Original Article

PALLIATIVE MEDICINE

Palliative Medicine
2023, Vol. 37(5) 760–770

DOI: 10.1177/02692163231157871
journals.sagepub.com/home/pmj
SAGE

Measurement equivalence of the paper-based and electronic version of the Integrated Palliative care Outcome Scale (IPOS): A randomised crossover trial

Anna Bolzani[1], Sophie Kupf[1], Farina Hodiamont[1], Isabel Burner-Fritsch[1], Claudia Bausewein[1] and Christina Ramsenthaler[2,3,4]

Abstract

Background: The Integrated Palliative Care Outcome Scale (IPOS) validly and reliably measures symptoms and concerns of those receiving palliative care.

Aim: To determine the equivalence of the paper version with an electronic version of the IPOS (eIPOS).

Design: Multicentre randomised crossover trial (NCT03879668) with a within-subject comparison of the two modes (washout period 30 min).

Setting/Participants: Convenience sample of specialist inpatient and palliative home care patients aged over 18 years with cancer and non-cancer conditions was recruited. Scores were compared using intraclass correlation coefficients (ICC), Bland-Altman plots and via a mixed-effects analysis of variance.

Results: Fifty patients were randomised to complete paper-electronic ($n = 24$) and electronic-paper ($n = 26$) IPOS with median age 69 years (range 24–95), 56% male, 16% non-cancer. The ICCs showed very high concordance for the total score (ICC 0.99, 95% CI 0.98–1.00), lowest ICCs being observed for symptoms 'Appetite loss' and 'Drowsiness' (ICC 0.95, 95% CI 0.92–0.97). Nine of seventeen items had ICCs above 0.98, as did all subscales. No statistically significant mode, order, age, and interaction effects were observed for IPOS total score and subscales, except for 'Communication' ($F_{mode} = 5.9$, $p = 0.019$). Fifty-eight percent preferred the electronic version. In the group 75+ years, 53% preferred the paper version. Only three entries in the free-text main problems differed between the versions.

Conclusion: The very high equivalence in scores and free text between the IPOS and the eIPOS demonstrates that eIPOS is feasible and reliable in an older palliative population.

Keywords

Palliative care, integrated palliative care outcome scale, mode of administration, agreement, crossover trial, patient-reported outcome measure

What is already known about the topic?

- The implementation of electronic self-reported versions may offer several advantages to palliative and hospice care, especially when fully integrated within an electronic patient health record or to help patients with sensory impairment.
- The typically older palliative population might pose a barrier to the electronic implementation of patient-centred outcome measures.
- Psychometric properties cannot be assumed stable across administration modes, necessitating a careful electronic adaptation of paper versions.

[1]Department of Palliative Medicine, University Hospital Munich, Ludwig-Maximilians-University Munich, Germany
[2]School of Health Professions, Zurich University of Applied Sciences ZHAW, Switzerland
[3]Cicely Saunders Institute of Palliative Care, Policy and Rehabilitation, King's College London, London, UK
[4]Wolfson Palliative Care Research Centre, Hull York Medical School, University of Hull, Hull, UK

Corresponding author:
Claudia Bausewein, University Hospital Munich, Department of Palliative Medicine, Marchioninistr. 15, 81377 Munich, Germany.
Email: claudia.bausewein@med.uni.muenchen.de

What this paper adds?

- This randomised crossover trial in palliative patients showed the equivalence of scores between the self-completed paper and electronic version of the Integrated Palliative care Outcome Scale with near perfect agreement reached for 17 out of 21 items.
- No version took longer to complete. Overall, 58% preferred the electronic version. Only in the group of 75 years or older, slightly more than half preferred the paper version for self-completion.

Implications for practice, theory or policy

- The high agreement and good acceptability of the electronic version of the Integrated Palliative care Outcome Scale was achieved via careful early involvement of patients and staff within a co-design framework.
- Implementation of electronic assessment of patient-centred outcomes in palliative care is feasible once setting-specific barriers and facilitators are acknowledged and addressed in close collaboration with all stakeholders.

Introduction

Outcome measures are instruments that allow the assessment of change in a patient's health status over time. Due to their patient-centredness, they play an increasingly important role in palliative care. Outcome measures are central for identifying patients' needs, aiming at enhancing quality of life and the relief of suffering, and support evaluating the standard of care received.[1–3] In palliative care, both patient and staff-completed versions of outcome measures exist to enable outcome measurement at the end of life.[4–6]

One such outcome measure for palliative care is the Integrated Palliative care Outcome Scale (IPOS) that has been studied extensively over the past years and has seen many developments and adaptations to different settings, conditions and populations.[6,7] Next to the Edmonton Symptom Assessment Scale,[8,9] it is one of the most widely used measures in the field.[4,5,10,11] The IPOS assesses how much a patient is affected by symptoms, emotional concerns as well as communication and practical issues. The validity of the measure (in terms of structural validity, content, construct and criterion validity), its reliability and sensitivity to change have been demonstrated in several international studies for the inpatient, hospice, and specialist home care setting.[6,7,12,13]

The implementation of electronic self-reported outcomes may offer several advantages to the palliative care and the specialist home care setting in particular, such as low administration cost, scalability, adaptability on devices for those with sensory impairments, immediate automated analysis, and the possibility for full integration into an electronic patient record system.[14] The implementation in routine care could lead to shorter reaction times of staff to emerging symptoms/problems or crises, and ultimately to improved patient outcomes.[3] A few such systems for electronic capture with full integration of automated symptom monitoring have been established in oncology in recent years.[15] Only a handful of systems have been implemented in palliative care.[3,16–20]

Despite these advantages and their usefulness, electronic outcome measures are not commonly implemented in the palliative care setting, with only 25% of studies in palliative care and advanced oncology reporting using electronic versions.[5] Fearing the loss of personal contact while judging face-to-face interactions as more suitable in routine palliative care are named as the most important barriers.[21] A further barrier to electronic assessment in the home care setting may be that patients are typically older, and thus belong to a population with associated lower levels of computer and internet usage.[22–24] Moreover, psychometric properties of questionnaires cannot be assumed stable across administration modes,[25] and it is thus recommended to empirically evaluate score equivalence and accordance of modes.

