The use of proxies and proxy-reported measures: a report of the international society for quality of life research (ISOQOL) proxy task force

Roydhouse, Jessica K.; Cohen, Matthew L.; Eshoj, Henrik R.; Corsini, Nadia; Yucel, Emre; Rutherford, Claudia; Wac, Katarzyna; Berrocal, Allan; Lanzi, Alyssa; Nowinski, Cindy; Roberts, Natasha; Kassianos, Angelos P.; Sebille, Veronique; King, Madeleine T. [**and 1 more**]

How to cite

ROYDHOUSE, Jessica K. et al. The use of proxies and proxy-reported measures: a report of the international society for quality of life research (ISOQOL) proxy task force. In: Quality of life research, 2021. doi: 10.1007/s11136-021-02937-8

This publication URL: https://archive-ouverte.unige.ch/unige:158548

Publication DOI: 10.1007/s11136-021-02937-8

Quality of Life Research
https://doi.org/10.1007/s11136-021-02937-8

REVIEW

The use of proxies and proxy-reported measures: a report of the international society for quality of life research (ISOQOL) proxy task force

Jessica K. Roydhouse[1,2] · Matthew L. Cohen[3] · Henrik R. Eshoj[4] · Nadia Corsini[5] · Emre Yucel[6,7] · Claudia Rutherford[8,9] · Katarzyna Wac[10,11] · Allan Berrocal[10] · Alyssa Lanzi[3] · Cindy Nowinski[12] · Natasha Roberts[13,14] · Angelos P. Kassianos[15] · Veronique Sebille[16,17] · Madeleine T. King[9] · Rebecca Mercieca-Bebber[18,19] on behalf of the ISOQOL Proxy Task Force and the ISOQOL Board of Directors

Accepted: 3 July 2021

Abstract

Aims Proxy reports are often used when patients are unable to self-report. It is unclear how proxy measures are currently in use in adult health care and research settings. We aimed to describe how proxy reports are used in these settings, including the use of measures developed specifically for proxy reporting in adult health populations.

Methods We systematically searched Medline, PsycINFO, PsycTESTS, CINAHL and EMBASE from database inception to February 2018. Search terms included a combination of terms for quality of life and health outcomes, proxy-reporters, and health condition terms. The data extracted included clinical context, the name of the proxy measure(s) used and other descriptive data. We determined whether the measures were developed specifically for proxy use or were existing measures adapted for proxy use.

Results The database search identified 17,677 possible articles, from which 14,098 abstracts were reviewed. Of these, 11,763 were excluded and 2335 articles were reviewed in full, with 880 included for data extraction. The most common clinical settings were dementia (30%), geriatrics (15%) and cancer (13%). A majority of articles (51%) were paired studies with proxy and patient responses for the same person on the same measure. Most paired studies (77%) were concordance studies comparing patient and proxy responses on these measures.

Discussion Most published research using proxies has focused on proxy-patient concordance. Relatively few measures used in research with proxies were specifically developed for proxy use. Future work is needed to examine the performance of measures specifically developed for proxies.

Systematic review registration PROSPERO No. CRD42018103179

Keywords Proxy measures · Proxy-reported outcomes · Outcome measures · Quality of life · Systematic review

Plain English summary

Questionnaires are often used to measure patient perceptions of their health and experiences. If patients cannot complete these questionnaires, other individuals ("proxies") may be asked to complete the questionnaires for them. These questionnaires may be designed for patient completion and then adapted for completion by proxies, or they may be designed specifically for proxies. Following a detailed literature search, we reviewed more than 800 articles to address the question of how questionnaires for proxies have been used. We found that questionnaires for proxies were most commonly used in the fields of dementia, geriatrics and cancer. We also found that few questionnaires were designed specifically for proxies. Most of the studies we reviewed involved comparing results from questionnaires completed by patients reporting their own health and proxies answering adapted versions of the same questionnaires about patients' health, to look at how similar proxy and patient

The present institution is Bristol Myers Squibb, however this work was conducted while based at Amgen.

✉ Jessica K. Roydhouse
jessica.roydhouse@utas.edu.au

Extended author information available on the last page of the article

Published online: 12 July 2021

responses were. Finally, we found that the term "proxy" was used inconsistently across studies, often referring to different people (e.g., a clinician, family member, or friend). Helpful areas to look at in the future include a clearer definition of what it means to be a "proxy," and encouraging consistent use of terms and processes for measuring proxy-reported outcomes.

Introduction

Patient-centred outcomes are increasingly important in research and clinical settings. A major challenge for the assessment of patient-centred health outcomes is how to assess them for individuals who are unable to reliably self-report their outcomes [1]. Examples include cognitive or linguistic impairment that inhibits comprehension of items, self-awareness or self-expression, or symptom burden and clinical deterioration in terminal illness [2]. Some aspects of patient health may be assessable via clinician observation or performance-based measurement, but others require self-report [3]. The latter include symptom experience, emotional wellbeing and quality of life. These aspects of health outcomes are challenging to assess in these individuals, but nevertheless remain very important in research and clinical settings.

Proxy-reported outcomes (ProxRO) provide a means of capturing such data from patients who cannot self-report, and have been used in research and surveys to avoid what would be otherwise missing data. A proxy is a person who reports an outcome on behalf of a patient [4, 5]. Typically the proxy is a family caregiver, but health care professionals (HCPs) may also act as proxies [2]. If clinical and professional judgment form part of the rating, such outcomes may be considered clinician-reported outcomes (ClinROs) [6]. ProxROs should also be differentiated from observer-reported outcomes (ObsROs), which are "limited to the assessment of observable signs and symptoms that can be reported from the perspective of a parent or caregiver" [7] (p. 17) per the Food and Drug Administration (FDA). In contrast, the FDA defines ProxRO instruments as proxies reporting "as if he or she were the patient" [7] (p. 17). Similarly, the European Medicines Agency (EMA) notes that a proxy "is a person who reports an outcome as if she/he was the patient him/herself" and defines ObsROs as "based on an observation by someone other than the patient or a health professional", noting that ObsRO reports "include only events or behaviours that can be observed [5] (p. 11, 12)." Importantly, in these situations self-report is preferred where possible and should not be discounted. The challenge of shifting to proxy report where and if needed is a prominent one in dementia research [8].

