ≥18	Assessment of construct validity (convergent: Pearson's

							correlation with EQ-5D-5L; Discriminant: univariate and multivariate analysis) and sensitivity to change
		5	Mitchell et al. (2017)[40], UK	V	General WB	English	Adults ≥18	Concept-mapping from condition-specific and capability items; Discriminant validity testing (Mann-Whitney *U* test using DASS-D and K10 data; Multivariable regression analysis using OLS)
		5	Linton et al. (2018)[36], Germany	V	General WB	German	Adults ≥18	Measurement of internal-consistency (Cronbach's Alpha), convergent (Pearson's correlation with EQ-5D-3L, SF-6D, SWLS scores), and construct validity (OLS regressions)
		5	Tang et al. (2018)[35], China	D/V	General WB	Chinese	Adults ≥18	*RT* translated original version into Chinese; *FG* evaluated appropriateness of the translation; pilot testing; backward translation; online-survey to check acceptability, reliability (item correlations), and validity (EFA and correlations with EQ-5D-3L and EQ-VAS)
7.	ICECAP-SCM	7	Sutton & Coast (2014) [17], UK	D	WB in end of life care	English	People at end of life	Interviews to determine conceptual elements of a good death; follow-up interviews to

								check conceptual attributes
8.	ICECAP-FC	10	Al-Janabi (2018) [18], UK	D	WB capabilities and functionings	English	Adults ≥18	ICECAP-A modified with additional question on functioning to each attribute by *RT*
9.	OCAP	64	Anand et al. (2009)[19], UK	D	General Capabilities (e.g. enjoying recreational time, political views, making friends bodily health and integrity)	English	Adults ≥18	Development of items based on Nussbaum criteria [27]
10.	OCAP-18	18	Lorgelly et al. (2015)[20], UK	D	General Capabilities (e.g. enjoying recreational time, political views, making friends bodily health and integrity)	English	Adults ≥18	Items, based on OCAP-questionnaire [19], reduced on analysis of *FG*, cognitive interviews, and factor analysis
11.	OxCAP-MH	16	Simon et al. (2013)[21], UK	D/V	General capabilities for mental health	English	Adults ≥18 with a mental illness	Adaption of the OCAP-18 [20] based on expert-*FG* and validation (correlation with GAF, EQ-5D-VAS, EQ-5D-3L)
		16	Vergunst et al. (2017) [33], UK	V	General capabilities for mental health	English	Adults ≥18 with a mental illness	Measurement of internal-consistency (Cronbach's alpha), test-retest (1-week apart; ICC), and construct validity (correlation with EQ-5D, BPRS, GAS, SIX)
		16	Simon et al. (2018)[70], UK	D/V	General capabilities for mental health	English	Adults ≥18 with a mental illness	Forward-backward-translation of OxCAP-MH into German and linguistic validation

								through German native speakers
		16	Laszewska et al. (2019) [34], Austria	C/V	General capabilities for mental health	German	Adults ≥18 with a mental illness	Comparison against the EQ-5D-5L (EFA). Measurement of responsiveness (anchor questionnaires and standardized response mean), discriminant validity (subgroup comparison using *t* test and one-way ANOVA), and test-retest (ICC; baseline - max 30 days after)
12.	CQ-CMH	104	Sacchetto et al. (2016) [22], Portugal	D/V	Mental Health	Portuguese	Consumers of mental health services	*FG* interview data analysis; development of item/rating scale by steering committee and additional comparison with Nussbaum criteria [27]; Assessment of face-validity
13.	ACQ-CMH-98	98	Sacchetto et al. (2018) [23], Portugal	D/V	Mental Health	Portuguese	Consumers of mental health services	Adaption of the CQ-CMH questionnaire [22] based on panel members judgement; Measurement of validity (correlation with WHOQOL-Bref, RAS, K6)
14.	Capability-based questionnaire	8	Kinghorn et al. (2015) [24], UK	D	WB	English	People suffering from chronic pain	*FG* interview and individual interviews to identify list of important capabilities; Development of questionnaire for self-completion based on identified capabilities by *RT*
15.	CADA	34	Ferrer et al. (2014)[25], USA	D	Physical Activity and Diet	English	Adults with obesity and diabetes	*FG* interviews were used to identify important themes; questionnaire

	created by *RT* based on themes

ADL=activities of daily living; ADRQL=Alzheimer's disease related Quality of life; BPRS=Brief Psychiatric Rating Scale; C=Comparison; CTM-3=3-Item Care Transition Measure; D=Development; EFA=exploratory factor analysis; *FG*=Focus group; GAF=Global Assessment of Functioning; ICC=Intra-class correlation coefficient; M=Measurement*;* OLS=ordinary least square; RAS=Recovery Assessment Scale; *RT*=Researcher Team; SIX=Objective Social Outcomes Index; V=Validation; WB= Well-Being