We therefore conducted a randomised crossover trial to test the score equivalence of paper and electronic versions of the IPOS for individual items, subscale scores and the total score. We hypothesised equivalence between the two administration modes.

Methods

Study design

This study was a multicentre, randomised, single-blinded, two-arm crossover trial of 5 months' duration and is part of the project Palli-MONITOR, a multicentre, sequential mixed-methods, two-phase development and feasibility study (NCT03879668; https://clinicaltrials.gov/show/NCT03879668).[26] The study is reported in accordance with the CONSORT crossover guideline.[27] No changes to the original study protocol were made.[26]

The crossover design is the preferred design for establishing measurement equivalence between different modes of administration as per guidelines.[25,28,29] Patients were randomly assigned to two groups defined by the order of mode of administration (paper version of the IPOS first, group 'P-E' vs. electronic completion of the eIPOS first, group 'E-P').

Trial participants and settings

Participants either received specialised (five services) or general palliative home care (two services) or specialist palliative inpatient care at the university hospital's inpatient unit. Eight study sites in Bavaria, Germany, participated from May to September 2019 representing both an urban and rural palliative population. Eligibility criteria were aged 18 years with advanced and incurable malignant or non-malignant disease, the capacity to give written informed consent and being sufficiently fluent in German to complete both questionnaire versions. Exclusion criteria were cognitive impairment or being in a poor general condition or actively dying, as judged by the patient's clinical team. All patients who agreed to participate gave written informed consent. The study was approved by the Medical Institutional Ethics Committee of the University Hospital Munich (REC ref no: 18-871). No further ethical issues emerged during the study.

Interventions and procedures

Patients were screened consecutively for inclusion in the study. Eligible patients, who agreed to share contact details, were contacted by the study team, and informed about the study. If they were interested to participate, a member of the study team contacted them in person to give more information and take written informed consent. Patients were assigned a seven-digit identification number. Allocation to administration order was done by the principal investigator, based on a computer-generated 1:1 schedule (enuvo).[30] The patient then completed both paper IPOS and eIPOS in randomised order, with a 30-min washout period between administrations. The washout period was optimised to the palliative care setting, striking a balance between fluctuating symptom burden, and mitigating recall bias and carry-over effects. Both modes were completed in one visit lasting 45–60 min.

Neither the clinical staff nor the participants were blinded to the result of the randomisation. The statistician conducting the analysis was blinded to group allocation. Assignment to trial arm was concealed on a paper until written informed consent was obtained from participants.

The IPOS is a short 17-item outcome instrument to assess palliative-care related symptoms and concerns in generalist and specialist PC settings.[6,7] The items cover physical symptoms (e.g. pain, shortness of breath, fatigue etc.), emotional concerns (patient and family anxiety, depression, feeling at peace), and communication & practical problems (sharing feelings, information needs, practical problems). Patients can designate their three main symptoms and problems as well as name and rate additional symptoms not included in the symptom list. All closed items are scored on a 5-point Likert scale ranging from 0 'not at all' to 5 'overwhelmingly'. A sum score of all items and three subscale scores can be calculated.[6,12] The paper IPOS can be obtained from www.pos-pal.org.

We used the paper-based, setting-specific 3-day or 7-day recall version of IPOS validated for the German context.[6] Its electronic adaptation was developed to resemble the paper version as closely as possible based on results from an interview study with patients testing a pilot version.[26] For free-text items, patients could enter symptoms and problems in short-answer boxes. Rated items on the IPOS could be answered by selecting the appropriate box on the five-point Likert scale. The eIPOS was provided on all operating systems (Apple, Windows, Android) and devices (e.g. laptop, computer, tablet, smartphone). Navigation buttons at the end of each screen allowed navigation through the questionnaire. Participants could progress to the next item without answers being mandated in eIPOS.

Data collection

For concordance, all answers to open-text and closed items on the IPOS were recorded for both the paper and eIPOS version. The time to complete was taken after each administration. Preference was asked using a closed question. Socio-demographic data included age, gender, nationality, main diagnosis, main care provider, device and operating system used and general use of electronic devices (daily, several times per week, once per week, less than once a week).

Statistical analysis

Sample size. With a power of 80%, a target intraclass correlation coefficient (ICC) of 0.9 and a significance level of $\alpha = 0.05$, the calculated sample size is 47. Considering possible dropouts, it was planned to include 50 participants.