Although definitions of proxy reporting emphasise the importance of the proxy taking on the patient's perspective, other perspectives are also used [9], including asking the proxy to report from their perspective rather than taking the patient's perspective, and the perspective sought when asking proxies to complete instruments is not always reported in the literature [9]. Matza and colleagues similarly distinguish between "observational measures" that focus on observable and observed behaviours, such as crying, where the observer does not make any judgment or interpretation, and "proxy measures," where interpretation is involved [10]. Although there are differences between the definitions, the aspects of health that requires judgment or interpretation appears to be a delineation between these types of measures. Additionally, in some cases the term "informant report" may be used to refer to proxy reports. Informants are often asked to report on symptoms and/or behaviour [11, 12]; one definition of informant uses the term interchangeably with proxy [13].

The frequency with which ProxROs are used in studies in adult health populations reflects health conditions. For example, a recent review showed that while only 3% of trials registered with the Australian New Zealand Clinical Trials Registry (ANZCTR) from 2005-March 2017 included proxy-reported endpoints, that increased to 11, 10 and 8% of registered mental health, stroke and neurological trials, respectively [14]. Proxy-reported data play an important role in palliative care research [15] and proxies have been utilized in numerous health outcome and care experience surveys [16–19].

Several previous studies have evaluated discrepancies between proxy- and self-report for pairs of individuals [20–22]. In these studies, the instruments used were typically developed for self-report; for example, Pickard and colleagues compared patient and proxy responses on the EQ-5D, a generic patient-reported outcome (PRO) measure (PROM); the five EQ-5D items, initially developed and evaluated in a patient population, were adapted for proxy completion [23]. However, in situations where proxy use is frequent, for example due to cognitive deficit or symptom burden in the target population, it may be more appropriate to use measures specifically developed for proxies. To our knowledge, there are no comprehensive reviews of proxy-reported measures for adults or their use in research and clinical practice. The international society for quality of life research (ISOQOL) proxy task force's review had the following aims: (1) descriptive summary of proxy-reported measures used in studies; (2) summary of how measures used in studies with proxies have been developed; and (3) review of proxy-specific measures against COSMIN criteria. In this paper, due to the volume of articles identified, we focus on the first aim. Other aims will be addressed in subsequent papers. In particular, in this paper we present a summary of how proxy measures are used in research and

categorise the types of measures used by proxies in adult health populations.

Methods

The methodology for this systematic review complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement [24]. The protocol is registered with the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42018103179).

Search strategy

A systematic search of Medline (OVID), PsycINFO, PsycTESTS, CINAHL and EMBASE from database inception until 22 February 2018 was conducted, based on search terms developed in collaboration with an academic librarian (Appendix 1: Search strategy). The search included terms for quality of life, health outcomes, proxy-reporters, and health conditions. Duplicate retrieved records were identified and deleted.

Eligibility criteria

Original research articles that described the use and/or development of a proxy-reported measure of adult patient health outcomes (e.g. symptoms, health-related quality of life), experience, health behaviours or health service usage were eligible for inclusion. Patients could have any health condition. Proxies could be informal caregivers (e.g. family members, friends) or health care professionals. The protocol stated that articles in any language were eligible, however in practice we limited the eligibility criteria to articles written in Danish, English, French, Greek, Italian, Norwegian, Polish, Spanish or Swedish, as native or fluent speakers of these languages were involved in our research team.

We excluded: studies of proxy measures in paediatric contexts (age < 18 years), systematic reviews, opinion pieces, dissertations, conference abstracts, articles evaluating hypothetical health states (i.e. lay people or individuals without a specific health condition ranking or valuing health or disease states), qualitative studies that did not address any aspect of proxy measurement development, studies of caregivers' health and/or experience, articles describing proxies as medical decision makers on behalf of patients (e.g. whether to proceed with surgery, begin treatment, turn off life support, divide personal estate, give advanced directives, provide informed consent, etc.), proxy reports being used to classify or diagnose patients (e.g. reporting if a patient did or did not have a specific health condition) and studies whose focus was on caregiver outcomes.

Abstract/title screening procedure

We conducted two pilot training exercises in which each reviewer screened 100 titles and abstracts per exercise. Within each exercise, two independent reviewers assessed each title and abstract independently, and a third reviewer assessed any discrepancies. Following the pilot, reviewers were allocated an equal number of abstracts to screen against eligibility criteria. Title and abstract screening were performed in Microsoft Excel spreadsheets, pre-populated with key details of the article (Authors, Title, Journal citation, Abstract). Reviewers indicated whether the article should proceed to the next stage of screening ("definitely eligible" or "likely to be eligible") or if the article should be "excluded", including the reason for exclusion from a drop-down list. Reviewers recorded the broad clinical area of the article from a pre-defined drop-down list for articles that are "definitely" or "possibly" relevant to the review and recorded the language of any non-English eligible articles.

Full text screening

Articles were divided among reviewers. Each reviewer obtained the articles in full text and screened them against eligibility criteria. Review decisions and key data were recorded in an Excel spreadsheet. When articles were excluded, reviewers selected an exclusion reason from a drop-down list.