Table 4: Psychometric properties of the identified tools

Qualitative/mixed methods

No.	Instrument	Author (year)	Sample size, description of study population; age; % male; country
1.	Interviews	Weaver et al, (2014)[45]	n=45; adults with diabetes; M=60; 42%; Canada
		Ndomoto et al. (2018)[46]	n=55; whole community; n.a.; n.a.; rural Kenya and urban deprived UK
		Sauter et al. (2018)[26]	n=26; older adults; 65+;38%; Germany
2.	Videography	Petros et al. (2016)[47]	n=12; adults with mental illness; n.a.; n.a.; USA

Mixed method

No.	Instrument	Author (year)	Sample size, description of study population; age; % male; country
3.	Questionnaire and Interviews	Bucki et al. (2016)[63]	n=62; adult care givers; M=59; 36%; Luxembourg

Quantitative

No.	Instrument	Author (year)	Sample size, description of study population; age; % male; country	Validity	Reliability/ Responsiveness/ Sensitivity
4.	Secondary Data Analysis	Abu-Zaineh & Woode (2018) [48]	N=25,180; young adults: M=21; 50%; Palestine		
		Anand et al. (2005)[14]	n=12,040; adults; 18+; 45%; UK		
		Douptcheva et al. (2014)[64]	n=2,814; women; 18+; 0%; Ghana		
		Tellez et al. (2016)[65]	n=8,841; older adults; 60+; n.a.; France		
5.	ICECAP/ ICECAP-O (5)	Coast et al. (2008)[15]	n=255; older adults; 65+; 56%; UK	n.a.	n.a.

Coast et al. (2008)[49]	n=314; older adults; 65+; 54%; UK	***Construct:*** EQ-5D overall value and Attachment χ^2 = .42 Security χ^2 = .008 (p <.01) Role χ^2 = <.001 (p <.01) Enjoyment χ^2 = <.001 (p <.01) Control χ^2 = <.001 (p <.01)	n.a.
Flynn et al. (2011)[50]	n = 809; older adults; 65+; 49%; UK	***Construct:*** comparison ICECAP-O tariff scores with qualitative interviews of attribute development [67] and subjective wellbeing literature provides construct validity.	
Couzner et al. (2012)[38]	n=82; older adults in rehabilitation; M=76; 50%; Australia	***Construct:*** EQ-5D overall value and Attachment χ^2 = .741 Security χ^2 = .088 Role χ^2 = .092 Enjoyment χ^2 = .058 Control χ^2 = .043 (p<.05) ICECAP-O and CTM-3 Spearman's r = .23; (p<.05) and EQ-5D Spearman's r = .44; (p<.001)	n.a.
Makai et al. (2012)[68]	n=122; older adults; M=82; 32%; Netherlands	***Convergent/discriminant:*** ICECAP-O and nursing version of EQ-5D r = .48 (p <.001) Overall life r = .52 (p <.001) ICECAP-O and family version of EQ-5D r = .57 (p<.001) Overall life r = .48 (p<.001)	n.a.
Davis et al. (2013)[53]	n=215; older adults post falls; M=79; n.a.; Canada	***Construct:*** Two factor analysis indicated two separate but correlated factors, supporting that the instruments provide complementary data with RMSEA (90% CI) = .05 (.00 – .09)	n.a.
Makai et al. (2013)[29]	n=275; older adults post hospitalization; 65+; 46%; Netherlands	***Convergent:*** Correlation ICECAP-O significant to Cantril's ladder r =.51(p<.001) SPF_IL r =.60(p<.001) EQ-5D r =.40(p<.001) SF-20 r =.47(p<.001) ***Discriminant:*** EQ5D Top 50% M=.90 (p <.01) Bottom 50% M=.80 Multimorbid Max. 1 chronic condition M=.89 (p <.01) More than 2 conditions	n.a.