Data analysis. Data management and analyses were conducted with SPSS 27[31] and R 4.0.[32] Data are described via means and standard deviations (SD) for continuous and absolute frequencies and percentages for categorical variables. All data is presented for the whole sample and separately per trial arm. The distribution of scores for quantitative items is compared graphically and via the percentage of floor or ceiling effects (>15% of participants scoring the lowest or highest response option)[33] between the modes. A mixed-effects 2×2 analysis of variance model was used to assess mode effect (within-subject factor), order effect (between-subject factor) and the mode $\times$ order interaction effect. A significant order and interaction would indicate carry-over effects. Additionally, age was fitted as a covariate to evaluate any statistically significant mode $\times$ age effect. Four univariate

mixed-effects models were run with IPOS Total Sum score, IPOS subscale scores and time to complete as dependent variables.

Following guidelines and other research,[25,34–36] the concordance of the IPOS and eIPOS was evaluated using intraclass correlation coefficients (two-way mixed effects model for absolute agreement) with 95% confidence intervals, ranging from 0 to 1, for all individual IPOS items and the subscales and total score. An ICC of >0.90 was considered indicating excellent agreement.[37,38] Prevalence- and bias-adjusted kappa coefficients were additionally used to take possible bias between modes and distributional floor and ceiling effects into account.[39] To assess the magnitude of possible systematic error, we also present the mean difference of scale scores between modes. Data were evaluated graphically by Bland-Altman plots.[40] The score difference (paper minus electronic) was plotted against the average paper and electronic score for each individual, including 95% limits of agreement calculated by $1.96 \times SD_{difference}$. Any systematic bias is thus separated from random measurement error.

A statistical significance level of 5% was used for all analyses. Missing items were imputed with the scale's median. A sensitivity analysis with the imputed data did not produce different results due to the very low rate of missing data.

Results

Participants

A total of 66 eligible patients were invited to participate. Of these, 50 accepted the invitation and were randomised to either 'paper-electronic (P-E)' or 'electronic-paper (E-P)' order. Of those participating in the trial, almost all patients completed all items in both versions. Only one score on the item *Poor mobility* was missing for the paper version, and one score for the item *Sharing feelings* was missing for eIPOS. The trial flow is shown in Figure 1.

There were no significant differences between those allocated to the two orders. The mean age of participants was 67.9 years (SD: 13.6), 56% were male. Demographic details are given in Table 1. Participants accessed the eIPOS most commonly on a tablet (68%), followed by laptop (16%) and PC or mobile phone (8%, respectively). Slightly over half of the sample used their device daily, 6% only used it less than once per week.

Descriptive statistics

Descriptive statistics for the distribution of scores on the IPOS and the eIPOS are provided in Table 2. The mean and median scores between the two modes only differed in the first decimal. Small consistent differences existed for the proportion of floor and ceiling effects with the eIPOS showing a slightly smaller proportion of floor effects in four symptom items, two of the four emotional subscale items, and all Communication & practical problems subscale items. The score distribution is presented graphically in Supplemental Figures 1 and 2 for the IPOS and the eIPOS. Fifty-eight percent of patients preferred the electronic version, 40% the paper version.

Mode equivalence

The mixed-effects analysis of variance with mode and order as the two main effects showed mean differences of −0.1 for the IPOS subscale scores and −0.3 for the IPOS total score when comparing the IPOS to the eIPOS. Table 3 shows that all mode and order effects as well as interaction effects were statistically non-significant, except for a statistically significant mode effect for the IPOS Communication & Practical problems subscale (F ($2,48$) = 5.9, $p < 0.019$). None of the mode $\times$ age interaction effects were statistically significant. Table 4 presents results for the outcome mean time to complete for both modes. There was no statistically significant order, mode, or interaction effect for time to complete. However, the paper and electronic versions differed significantly between age groups with patients aged 60 years or younger requiring the shortest time to complete both modes.

All ICCs for the comparison of IPOS and eIPOS were $\geqslant 0.95$ (see Table 5). *Appetite loss* and *Drowsiness* presented the lowest ICCs with 0.95 (95% CI: 0.92; 0.97). *Constipation* and *Feeling not at peace* (spiritual concerns) had ICCs of 0.96 (95% CI: 0.94; 0.98). Four items (*Family anxiety*, *Feeling depressed*, *Fatigue*, *Nausea*) had ICCs of 0.97 (95% CI: 0.95; 0.98). All other items showed ICCs of 0.98 or above. The lowest kappa score was found for *Drowsiness* ($\kappa = 0.78$), followed by *Appetite loss* and the IPOS Emotional concerns subscale with $\kappa = 0.82$, respectively. Nine items, two subscales and the IPOS total score showed κ above 0.92.

Bland-Altman plots for the three subscales and the IPOS total scores for the comparison of paper version and eIPOS are presented in Supplemental Appendix 2. The systematic bias was largest for the IPOS total score with mean difference −0.13 (limits of agreement: −3.14; 2.89), followed by the IPOS Communication & Practical problems subscale ($M_{diff} = -0.10$, limits of agreement: −0.69; 0.49). The IPOS Emotional concerns subscale showed a systematic bias of −0.08 (limits of agreement: −1.61; 1.45). The IPOS Physical symptoms subscale was measured without a systematic bias between the modes ($M_{diff} = 0.00$, limits of agreement: −1.94; 1.94).