Full text data extraction

Key information about each article was pre-populated in the Excel spreadsheet (Authors, Title, Journal citation, Abstract, Broad clinical area). Reviewers were required to extract data on the following areas: (1) proxy measure title (i.e. title of measure[s] used by proxy in the study), (2) proxy measure acronym, (3) article type (development paper for proxy measure, application or use of proxy measure), (4) development paper reference for the proxy measure, (5) study population, (6) study design (e.g. RCT), (7) proxy-patient relationship (e.g. family member), (8) reason for proxy use, (9) how proxies were used (e.g. paired study with responses for the same individual), (10) the % of patient and proxy participants, for unpaired studies. The first author reviewed ~ 10% of the extraction decisions of the other reviewers.

Classification of identified measures

Each identified measure (i.e. questionnaire) completed by proxies was entered into an Excel spreadsheet and given a unique identifying number. For each measure, the paper that reported its development was sought, and where available, was reviewed to determine the original population in which

it was developed and evaluated. If a development paper was not cited (i.e., a study-specific or ad-hoc measure), then the type of population in the article using the measure was used for classification. For example, a study-specific or ad-hoc questionnaire without a cited development paper used in a case–control study that interviewed the next of kin of decedents to evaluate health behavior would be classified as a proxy-specific measure. If a development paper was cited but could not be obtained, or the study population did not allow for the determination we sought (as described above), then the measures were classified as 'unclear.' We also considered measures that were self-described proxy measures, but then classified these based on the original population in which the measure had been developed and evaluated.

For proxy-specific measures (ProxROM), where possible we further classified them by their context of use, e.g., post-death/bereavement measures, parallel measures (e.g., a PROM and ProxROM were developed at the same time), or informant measures.

Data synthesis

In this paper, the primary level of analysis is study unless otherwise specified. We summarise how proxy measures were used in research (e.g., clinical context of study, study design, etc.). The number of studies in which each measure was used was tallied, both overall and by health condition. We also summarised the papers using proxy-reported measures descriptively (e.g., health conditions: cancer, Parkinson's disease, etc.; study designs: RCT, cohort, cross-sectional) as per the data extraction fields noted above. Finally, we described the classification of the measures identified or described as developed specifically for proxy use.

Results

The search identified 17,677 possible articles, from which 14,098 abstracts and titles were screened. Of these, 2335 articles were eligible and reviewed in full, of which 880 were included for data extraction (Fig. 1).

Clinical areas

The most common clinical area in which studies were conducted was dementia (all cause, but mostly from Alzheimer's disease or related disorders) (264/880, 30%), followed by geriatrics (130/880, 15%) and cancer (116/880, 13%) (Table 1). Geriatrics included, for example, centenarian studies where the focus was not necessarily on patients with a dementia diagnosis; however, it is likely that there was some overlap. There were 64/880 studies (7%) categorised as "other neurological conditions." The most common sub-areas for other neurological conditions included brain injury (22/64, 34%), amyotrophic lateral sclerosis (ALS; 9/34, 14%), post-stroke aphasia (7/34, 11%) and epilepsy (7/34, 11%). For these other neurological conditions, participants either did not have dementia or it was not specified.

Types of proxies

In a majority of the included articles (476/880, 54%), the proxies were the patient's family member. Only a minority of articles (67/880, 8%) had a health care professional as the proxy and very few (12/880, 1%) had another caregiver (i.e., not a family member or health care professional) as the proxy. Nearly one fifth of articles (169/880, 19%) did not report the type of proxy-patient relationship, and some articles included multiple types of proxies (116/880, 13%) (Table 2).

Study designs

The most common type of study design was a cross-sectional assessment (404/880, 45.9% of articles), followed by development/validation studies (145/880, 16%) and longitudinal studies (139/880, 16%) (Table 3).

Study use of proxies

As Table 3 shows, most articles (452/880, 51%) had both proxy and patient reports for the same patient on the same questionnaire (for example, both completed the EQ-5D, albeit with different wording). A smaller proportion of studies (66/880, 8%) had both proxy and patient responses for the same person but used different questionnaires (i.e., the proxy completed a specific questionnaire about the patient, and the patient completed a different questionnaire about the patient). A very small proportion of studies had both proxy and patient reports for the same person at one point in time, and then only a proxy report for those patients who could not self-report (2/880, < 1%). Only a few studies (129/880, 15%) had some patients who self-reported and proxies for other patients who did not self-report (Table 3). For studies that had both patient and proxies reporting for the same person on the same measure, 346/452, 77% were concordance studies comparing patient and proxy reports on these measures.

There were differences across the three most common clinical areas (dementia, geriatrics, cancer) in terms of how proxies were used (Table 4). In cancer, over half of the articles (61/116, 53% of articles in cancer) had both proxy and patient reports for the same individual on the same measure. For dementia and geriatrics, this was < 50% (116/264, 44% and 52/130, 40%, respectively). Having some patients self-report and proxies report for other patients was common in geriatrics (44/130, 34%), but less