				M=.82
Horwood et al. (2014)[51]	n=20; older adults with hip/knee replacement; M=70; 30%; UK	***Face:*** Majority of participants had no problems completing the measure		n.a.
Hörder et al. (2016)[42]	n=40; older adults; 70; 48%; Sweden	n.a.		***Test-retest:*** ICC=.80 systematic disagreement Cohen's κ (95% CI) attachment κ=.34; -.17 (-.35 – -.03) (significant) Security κ=.22; .05 (-.11–.20) Role κ=.41; .00 (-.16–.16) Enjoyment κ=.24; -.02 (-.19 – .14) Control κ=.17; -.13 (-.32 – .05)
Davis et al. (2017)[52]	n=247; older adults with impaired mobility; 80±7; 37%; Canada	n.a.		***Responsiveness:*** Change Baseline to 12-Month follow up: M = -.016 (p < .05); r=.50 Relation of change divided by faller status: Change Baseline to 12-Month follow up, faller: M =-.13 Change Baseline to 12-Month follow up, non-faller: M = .00
Sarabia-Cobo et al. (2017)[30]	n=217; older adults with dementia; M= 87; 19%; Spain	***Convergent:*** Correlation EQ-5D+C to ICECAP-O tariff: r = .62 (p<.01) Attachment r = .11 (p<.05) Security r = .32 (p<.05) Role r = .71 (p<.01) Enjoyment r = .56 (p<.01) Control r = .41 (p<.01)	***Discriminant:*** Depression severity Mildmean M = .72 (p <.01) Moderate M = .63 Severe M = .50 Care level Low M = .70 (p <.01) Medium M = .59 High M = .39	***Internal consistency:*** α= .820
Franklin et al. (2018)[43]	n=584; older adults; 65+; 38%; UK	***Construct:*** OLS model with EQ-5D-3L items as discrete variables, including age, sex and care home explanatory variables		

				produced best overall model: RMSE=.16; R²=.35	
		Milte et al. (2018)[44]	n=87; older adults following a hip fractur; 60+; 30%; Australia	***Convergent:*** Spearman Correlation EQ-5D-3L scores to (95% CI; p-values) Attachment: r =.27 (.07 – .43; .013) Security: r =.51 (.32 – .67; <.001) Role: r =.34 (.12 – .52; .002) Enjoyment: r =.26 (.03 – .46; .016) Control: r =.46 (.23 – .62; .000)	n.a.
6.	ICECAP-A	Al-Janabi et al. (2012)[16]	n=36; adults; 18+; 41%; UK	n.a.	n.a.
		Al-Janabi et al. (2013)[54]	n=34; adults; 47%; UK	***Face:*** Individuals largely responded to questions in intended manner and encountered problems on fewer than 10% of the items.	n.a.
		Al-Janabi et al. (2013)[39]	n=418; adults; M=51.7; 38%; UK	***Construct:*** Associations between EQ-5D and Stability χ^2 = <.001 (p <.01) Attachment χ^2 = .34 Autonomy χ^2 = <.001 (p <.01) Achievement χ^2 = <.001 (p <.01) Enjoyment χ^2 = <.001 (p <.01)	n.a.
		Al-Janabi et al. (2016)[69]	n=237; adults; 18+; 52%; UK	n.a.	***Test-retest***: Stability 89.8%; κ= .61 Attachment 88.8%; κ= .57 Autonomy 87.8%; κ= .52 Achievement 88.1%; κ= .53 Enjoyment 88.1%; κ= .54 ICC=.72
		Keeley et al. (2015)[37]	n=357; adults with knee pain; M=64; 49%; UK	***Construct:*** Correlation ICECAP-A to EQ-5D-3L r=.255 GAD-7 r=-.205 PHQ-8 r=-.190	***Responsiveness:*** Anchor-based analysis (baseline and 6-months follow-up) Mean ICECAP-A change (95% CI) EQ-5D-3L Improved .02 (.002–.042) (p <.05) EQ-5D-3L no change -.003 (-.128–.007) EQ-5D-3L worsened -.54 (-.084–-.024) (p <.01) GAD-7 Improved .020 (.002–.042) GAD-7 no change -.004 (-.003–-.011) GAD-7 worsened -.07 (-.11–-.032)