Concordance of free-text answers

Thirteen participants overall did not volunteer any free-text main problems. Of those indicating main problems, 81% showed exact concordance of answers. Volunteered

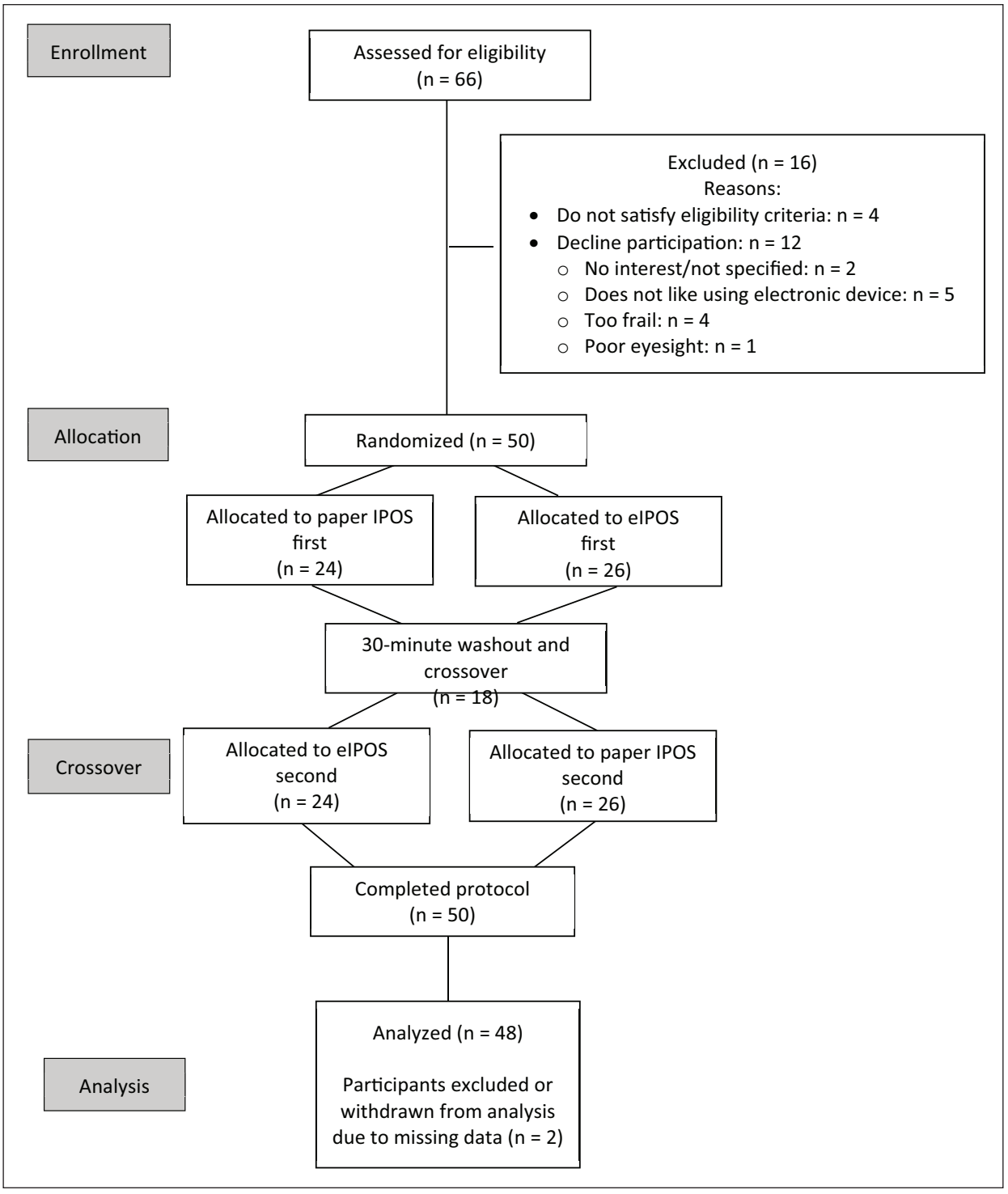

Figure 1. Flow diagram for crossover trial of paper versus electronic version of IPOS in palliative care.

problems were mainly physical symptoms. Four persons volunteered different problems in the electronic version and three changed the order of main problems.

Twelve patients volunteered additional symptoms at the end of the IPOS symptom list. The concordance was nearly perfect, with one person scoring one symptom differently.

Discussion

Main findings/results of the study

The results of this randomised crossover trial indicate highly comparable and concordant responses between the paper version of IPOS and the eIPOS, at the total

Table 1. Demographic and clinical characteristics of study participants (*n* = 50).