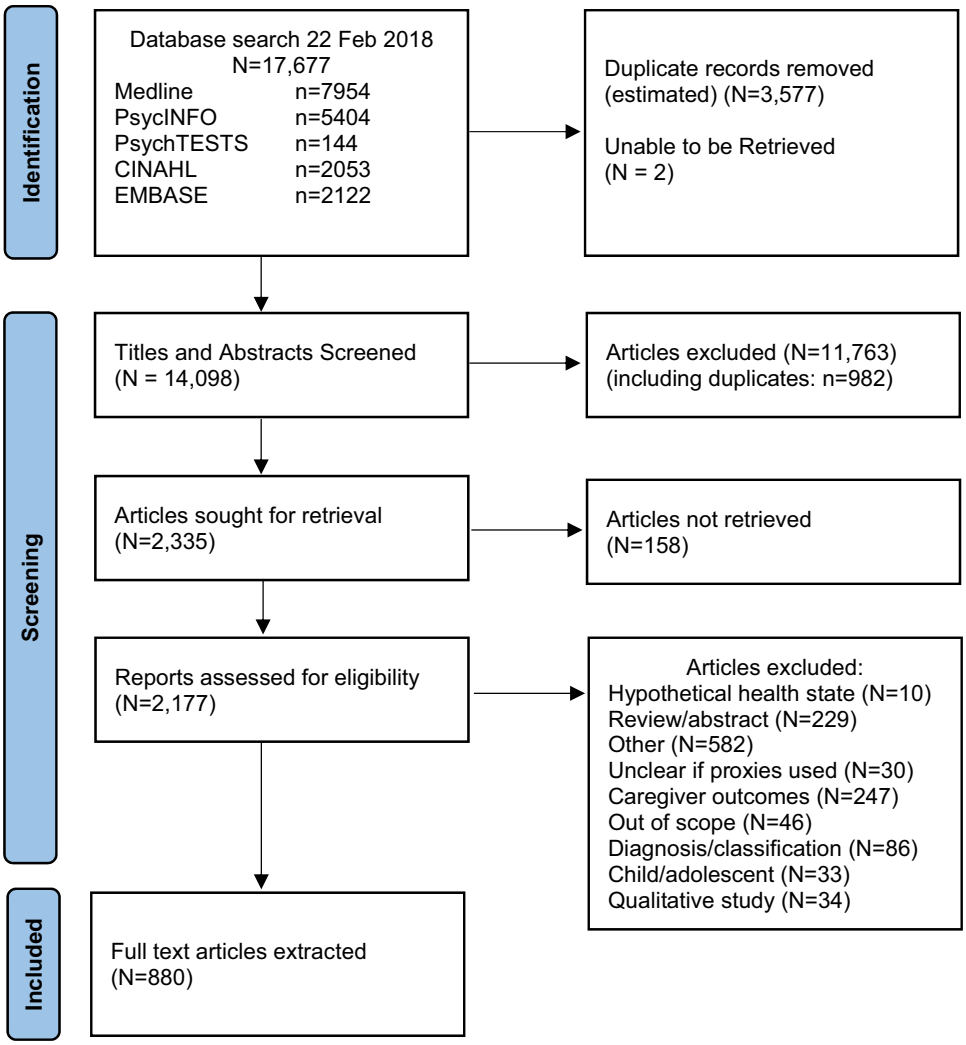

Fig. 1 PRISMA diagram

Table 1 Primary clinical areas of articles using proxies ($n=880$)

Area	*N* (%) articles
AIDS/HIV	5/880 (1)
Cancer	116/880 (13)
Critical care	18/880 (2)
Dementia (all cause, but almost always Alzheimer's disease or related disorder)	264/880 (30)
General population	9/880 (1)
Geriatrics (without dementia or not specified)	130/880 (15)
Intellectual disabilities	33/880 (4)
Multiple sclerosis	12/880 (1)
Other	107/880 (12)
Other neurological conditions (e.g. ALS) (without dementia or not specified)	64/880 (7)
Palliative care	37/880 (4)
Parkinson's (without dementia or not specified)	13/880 (1)
Physical injury/disability	22/880 (3)
Stroke (without dementia or not specified)	50/880 (6)

Table 2 Proxies' relationship to patients in the included articles ($N=880$)

Type of use	*N* (%) articles
Family member	476 (54)
Health care professional	67 (8)
Multiple proxy types	116 (13)
Not stated	169 (19)
Other	40 (5)
Other caregiver	12 (1)

common in dementia (10/264, 4%) and cancer (18/116, 16%). Relying only on proxy reports was most common in dementia (84/264, 32%), followed by cancer (32/116, 28%); this was less common in geriatrics (16/130, 12%). Having proxies and patients report for the same individual but on different measures was much more common in dementia (41/264, 16%) compared to geriatrics (7/130, 5%) or cancer (2/116, 2%).

Table 3 Study use of proxies and study designs of articles using proxies ($N = 880$)

Aspect	*N* (%) articles
Type of use	
Paired study with proxy+patient reports for the same patient using the same questionnaire(s)[a]	452 (51)
Paired study with proxy+patient reports for the same patient using different questionnaire(s)[b]	66 (8)
Proxy and patient report for the same patient at one point in time, then only the proxy reports when the patient can't self-report	2 (< 1)
Proxies report for patients who cannot self-report and patients who can self-report do	129 (15)
Only proxy reports (no patient self-reports)	192 (22)
Other	39 (4)
Study design	
National/international health or experience survey	57 (6)
Randomised controlled trial of an intervention	31 (4)
Non-randomised controlled study of an intervention	17 (2)
Cross-sectional assessment	404 (46)
Longitudinal assessment	139 (16)
Development/validation study	145 (16)
Case–control study	48 (5)
Other	39 (4)

[a]This includes questionnaires designed for self-report that may have had their wording changed for proxy completion (e.g. EQ-5D) and questionnaires with different patient/proxy versions (e.g. DEMQOL and DEMQOL-Proxy)

[b]This includes self-report questionnaires for the patient and a different questionnaire for the proxy that is not a reworded or different version of the patient's questionnaire

Table 4 Study use of proxies: comparison across three most common clinical areas

Proxy use in studies	Dementia $N = 264$ articles (%)	Geriatrics $N = 130$ articles (%)	Cancer $N = 116$ articles (%)
Proxy and patient reports for the same patient using the same questionnaire(s)[a]	116 (44)	52 (40)	61 (53)
Proxy and patient reports for the same patient using different measure(s)[b]	41 (16)	7 (5)	2 (2)
Proxy and patient report for the same patient at one point in time, then only the proxy reports when the patient can't self-report	0 (0)	1 (1)	0 (0)
Proxies report for patients who cannot self-report and patients who can self-report do	10 (4)	44 (34)	18 (16)
Only proxy report (no patient self-report)	84 (32)	16 (12)	32 (28)
Other	13 (5)	10 (8)	3 (3)

[a]This includes questionnaires designed for self-report that may have had their wording changed for proxy completion (e.g. EQ-5D) and questionnaires with different patient/proxy versions (e.g. DEMQOL and DEMQOL-Proxy)