				(p <.01) PHQ-8 Improved .014 (-.005–.032) PHQ-8 no change .003 (-.006–.011) PHQ-8 worsened -.048 (-.078–-.017) (p <.01)
Goranitis et al (2016)[31]	n=478; women with irritative lower urinary tract syndrome; M=55; 0%; UK	***Convergent:*** EQ-5D correlated (p<.01) to Stability r= .38 Attachment r= .21 Autonomy r= .48 Achievement r= .45 Enjoyment r= .40 ICIQ-OAB correlated (p<.01) to Stability r= -.23 Attachment r= -.12 Autonomy r= -.19 Achievement r= -.21 Enjoyment r= -.25	***Discriminant*** ICECAP-A mean score (SD) Total impact of symptoms Low M=.86(.14) (p<.01) Moderate M=.87 (.13) High M=.81 (.18)	***Responsiveness:*** ICECAP-A Score change baseline to follow-up (SD) Symptoms' bother Increased bother -.05 (.15) (p <.01) Same bother -.03 (.17) Lower bother .00 (.15) Symptoms' frequency Improved .00 (.15) Same level -.039 (.13) Deteriorated .06(.18) (p <.01)
Goranitis et al (2016)[32]	n=83; adults with opiate dependence; M=37; 87%; UK	***Convergent:*** Correlation of ICECAP-A to Psychological health r=.55 (p <.01) Physical health r=.36 (p <.01) Quality of Life r=.55 (p <.01)	***Discriminant:*** ICECAP-A mean score; ±SD: Psychological health High M=.57; ±.19 (p <.01) Low M=.74; ±.15 Physical health status High M=.59; ±.20 (p <.01) Low M=.71; ±.17 Overall Quality of life High M=.58; ±.19 (p <.01) Low M=.75; ±.14	***Sensitivity:*** ICECAP-A mean change baseline to 3-months follow-up; ±SD: Psychological health Not improved M=.00; ±.19 Improved M=.08; ±.13 (p <.01) Physical health status Not improved M=.04; ±.19 Improved M=.05; ±.13 (p <.05) Overall Quality of life Not improved M=.02; ±.17 Improved M=.07; ±.15(p <.05)
Mitchell et al.	n=617; adults		***Discriminant:***	n.a.

		(2017)[40]	with depression; 18+; 33%; UK	ICECAP-A mean score to DASS-D Normal/well M=.84 Mild M=.71 Moderate M=0.71 Severe M=.64 Very severe M=.47	
		Linton et al. (2018)[36]	n=2,501; adults (healthy or with Arthritis, Asthma, Cancer, Depression, Diabetes, hearing problems, heart disease); 18+; 52%; Germany, UK	*Convergent:* Correlation of ICECAP-A and EQ-5D-5L Germany r= .62 UK r= .61 SWLS Germany r= .66 UK r= .68 SF-6D Germany r= .64 UK r= .65	*Internal-consistency:* (Cronbach's α) across subsamples Germany; UK Overall sample α=.83; .85 Healthy α=.78; .80 Arthritis α=.74; .78 Asthma α=.77; .83 Cancer α=.86; .83 Depression α=.78;.79 Diabetes α=.83;.86 Hearing loss α=.74;.84 Heart disease α=.83;.85
		Tang et al. (2018)[35]	n=975; adults; 18+/M=34; 47%; China	*Construct:* Two factor-analysis indicate a different construct between ICECAP-A and EQ-5D-3L Correlation of ICECAP-A and EQ-5D-3L Stability r= .39 (p<.01) Attachment r= .34 (p<.01) Autonomy r= .38 (p<.01) Achievement r= .27 (p<.01) Enjoyment r= .38 (p<.01)	*Internal-consistency*: Cronbach's α=.799
7.	ICECAP-SCM	Sutton & Coast (2014) [17]	n=23; older adults; 65+; n.a.; UK	n.a.	n.a.
8.	ICECAP-FC	Al-Janabi (2018) [18]	n=943; adults with long-term after-effects of meningitis; M=53; 25%; UK	n.a.	n.a.
9.	OCAP	Anand et al. (2009)[19]	n=1,048; adults; 18+;	n.a.	n.a.