	All (*n* = 50)	Paper IPOS first (*n* = 24)	eIPOS first (*n* = 26)
Age, years			
Mean ± SD	67.9 (13.6)	68.4 (13.3)	67.4 (14.2)
Median (IQR)	69 (60–77)	67 (60.25–78.5)	71 (56.5–76.25)
⩽ 60 years	13 (26)	6 (25.0)	7 (26.9)
61–74 years	20 (40)	11 (45.8)	9 (34.6)
75 + years	17 (34)	7 (29.2)	10 (38.5)
Sex			
Male	28 (56.0)	14 (58.3)	14 (53.8)
Female	22 (44.0)	10 (41.7)	12 (46.2)
Nationality			
German	44 (88.0)	21 (87.5)	23 (88.5)
Other	6 (12.0)	3 (12.5)	3 (11.5)
Diagnosis			
Cancer	42 (84.0)	20 (83.3)	22 (84.6)
Non-cancer	8 (16.0)	4 (16.7)	4 (15.4)
Cancer: Digestive organs	10 (20.0)	6 (25.0)	4 (15.4)
Respiratory tract	5 (10.0)	3 (12.5)	2 (7.7)
Genitourinary tract	9 (18.0)	5 (20.8)	4 (15.4)
Breast	8 (16.0)	4 (16.7)	4 (15.4)
Lymph/Haematopoietic	1 (2.0)	0 (0)	1 (3.8)
Brain	4 (8.0)	1 (4.2)	3 (11.5)
Other cancer[a]	5 (10.0)	1 (4.2)	4 (15.4)
Non-cancer: COPD or ILD	3 (6.0)	1 (4.2)	2 (7.7)
Renal failure	2 (4.0)	1 (4.2)	1 (3.8)
Heart failure	2 (4.0)	2 (8.3)	0 (0)
Other non-cancer[b]	1 (2.0)	0 (0)	1 (3.8)
Setting			
Inpatient palliative care	22 (44.0)	11 (45.8)	11 (42.3)
Specialist palliative home care	22 (44.0)	10 (41.7)	12 (46.2)
General home care service	6 (12.0)	3 (12.5)	3 (11.5)
Device			
PC	4 (8.0)	1 (4.2)	3 (11.5)
Laptop	8 (16.0)	2 (8.3)	6 (23.1)
Tablet	34 (68.0)	18 (75.0)	16 (61.5)
Mobile phone	4 (8.0)	3 (12.5)	1 (3.8)
Operating system			
Apple	34 (68.0)	19 (79.2)	15 (57.7)
Windows	11 (22.0)	3 (12.5)	8 (30.8)
Android	5 (10.0)	2 (8.3)	3 (11.5)
Use of electronic or mobile devices			
Daily	26 (52.0)	13 (54.2)	13 (50.0)
Several times per week	9 (18.0)	4 (16.7)	5 (19.2)
Once per week	2 (4.0)	0 (0)	2 (7.7)
Less than once per week	3 (6.0)	1 (4.2)	2 (7.7)
Never	10 (20.0)	6 (25.0)	4 (15.4)

[a]Other cancers: Thyroid cancer, Ewing sarcoma, Pharynx cancer, Melanoma, [b]Other non-cancer: Unspecified non-cancer condition.
COPD: chronic obstructive pulmonary disease; eIPOS: electronic version of IPOS; ILD: interstitial lung disease; IQR: interquartile range; IPOS: Integrated Palliative care Outcome Scale; PC: personal computer; SD: standard deviation.

score, the subscale scores, and the individual item level. Mean summary and subscale score differences were very small (<1% of score ranges) and non-significant throughout all analyses. ICCs between paper and electronic scores were very high, significant and all exceeded the cut-off of >0.90.[24,25,28,29] *Appetite loss*, *Constipation*, *Drowsiness* and *Feeling not at peace* showed the lowest ICCs in comparison to other items, the subscales, and the total score.

Table 2. Sample statistics for the paper and electronic IPOS versions (n = 50).

	Paper IPOS				eIPOS			
	Mean (SD)	Median	% floor	% ceiling	Mean (SD)	Median	% floor	% ceiling
Pain	1.9 (1.1)	2	12	8	1.9 (1.2)	2	14	8
Shortness of breath	1.2 (1.3)	1	44	2	1.3 (1.3)	1	42	4
Fatigue	2.6 (1.1)	3	8	22	2.6 (1.1)	3	6	24
Nausea	1.3 (1.2)	1	36	6	1.2 (1.2)	1	36	6
Vomiting	0.7 (1.1)	0	64	4	0.7 (1.1)	0	64	4
Appetite loss	1.6 (1.3)	2	30	8	1.6 (1.3)	2	26	8
Constipation	1.2 (1.4)	0	52	8	1.3 (1.4)	1	46	8
Dry mouth	1.4 (1.3)	1	34	6	1.4 (1.3)	1	34	6
Drowsiness	2.2 (1.1)	2	8	8	2.2 (1.1)	2	10	10
Poor mobility	2.9 (1.1)	3	4	34	2.8 (1.1)	3	4	30
Patient anxiety	2.4 (1.3)	2	12	22	2.4 (1.2)	2,5	6	20
Family anxiety	2.9 (1.0)	3	2	32	2.9 (1.0)	3	2	34
Depressed	2.2 (1.3)	2	12	18	2.2 (1.2)	2	12	14
Not at peace	1.2 (1.0)	1	18	4	1.3 (0.9)	1	12	2
Sharing feelings	0.8 (0.9)	1	42	2	0.9 (0.9)	1	40	2
Information needs	0.9 (0.9)	1	34	4	0.9 (0.9)	1	32	4
Practical problems	1.1 (1.2)	1	40	6	1.1 (1.2)	1	38	6
IPOS total score	28.4 (9.0)	28	4	0	28.7 (8.8)	29	4	0
IPOS Physical symptoms	17.0 (6.9)	16	8	2	17.1 (6.9)	17	8	2
IPOS Emotional concerns	8.7 (3.8)	8,5	10	18	8.8 (3.6)	9	10	14
IPOS Communication & Practical problems	2.8 (2.3)	3	44	2	2.9 (2.3)	3	40	2

eIPOS: electronic version of IPOS; IPOS: Integrated Palliative care Outcome Scale; SD: standard deviation.

Table 3. Results of mixed-effects 2 × 2 analysis of variance (Mode: Paper or electronic, Order: P-E vs. E-P, Interaction, Covariate: age) in n = 50 palliative patients.