[b]This includes self-report questionnaires for the patient and a different questionnaire for the proxy that is not a reworded or different version of the patient's questionnaire

Categories of measures

We identified 527 measures used in the 880 included studies. The most common measure type used was one originally designed for patient self-report (243/527, 46%). Of the 527 measures, 177 (34%) were described as 'proxy' measures or could be classified as proxy measures based on the development paper (if available) or the paper that used these measures (if a development paper was not available). As Table 5 shows, 53/177 (30%) were ad hoc, study-specific or unclear in terms of design. Furthermore, 16/177 (9%) appeared to be miscategorised, that is the authors of the included studies called them 'proxy measures', but when checked, the development papers did not corroborate this. This included a small group of measures designed for clinician completion which could potentially be classified as ClinROs (7/177, 4%). Additionally, several measures were described in included studies as proxy measures, but we determined that they were adaptations of PRO measures for proxy use (8/177, 5%).

Nonetheless, most of the measures described as 'proxy' measures (107/177, 60%) were designed originally for proxy

Table 5 Use of proxy-specific measures ($N = 177$ measures used in $N = 452$ articles)

Measure category	*N* measures (% of 177 proxy measures identified)	*N* articles (% of 452 articles that used proxy-specific measures)
Ad hoc, study-specific or unclear	53 (30)	69 (15)
Not designed originally for proxy completion	16 (9)	22 (5)
PRO[a] with some adaptation for proxy	8 (5)	11 (2)
ClinRO[b]	7 (4)	9 (2)
ObsRO[c]	1 (1)	2 (< 1)
Designed originally for proxy completion	107 (60)	360 (80)
ProxRO[d]: Informant (including informant measures with a parallel component, i.e. informant and patient versions)	41 (23)	147 (33)
ProxRO[d]: Parallel[e]	36 (20)	153 (34)
ProxRO[d]: Bereavement/post-death	14 (8)	27 (6)
ProxRO[d]: Non-bereavement	16 (9)	33 (7)
Combined measure (designed for both patient and proxy to complete together)	1(1)	1 (< 1)

[a]PRO described as proxy measure but is an adapted PRO measure with wording changed

[b]ClinRO designed for clinicians to complete

[c]ObsRO described as proxy measure but focuses only on observable behaviours

[d]ProxRO designed for non-clinician proxy to complete

[e]Parallel self- and proxy-versions for an instrument

completion. This included informant measures (41/177, 23%) and parallel (i.e., both self- and proxy-report versions for a measure) measures (36/177, 20%). Measures for proxy completion that were not parallel or informant measures but did not pertain to bereavement were relatively infrequent as a category (16/177, 9%).

These 177 'proxy' measures were included in 452 articles. Although 15% (69/452) articles used ad hoc, study-specific or unclear measures, most articles (360/452, 80%) used measures designed originally for proxy completion. The most common measure categories used were informant (147/452, 33%) and parallel (153/452, 34%).

Discussion

A large number of studies (> 800) that included proxy-reported measures were identified; the majority of these studies had patient and proxy responses for the same individual patient. Furthermore, most such studies focussed on patient-proxy concordance. Clinical areas in which proxies were commonly used in studies included dementia, geriatrics and cancer. The most common type of measure used in studies was designed for patient self-report, i.e., a PROM, and less than half of the measures used were developed specifically for proxy use or described as such. Of the measures described in included studies as being for proxies, several were arguably better classified as ClinROs as they were designed for clinician completion, which may reflect that the latter term has entered the health outcomes research lexicon more recently and may be unfamiliar to some researchers.

Previous reviews in this area have tended to focus on the issue of proxy-patient discrepancy, primarily by using or reviewing data from paired studies [25–31]. In general, these studies have found better concordance for more observable domains of health (e.g., physical function) compared to less observable domains (e.g., emotional function) [25, 31]. Although concordance studies can provide valuable information, to date there has been limited advice regarding when or if to switch from patient to proxy report if both reports are collected. Additionally, assessment of concordance alone is likely insufficient for measure evaluation; other aspects, such as psychometric properties, are also likely to be an important consideration.

Furthermore, several of the reviews of concordance studies have focussed on proxy-patient discrepancy in cancer [25, 29, 31]; interestingly, our review identified that measures developed specifically for proxies were more commonly used in dementia, rather than in cancer. It is likely that the instruments evaluated in these studies are PRO instruments that have been adapted for proxy use by rewording, e.g., changing 'I feel' to 'The patient feels.' Evaluating discrepancies between patients and proxies can contribute to our understanding of the validity of substituting proxies for patients if patients are unable to self-report. Importantly, however, the use of measures in a way in which they were not originally designed, without adequate psychometric performance assessment, may not necessarily be appropriate. This issue is not limited to the adaptation of PROMs for

proxy report; for example, Liebzeit and colleagues found that instruments frequently used to measure functional status in elderly patients transitioning between care settings were not being administered as originally designed and tested [32]. This suggests potential areas for future research regarding proxies, specifically a closer assessment of proxy-specific measures, or if patient-reported measures are to be adapted then a careful validation and evaluation of their psychometric properties in a new population should be undertaken.

Additionally, the findings of this review suggest potential for misclassification of proxy measures relative to clinician measures. In a small number of cases, several proxy-specific measures were evaluated only with clinician respondents. It is not completely clear if these measures are appropriately classified as 'proxy' since clinicians may rely on their clinical judgment when evaluating aspects of patient health and behaviour. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Clinical Outcomes Assessment Emerging Good Practices Task Force defines a ClinRO assessment as one in which an individual uses their professional training in making a judgment [6]. Similarly, the Montreal Accord on Patient-Reported Outcomes defines ClinROs as clinical judgment using professional training following observation of an individual [33]. The Accord differentiates ClinROs from ObsROs and ProxROs. Although both involve observation, the Accord defines proxies as a "special kind of observer" whose "shared experience" allows them to report on patient health and behaviour [33] (p. 122). It is unclear how best to categorise proxy-specific measures designed for clinician completion. It seems unlikely that a clinician would not use their professional judgment when making an evaluation about patient health or the presence of symptoms. Further clarification of this issue would be helpful for future research. In addition, clarification regarding the perspective taken when respondents are completing these measures [9] may be helpful in further elucidating this issue.