			n.a.; UK			
10.	OCAP-18	Lorgelly et al. (2015)[20]	n=198 (qualitative), n=1,048 (quantitative); adults; M=46; 63%; UK	***Construct:*** Pairwise correlation with EQ-5D-3L =.576 (p<.001)		n.a.
11.	OxCAP-MH	Simon et al. (2013)[21]	n=333; adults with a mental illness; M=40; 67%; UK	***Convergent:*** Significant correlation of OxCAP-MH scores with GAF r=.25 EQ-5D VAS r=.51 EQ-5D-3L r=.41		n.a.
		Vergunst et al. (2017)[33]	n=172; adults with psychosis; M=38; 72%; UK	***Convergent:*** Correlation of OxCAP-MH with EQ-5D-3L r=.452 (p<.001) EQ-5D VAS r=.522 (p<.001) BPRS r=-.413 (p<.001) GAF r=.240 (p<.001) SIX r=.118		***Internal consistency*** Cronbach's α=.79 ***Test-retest (1-week apart):*** ICC= .86 (p<.001) Adjusted R²=.73 ***Sensitivity:*** Baseline (T1) M=67.7 (13.8) 12 months follow up (T2) M=70.8 (11.85) One-SEM values T1=6.47; T2=6.49
		Simon et al. (2018)[70]	n=10; adults with mental illness; M= 37; 40%; UK	n.a.		n.a.
		Laszewska et al. (2019)[34]	N=159; adults with mental illness; M=45; 36%; Austria	***Convergent:*** Correlation of OxCAP-MH change scores with EQ-5D-3L r=.30 (p<.05) EQ-5D VAS r=.31 (p<.05) BSI-18 r=-.42 (p<.05) GAF r=.15 (p<.05) Mini-ICF-APP r=-.10	***Discriminant:*** OxCAP-MH mean score (SD): Multi-morbidity one Axis diagnosis M=68.2(14.4) ≥2 Axis diagnoses M=56.0(16.8) (p<.001) Rating of QoL Very poor/poor M=48.0(15.4) Neither poor or good M=65.3(11.5) (p<.001) Good/very good M=74.3(11.2) (p<.001)	***Test-retest*** (after 30 days) Cronbach's α=.85 ICC=.80 (95%CI .69-.87)
12.	CQ-CMH	Sacchetto et al. (2016)[22]	N=50; adults with mental illness; M=42; 70%; Portugal	***Face:*** 15 participants confirmed familiarity with language used and relevance of addressed issues. Questionnaire rated as understandable and easy to fill out but too extensive.		n.a.

13.	ACQ-CMH-98	Sacchetto et al. (2018)[23]	n=332; adults with mental illness; M=44; 59%; Portugal	***Content:*** Participants (n=15) CVI: .89 *Convergent:* Pearson's correlation with WHOQOL-Bref r=.60 (p<.001) K6 r=.46 (p<.001)	***Discriminant:*** Pearson's correlation with RAS r=-.17 (p=.046)	***Test-retest:*** 55% of items high (r=.9 to ≥6) 45% of items low (r= <6) ANOVA test significant for 5 items (p<.05) ***Internal consistency:*** Optimism α=.91 Affiliation α=.84 Activism α=.84 Practical Reason α=.76 Self-sufficiency and Self-determination α=.76 Family α=.78
14.	Capability-based questionnaire	Kinghorn et al. (2015)[24]	n=16; adults with chronic pain; 33+; 43%; UK	n.a.		n.a.
15.	CADA	Ferrer et al. (2014)[25]	n=109; adults with obesity and diabetes mellitus; M=49; 22%; USA	n.a.		***Internal-consistency***: Convenience, cost: α=.78 Neighborhood opportunity: α=.78 Barriers: α=.75 Knowledge: α=.83 Time Pressure: α=.75 Family support: α=.62 Spouse/partner: α=.65 Nonfamily support: α=.80

BPRS=Brief Psychiatric Rating Scale; CTM-3=3-Item Care Transition Measure; DASS-D= Depression Anxiety Stress Scales; GAF=Global Assessment of Functioning; ICC=Intra-class correlation coefficient; K6= Kessler Psychological Distress Scale; M=Mean; OLS=ordinary least square; PHQ-8= Patient Health Questionnaire depression scale; RAS=Recovery Assessment Scale; SEM= Structural equation modeling; SIX=Objective Social Outcomes Index