	Paper IPOS	eIPOS	Mean of paired differences	F^a, p			
	Mean (SD)	Mean (SD)	P–E (SD_{diff})	Mode effect	Order effect	Mode × order	Mode × age
IPOS total score	28.4 (9.0)	28.7 (8.8)	−0.3 (1.9)	F = 1.3 p = 0.258	F = 0.8 p = 0.381	F = 0.3 p = 0.572	F = 0.8 p = 0.373
IPOS Physical symptoms	17.0 (6.9)	17.1 (6.9)	−0.1 (1.3)	F = 0.4 p = 0.511	F = 1.4 p = 0.244	F = 0.1 p = 0.808	F = 0.8 p = 0.377
IPOS Emotional concerns	8.7 (3.8)	8.8 (3.6)	−0.1 (0.7)	F = 0.5 p = 0.466	F = 0.1 p = 0.984	F = 0.2 p = 0.699	F = 0.0 p = 0.984
IPOS Communication and Practical problems	2.8 (2.3)	2.9 (2.3)	−0.1 (0.3)	F = 5.9 p = 0.019	F = 0.1 p = 0.957	F = 2.3 p = 0.137	F = 3.6 p = 0.062

[a]All degrees of freedom for the F ratio: 2, 48 (except for Mode × age interaction).
F ratios and p-values in bold are statistically significant on the 5% level.
eIPOS: electronic version of IPOS; F: F ratio; IPOS: Integrated Palliative care Outcome Scale; p: p-value; P-E: Mean difference between paper and electronic version of the IPOS; SD: standard deviation; SD_{diff}: Standard deviation of the difference.

The concordance extends to volunteered main problems and concerns as well as volunteered symptoms. The preference for the eIPOS was higher than for the paper version. A mode × age interaction effect was shown for the IPOS Communications & Practical problems subscale, and a significant age effect was shown for completion time.

What this study adds

Compared to studies assessing the concordance of paper and electronic versions of outcome measures in populations of patients with advanced diseases (mostly cancer),[36,41–44] our sample comprised older patients, an equal gender distribution with a more heterogenous disease variety due to its palliative sampling frame. Contrary to the age and gender bias in equivalence studies reported in general clinical populations,[24,45] a significant mode × age interaction effect was only observed for items *Sharing feelings with family/friends*, *Information needs* and *Practical problems*. Symptoms and emotional concerns were reported in equal manner between the modes. This shows that electronic adaptations of

Table 4. Results of mixed-effects 2 × 2 analysis of variance for outcome mean time to complete in *n* = 50 palliative patients.

		Paper IPOS		eIPOS			
	No. of patients	Mean Time	95% CI	Mean Time	95%	M_{diff}	95% CI
All patients	50	5.82	5.28, 6.36	5.81	5.19, 6.42	0.01	−0.22, 0.24
Patients by order of administration							
Order P-E	24	6.35	5.46, 7.24	6.21	5.19, 7.22	0.15	−0.26, 0.55
Order E-P	26	5.33	4.68, 5.97	5.44	4.68, 6.20	−0.12	−0.37, 0.14
Patients by age, in years*							
⩽60	13	4.35	3.39, 5.31	3.81	2.80, 4.81		
61–74	20	6.15	5.38, 6.92	6.13	5.32, 6.93		
75+	17	6.56	5.72, 7.40	6.97	6.09, 7.85		

CI: confidence interval; E-P: eIPOS first, then paper version; eIPOS: electronic version of IPOS; IPOS: Integrated Palliative care Outcome Scale; M_{diff}: Mean of the difference; P-E: paper version of IPOS first, then eIPOS.
*Paper version: Differences for time to complete among age groups: *F(2, 47)* = 6.7, *p* = 0.003; Electronic version: *F(2, 47)* = 11.9, $p < 0.001$.

Table 5. Spearman correlations, prevalence and bias adjusted kappa coefficients (PABAK), mean differences (with 95% confidence interval), intraclass correlation coefficients for the agreement between paper IPOS and eIPOS (*n* = 50).

	r_s	PABAK	M_{diff} (95% CI)	ICC (95% CI)
Pain	0.97	0.92	0.00 (−0.08; 0.08)	0.98 (0.98; 0.99)
Shortness of breath	0.99	0.96	−0.04 (−0.09; 0.02)	0.99 (0.99; 1.00)
Fatigue	0.95	0.88	0.00 (−0.10; 0.10)	0.97 (0.96; 0.98)
Nausea	0.98	0.92	0.04 (−0.04; 0.12)	0.97 (0.96; 0.98)
Vomiting	0.99	0.96	0.00 (−0.06; 0.06)	0.98 (0.97; 0.99)
Appetite loss	0.94	0.82	−0.06 (−0.18; 0.06)	0.95 (0.92; 0.97)
Constipation	0.96	0.86	−0.06 (−0.17; 0.05)	0.96 (0.94; 0.98)
Dry mouth	0.99	0.94	–0.02 (−0.09; 0.05)	0.98 (0.97; 0.99)
Drowsiness	0.89	0.78	–0.02 (–0.15; 0.11)	0.95 (0.92; 0.97)
Poor mobility	0.95	0.92	0.04 (−0.04; 0.12)	0.98 (0.97; 0.99)
Patient anxiety	0.97	0.88	−0.04 (−0.14; 0.06)	0.98 (0.97; 0.99)
Family anxiety	0.92	0.86	−0.02 (−0.13; 0.09)	0.97 (0.95; 0.98)
Depressed	0.93	0.88	0.02 (−0.10; 0.14)	0.97 (0.95; 0.98)
Not at peace	0.94	0.88	−0.04 (−0.14; 0.06)	0.96 (0.94; 0.98)
Sharing feelings	0.96	0.94	−0.06 (−0.13; 0.01)	0.98 (0.97; 0.99)
Information needs	0.97	0.98	−0.02 (−0.06; 0.02)	0.99 (0.98; 1.00)
Practical problems	0.98	0.98	−0.02 (−0.06; 0.02)	0.99 (0.99; 1.00)
IPOS total score	0.98	0.99	−0.13 (−0.57; 0.32)	0.99 (0.99; 1.00)
IPOS Physical symptoms	0.99	0.94	0.00 (−0.29; 0.29)	0.99 (0.98; 1.00)
IPOS Emotional concerns	0.98	0.82	−0.08 (–0.30; 0.14)	0.99 (0.98; 0.99)
IPOS Communication & Practical problems	0.99	0.94	−0.10 (−0.19; −0.01)	0.99 (0.99; 1.00)