Another issue pertains to differentiating ObsROs and ProxROs. As noted previously, both the EMA and FDA definitions discuss perspective-taking when referring to ProxROs. The International Society for Quality of Life Research (ISOQOL) Translation and Cultural Adaptation Special Interest Group (TCA-SIG) cited the FDA's definition and differentiated ProxROs from ObsROs by this perspective-taking [34]. Furthermore, they emphasize that ObsROs are limited to fully observable behaviours and/or events. The EMA definition clearly differentiates ObsROs from ClinROs by specifying that health professionals cannot be observers who provide ObsROs [5]. The ISOQOL dictionary definition for ObsROs also discusses who can be an observer, stating that observers are "[people], not necessarily with any expert training" who report on observable behaviours and not feelings or emotions [35] (p. 97). In contrast, Cappelleri et al. include clinicians as possible observers for ObsROs, but differentiate ProxROs and ObsROs by the proxy's perspective-taking and the addition of the proxy's interpretation or judgment to the observation [36]. It is not completely clear how best to classify measures that may lack a specific perspective, or use one other than "reporting as patient," but report on events and behaviours that are not fully observable. Greater clarification of these issues will be beneficial for future research.

An additional issue that would benefit from greater research and clarity is the question of how best to develop proxy measures going forward. This review identified relatively few measures developed specifically for proxy use, and several papers evaluating PROs adapted for proxy use. Adapted measures may be preferable in some cases given the goal of a proxy report substituting for an otherwise unavailable patient report; however, such adaptation will likely need to go beyond simple rewording and at a minimum clear instructions will likely be required [37]. Recently, the ISOQOL TCA-SIG developed good practices and process recommendations for translation and cultural adaptation of non-PRO clinical outcome assessments [34]. Developing such recommendations for proxy measures may be useful.

Study limitations

Several limitations should be noted. First, we did not include articles in languages other than the nine spoken by members of our team; there may have been relevant studies in other languages. However, prior evidence suggests exclusion of non-English studies is unlikely to affect results [38]. Second, the search was conducted in 2018, and relevant papers may have been published since. Nevertheless, a very large number of articles were reviewed and extracted and repeated reviews of this breadth would be challenging.

Since our search in 2018, there have been several new studies focused more specifically on proxies and proxy reporting. This includes a qualitative study as part of the development of a proxy version of ASCOT [37], and a comparison of responses on the ICECAP-A for multiple proxy raters without a patient rater [39]. There were also analyses of proxy reporting for patient care experience [40], quality of life [41] and shared decision-making [42]. This suggests some work in new areas in the field of proxy reporting, but given the extensive body of evidence covered by our review, we feel the description of the state of the field based on the initial search remains accurate.

We coded clinical area (e.g., dementia, cancer) according to how it was described in each publication. However, this resulted in some overlapping classifications. For example, "dementia" describes a state at which an individual has lost the ability to live independently because of cognitive decline [43]. Alzheimer's disease is the most common cause of dementia in older adults [44], and in clinical settings most

individuals who are diagnosed with Alzheimer's disease present for care because it has resulted in dementia. However, dementia may also be caused by stroke, Parkinson's disease, ALS, traumatic brain injury, or other conditions. For example, about 30% of people with Parkinson's disease have dementia because of it [45]. For this review, we coded the clinical area in line with how the authors described their sample but acknowledge that the cause of dementia was not always clear, and the presence of dementia in cognitively mixed samples like Parkinson's disease or geriatrics was generally underspecified.

Conclusion

In conclusion, although there are numerous studies using proxies, these articles infrequently use measures specifically developed for proxy report. Furthermore, the term "proxy measure" is applied to a diversity of measures, including measures designed for patient self-report that have been adapted for proxy-report, ad hoc measures and measures specifically developed for proxy report. Most studies involving proxies tend to focus on proxy-patient concordance. Future work examining the performance of measures specifically developed for proxies may help advance the field.

This paper was reviewed and endorsed by the International Society for Quality of Life Research (ISOQOL) Board of Directors as an ISOQOL publication and does not reflect an endorsement of the ISOQOL membership.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11136-021-02937-8.