Figures

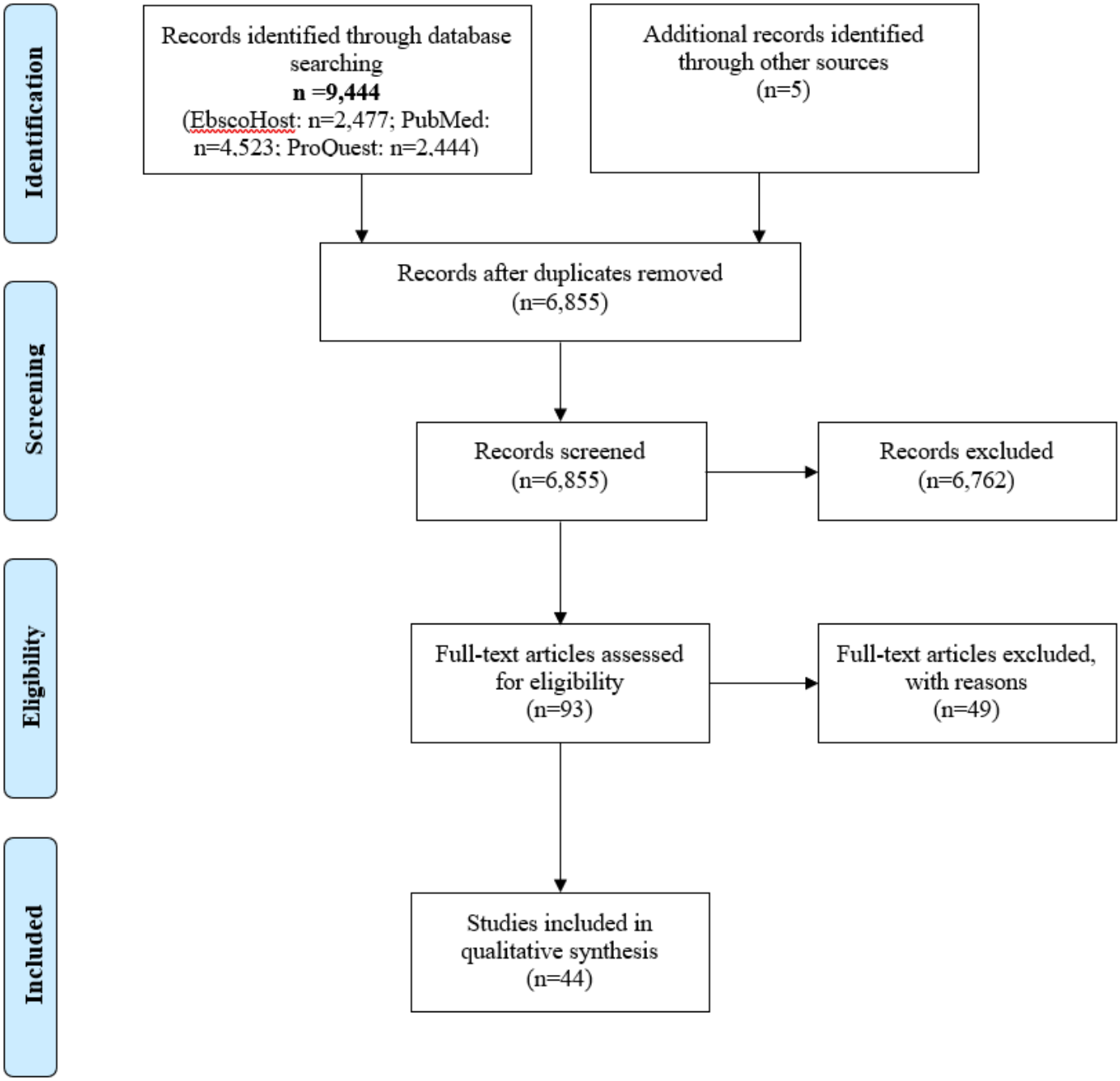

Figure 1

Literature search flow chart based on PRISMA [13]

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix.docx