CI: confidence interval; ICC: intraclass correlation coefficient; IPOS: Integrated Palliative care Outcome Scale; M_{diff}: Mean difference; PABAK: prevalence and bias-adjusted Kappa; r_s: Spearman's rho.

measures are possible despite the challenging palliative setting.

With the exception of one study in cancer patients,[44] equivalence studies report a higher preference for electronic versions of PROs for 52%–67% of the sample.[36,41–43] Since qualitative data is missing, this preference is not explained. Equivalence studies in general populations with non-advanced disease indicate that the preference of electronic outcome measures is strongly a function of age.[24,45] In palliative care, however, advanced illness and a traditional focus on delivering interventions via expert face-to-face communication coupled with a generally older population may hinder the successful implementation of electronic versions. A successful inclusion of self and proxy-reported electronic measures within an outpatient hospice population has been shown in the past.[17,46] It is also worth pointing out that electronic completion of outcome measures does not preclude face-to-face interaction and follow-up communication.

The level of concordance of the paper and electronic versions of a self-reported outcomes found in the present study was excellent. The agreement found was higher

than in similar studies testing the equivalence for quality of life and/or morbidity measures and showing acceptable to good ICCs of ⩾0.7,[41,43] or moderate to good agreement based on weighted kappa coefficients.[36,42,44] None of these cited studies, however, reached consistently high ICCs across both the total score and almost all subscale scores as we did in our study. Systematically reviewed features leading to high agreement have been coupled to randomised designs of shorter duration,[22,24,47] features clearly met in our study. High agreement may also be attributable to the deliberate early involvement of patients in the development of the electronic version via co-design. Additional research is needed to understand how visual factors contribute to high agreement between paper and electronic versions. To help older adults and/or those with peripheral neuropathy, qualitative evidence has also supported stylus or pen entry of data into electronic devices instead of the more common swipe-and-touch techniques.[48]

With the demonstration of high reproducibility and concordance between the two versions, the regular use of electronic IPOS in palliative home care may help harness the power of rapid, real-time assessment and feedback to patients and clinicians. This might also enhance interdisciplinary communication and care.[49,50] The setting itself need not be a barrier for the successful adoption of electronic versions.[51–53] However, implementation strategies need to recognise barriers and facilitators specific to the setting and a close collaboration with care teams is paramount.[49,54,55]

Limitations of the study

First, including only cognitively able patients might have resulted in a sampling bias, as up to 90% of palliative patients demonstrate some form of cognitive impairment before death.[56] Proxy-rated staff versions are available, but no proxy version for informal caregivers exists yet. Both should be tested for measurement equivalence when migrating to an electronic version. Second, the timing of assessments and selecting the appropriate washout period is a challenge in PC due to the fast-changing symptom burden specifically in inpatient populations as evidenced by often low to moderate test-retest reliability of measures.[6] Albeit we could not detect a significant order effect, these carryover effects cannot be excluded. The lower kappa values for the emotional subscale may point towards differences in interpretation of the underlying constructs being measured by the items and should be addressed in future studies on content and cross-cultural validity.[57]

Conclusion

Following the recommendations of the ISPOR guidelines, the results show that eIPOS is a valid and reliable measure in the palliative setting. Paper and electronic versions of the IPOS can be considered equivalent and interchangeable. This means a fundamental step towards a more widespread routine implementation of measures and their positive effects for the palliative home care setting. The challenge of using data from electronically implemented outcome measures effectively in routine clinical care remains, so that these measures can foster the patient-professional dialogue and help professionals deliver high-quality care.

Acknowledgements

We want to thank all collaborating services (specialist and general PC home services), their staff and their patients for the collaboration and for their efforts in collecting data and supporting the study. A special thanks goes to all patients and their families for participating in our study and giving us their time.

Author contribution

CB is the chief investigator and responsible for the design of this study. AB is the principle investigator and responsible for the conduct of this study. CB, FH, CR and AB conceptualised the study. SK, IBF and AB collected the data. SK and CR analysed the data. CR drafted the manuscript, with the support of AB, IBF, and SK. All authors provided critical feedback to the manuscript, read and approved the final draft.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Federal Joint Committee German Innovation Fund (Innovationsausschuss des Gemeinsamen Bundesausschuss), grant number 01VSF17014. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Ethics and consent

This crossover trial received research ethics approval from the local research ethics committee of the LMU University Hospital Munich (REC ref no: 18-871). All participants provided informed written consent. All study procedures adhere to the World Medical Association Declaration of Helsinki.