Data availability Details of the included studies can be found on Figshare: https://doi.org/10.6084/m9.figshare.14888907.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Rand, S., Caiels, J. (2015). Using proxies to assess quality of life: A review of the issues and challenges. Discussion paper. Quality and outcomes of person-centred care policy research unit (QORU). University of Kent. Retrieved December 1, 2020, from https://kar.kent.ac.uk/55009/1/Proxy%20literature%20review%20report%20FINAL.pdf
2. Jones, J. M., McPherson, C. J., Zimmermann, C., Rodin, G., Le, L. W., & Cohen, S. R. (2011). Assessing agreement between terminally ill cancer patients' reports of their quality of life and family caregiver and palliative care physician proxy ratings. *Journal of Pain and Symptom Management, 42*(3), 354–365.
3. Cohen, M. L., & Hula, W. D. (2020). Patient-reported outcomes and evidence-based practice in speech-language pathology. *American Journal of Speech-Language Pathology, 29*(1), 357–370.
4. Food and Drug Administration. (2009). Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims. Silver Spring, MD: Food and Drug Administration. Retrieved December 1, 2020, from https://www.fda.gov/media/77832/download
5. European Medicines Agency. (2016). Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man: The use of patient-reported outcome (PRO) measures in oncology studies. London, UK: European Medicines Agency. Retrieve December 1, 2020, from https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf
6. Powers, J. H., 3rd., Patrick, D. L., Walton, M. K., Marquis, P., Cano, S., Hobart, J., Isaac, M., Vamvakas, S., Slagle, A., Molsen, E., & Burke, L. B. (2017). Clinician-reported outcome assessments of treatment benefit: Report of the ISPOR clinical outcome assessment emerging good practices task force. *Value in Health, 20*(1), 2–14.
7. Food and Drug Administration. (2018). Attachment to discussion document for patient-focused drug development public workshop on guidance 3: Select, develop or modify fit-for-purpose clinical outcome assessments—Appendices 2018. Silver Spring, MD: Food and Drug Administration. Retrieved August 4, 2020, from https://www.fda.gov/media/116281/download
8. Hanson, L. C., Bennett, A. V., Jonsson, M., Kelley, A., Ritchie, C., Saliba, D., Teno, J., & Zimmerman, S. (2020). Selecting outcomes to ensure pragmatic trials are relevant to people living with dementia. *Jounal of American Geriatrics Society, 68*(Suppl 2), S55–S61.
9. Pickard, A. S., & Knight, S. J. (2005). Proxy evaluation of health-related quality of life: A conceptual framework for understanding multiple proxy perspectives. *Medical Care, 43*(5), 493–499.
10. Matza, L. S., Patrick, D. L., Riley, A. W., Alexander, J. J., Rajmil, L., Pleil, A. M., & Bullinger, M. (2013). Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value in Health, 16*(4), 461–479.
11. De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology, 9*, 123–149.
12. Grill, J. D., Zhou, Y., Karlawish, J., & Elashoff, D. (2015). Frequency and impact of informant replacement in Alzheimer disease research. *Alzheimer Disease and Associated Disorders, 29*(3), 242–248.
13. Gruber-Baldini, A. L., Shardell, M., Lloyd, K. D., & Magaziner, J. (2012). Use of proxies and informants. In A. B. Newman & J. A. Cauley (Eds.), *The epidemiology of aging* (pp. 81–90). Springer.
14. Mercieca-Bebber, R., Williams, D., Tait, M. A., Rutherford, C., Busija, L., Roberts, N., Wilson, M., Sundaram, C. S., & Roydhouse, J. (2019). Trials with proxy-reported outcomes registered on the Australian New Zealand clinical trials registry (ANZCTR). *Quality of Life Research, 28*(4), 955–962.
15. Evans, C. J., Benalia, H., Preston, N. J., Grande, G., Gysels, M., Short, V., Daveson, B. A., Bausewein, C., Todd, C., & Higginson, I. J. (2013). The selection and use of outcome measures in palliative and end-of-life care research: the MORECare international consensus workshop. *Journal of Pain and Symptom Management, 46*(6), 925–937.
16. Graham, C. (2016). Incidence and impact of proxy response in measuring patient experience: Secondary analysis of a large postal survey using propensity score matching. *International Journal for Quality in Health Care, 28*(2), 246–252.

17. Bjertnaes, O. (2014). Patient-reported experiences with hospitals: Comparison of proxy and patient scores using propensity-score matching. *International Journal for Quality in Health Care, 26*(1), 34–40.
18. Todorov, A., & Kirchner, C. (2000). Bias in proxies' reports of disability: Data from the National Health Interview Survey on disability. *American Journal of Public Health, 90*(8), 1248–1253.
19. Chawla, N., Urato, M., Ambs, A., Schussler, N., Hays, R. D., Clauser, S. B., Zaslavsky, A. M., Walsh, K., Schwartz, M., Halpern, M., Gaillot, S., Goldstein, E. H., & Arora, N. H. (2015). Unveiling SEER-CAHPS(R): a new data resource for quality of care research. *Jounal of General Internal Medicine, 30*(5), 641–650.
20. Sneeuw, K. C., Aaronson, N. K., Sprangers, M. A., Detmar, S. B., Wever, L. D., & Schornagel, J. H. (1998). Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. *Journal of Clinical Epidemiology, 51*(7), 617–631.
21. Sneeuw, K. C., Aaronson, N. K., de Haan, R. J., & Limburg, M. (1997). Assessing quality of life after stroke. The value and limitations of proxy ratings. *Stroke, 28*(8), 1541–1549.
22. Sneeuw, K. C., Aaronson, N. K., Osoba, D., Muller, M. J., Hsu, M. A., Yung, W. K. A., Brada, M., & Newlands, E. S. (1997). The use of significant others as proxy raters of the quality of life of patients with brain cancer. *Medical Care, 35*(5), 490–506.
23. Pickard, A. S., Johnson, J. A., Feeny, D. H., Shuaib, A., Carriere, K. C., & Nasser, A. M. (2004). Agreement between patient and proxy assessments of health-related quality of life after stroke using the EQ-5D and health utilities index. *Stroke, 35*(2), 607–612.
24. Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*, 1.
25. Lobchuk, M. M., & Degner, L. F. (2002). Patients with cancer and next-of-kin response comparability on physical and psychological symptom well-being: Trends and measurement issues. *Cancer Nursing, 25*(5), 358–374.
26. McPherson, C. J., & Addington-Hall, J. M. (2003). Judging the quality of care at the end of life: Can proxies provide reliable information? *Social Science and Medicine, 56*(1), 95–109.
27. Sneeuw, K. C., Sprangers, M. A., & Aaronson, N. K. (2002). The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *Journal of Clinical Epidemiology, 55*(11), 1130–1143.
28. Sprangers, M. A., & Aaronson, N. K. (1992). The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: A review. *Journal of Clinical Epidemiology, 45*(7), 743–760.
29. Tang, S. T., & McCorkle, R. (2002). Use of family proxies in quality of life research for cancer patients at the end of life: A literature review. *Cancer Investigation, 20*(7–8), 1086–1104.
30. von Essen, L. (2004). Proxy ratings of patient quality of life–factors related to patient-proxy agreement. *Acta Oncologica, 43*(3), 229–234.
31. Roydhouse, J. K., & Wilson, I. B. (2017). Systematic review of caregiver responses for patient health-related quality of life in adult cancer care. *Quality of Life Research, 26*(8), 1925–1954.
32. Liebzeit, D., King, B., & Bratzke, L. (2018). Measurement of function in older adults transitioning from hospital to home: An integrative review. *Geriatric Nursing, 39*(3), 336–343.
33. Mayo, N. E., Figueiredo, S., Ahmed, S., & Bartlett, S. J. (2017). Montreal accord on patient-reported outcomes (PROs) use series—Paper 2: Terminology proposed to measure what matters in health. *Journal of Clinical Epidemiology, 89*, 119–124.
34. McKown, S., Acquadro, C., Anfray, C., Arnold, B., Eremenco, S., Giroudet, C., Martin, M., & Weiss, D. (2020). Good practices for the translation, cultural adaptation, and linguistic validation of clinician-reported outcome, observer-reported outcome, and performance outcome measures. *Journal of Patient Reported Outcomes, 4*(1), 89.
35. Mayo, N. E. (Ed.). (2015). *Dictionary of quality of life and health outcomes measurement*. ISOQOL.
36. Cappelleri, J. C., Zou, K. H., Bushmakin, A. G., Alvir, J. M. J., Alemayehu, D., & Symonds, T. (2014). *Patient-reported outcomes: measurement, implementation and interpretation*. CRC Press.
37. Caiels, J., Rand, S., Crowther, T., Collins, G., & Forder, J. (2019). Exploring the views of being a proxy from the perspective of unpaid carers and paid carers: Developing a proxy version of the adult social care outcomes toolkit (ASCOT). *BMC Health Services Research, 19*, 201.
38. Moher, D., Pham, B., Klassen, T. P., Schulz, K. F., Berlin, J. A., Jadad, A. R., & Liberati, A. (2000). What contributions do languages other than English make on the results of meta-analyses? *Journal of Clinical Epidemiology, 53*(9), 964–972.
39. Kinghorn, P., & Afentou, N. (2020). Proxy responses to ICECAP-A: exploring variation across multiple proxy assessments of capability well-being for the same individuals. *PLoS ONE, 15*(7), e0236584.
40. Roydhouse, J. K., Gutman, R., Keating, N. L., Mor, V., & Wilson, I. B. (2018). Differences between proxy and patient assessments of cancer care experiences and quality ratings. *Health Services Research, 53*(2), 919–943.
41. Roydhouse, J. K., Gutman, R., Keating, N. L., Mor, V., & Wilson, I. B. (2018). Proxy and patient reports of health-related quality of life in a national cancer survey. *Health and Quality of Life Outcomes, 16*(1), 6.
42. Roydhouse, J. K., Gutman, R., Wilson, I. B., Kehl, K. L., & Keating, N. L. (2020). Patient and proxy reports regarding the experience of treatment decision-making in cancer care. *Psycho-Oncology, 29*(11), 1943–1950.
43. Smith, G. E., & Bondi, M. W. (2013). *Mild cognitive impairment and dementia: definitions, diagnosis, and treatment*. Oxford University Press.
44. Alzheimers Association. (2020). *2020 Alzheimer's disease facts and figures*. Alzheimers Dement.
45. Aarsland, D., Zaccai, J., & Brayne, C. (2005). A systematic review of prevalence studies of dementia in Parkinson's disease. *Movement Disorders, 20*(10), 1255–1263.

Authors and Affiliations

Jessica K. Roydhouse[1,2] · Matthew L. Cohen[3] · Henrik R. Eshoj[4] · Nadia Corsini[5] · Emre Yucel[6,7] · Claudia Rutherford[8,9] · Katarzyna Wac[10,11] · Allan Berrocal[10] · Alyssa Lanzi[3] · Cindy Nowinski[12] · Natasha Roberts[13,14] · Angelos P. Kassianos[15] · Veronique Sebille[16,17] · Madeleine T. King[9] · Rebecca Mercieca-Bebber[18,19] on behalf of the ISOQOL Proxy Task Force and the ISOQOL Board of Directors

1 Menzies Institute for Medical Research, University of Tasmania, 17 Liverpool Street, Hobart, TAS 7000, Australia

2 Department of Health Services, Policy & Practice, Brown University School of Public Health, Providence, RI, USA

3 Department of Communication Sciences and Disorders, University of Delaware, Newark, DE, USA

4 Department of Hematology, Quality of Life Research Center, Odense University Hospital, Odense, Denmark

5 Rosemary Bryant AO Research Centre, University of South Australia, Adelaide, SA, Australia

6 Amgen, Global Health Economics, Thousand Oaks, CA, USA

7 Present Address: Bristol Myers Squibb, New York, NY, USA

8 Cancer Nursing Research Unit (CNRU), Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia

9 School of Psychology, Faculty of Science, The University of Sydney, Sydney, NSW, Australia

10 Quality of Life Technologies Lab, University of Geneva, Geneva, Switzerland

11 Quality of Life Technologies Lab, University of Copenhagen, Copenhagen, Denmark

12 Departments of Medical Social Sciences and Neurology, Northwestern University Feinberg School of Medicine, Evanston, IL, USA

13 Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia

14 Queensland University of Technology, Brisbane, QLD, Australia

15 Department of Applied Health Research, University College London, London, UK

16 SPHERE, University of Nantes, University of Tours, INSERM, Nantes, France

17 Department of Methodology and Biostatistics, Nantes University Hospital, Nantes, France

18 Faculty of Medicine, Sydney Medical School, Central Clinical School, The University of Sydney, Sydney, NSW, Australia

19 NHMRC Clinical Trials Centre, The University of Sydney, Sydney, NSW, Australia