Data management and sharing

Due to data sharing restrictions set out by the ethics committee and the local data governance department, the full datasets used and analysed are only available from the corresponding author on reasonable request. R code for analyses can be obtained via contacting the corresponding author.

ORCID iDs

Sophie Kupf https://orcid.org/0000-0003-3880-6651
Farina Hodiamont https://orcid.org/0000-0001-7351-7092
Isabel Burner-Fritsch https://orcid.org/0000-0002-8356-3460
Claudia Bausewein https://orcid.org/0000-0002-0958-3041
Christina Ramsenthaler https://orcid.org/0000-0002-9996-1818

Supplemental material

Supplemental material for this article is available online.

References

1. Higginson IJ and Carr AJ. Measuring quality of life: using quality of life measures in the clinical setting. *BMJ* 2001; 322: 1297–1300.
2. O'Connor R. *Measuring quality of life in health*. London: Churchill Livingstone, 2004.
3. Currow DC, Allingham S, Yates P, et al. Improving national hospice/palliative care service symptom outcomes systematically through point-of-care data collection, structured feedback and benchmarking. *Support Care Cancer* 2015; 23: 307–315.
4. Albers G, Echteld MA, de Vet HC, et al. Evaluation of quality-of-life measures for use in palliative care: a systematic review. *Palliat Med* 2010; 24: 17–37.
5. Etkind SN, Daveson BA, Kwok W, et al. Capture, transfer, and feedback of patient-centered outcomes data in palliative care populations: does it make a difference? A systematic review. *J Pain Symptom Manag* 2015; 49: 611–624.
6. Murtagh FE, Ramsenthaler C, Firth A, et al. A brief, patient- and proxy-reported outcome measure in advanced illness: validity, reliability and responsiveness of the Integrated Palliative care Outcome Scale (IPOS). *Palliat Med* 2019; 33: 1045–1057.
7. Schildmann EK, Groeneveld EI, Denzel J, et al. Discovering the hidden benefits of cognitive interviewing in two languages: the first phase of a validation study of the Integrated Palliative care Outcome Scale. *Palliat Med* 2016; 30: 599–610.
8. Chang VT, Hwang SS and Feuerman M. Validation of the Edmonton Symptom Assessment Scale. *Cancer* 2000; 88: 2164–2171.
9. Nekolaichuk C, Watanabe S and Beaumont C. The Edmonton Symptom Assessment System: a 15-year retrospective review of validation studies (1991–2006). *Palliat Med* 2008; 22: 111–122.
10. Bausewein C, Simon ST, Benalia H, et al. Implementing patient reported outcome measures (PROMs) in palliative care—users' cry for help. *Health Qual Life Outcomes* 2011; 9: 27.
11. Collins ES, Witt J, Bausewein C, et al. A systematic review of the use of the palliative care outcome scale and the support team assessment schedule in palliative care. *J Pain Symptom Manag* 2015; 50: 842–853.e19.
12. Ramsenthaler C, Davies JM, Higginson IJ, et al. The internal structure of the Integrated Palliative Care Outcome Scale (IPOS): Evidence for a general palliative care factor in addition to symptoms, emotional well-being and quality of care as domains of palliative care. *Palliat Med* 2019; 33: 32.
13. Sandham MH, Medvedev ON, Hedgecock E, et al. A Rasch analysis of the integrated palliative care outcome scale. *J Pain Symptom Manag* 2019; 57: 290–296.
14. Kotronoulas G, Kearney N, Maguire R, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. *J Clin Oncol* 2014; 32: 1480–1501.
15. Bennett AV, Jensen RE and Basch E. Electronic patient-reported outcome systems in oncology clinical practice. *CA Cancer J Clin* 2012; 62: 337–347.
16. Aapro M, Bossi P, Dasari A, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. *Support Care Cancer* 2020; 28: 4589–4612.
17. Dy SM, Roy J, Ott GE, et al. Tell Us™: a web-based tool for improving communication among patients, families, and providers in hospice and palliative care through systematic data specification, collection, and use. *J Pain Symptom Manage* 2011; 42: 526–534.
18. Kyte D, Anderson N, Auti R, et al. Development of an electronic patient-reported outcome measure (ePROM) system to aid the management of patients with advanced chronic kidney disease. *J Patient Rep Outcomes* 2020; 4: 55.
19. Maguire R, Fox PA, McCann L, et al. The eSMART study protocol: a randomised controlled trial to evaluate electronic symptom management using the advanced symptom management system (ASyMS) remote technology for patients with cancer. *BMJ Open* 2017; 7: e015016.
20. McCANN L, Maguire R, Miller M, et al. Patients' perceptions and experiences of using a mobile phone-based advanced symptom management system (ASyMS©) to monitor and manage chemotherapy related toxicity. *Eur J Cancer Care* 2009; 18: 156–164.
21. Radionova N, Becker G, Mayer-Steinacker R, et al. The views of physicians and nurses on the potentials of an electronic assessment system for recognizing the needs of patients in palliative care. *BMC Palliat Care* 2020; 19: 45.
22. Gwaltney CJ, Shields AL and Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health* 2008; 11: 322–333.
23. Loh KP, McHugh C, Mohile SG, et al. Using information technology in the assessment and monitoring of geriatric oncology patients. *Curr Oncol Rep* 2018; 20: 25.
24. Muehlhausen W, Doll H, Quadri N, et al. Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health Qual Life Outcomes* 2015; 13: 167.
25. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009; 12: 419–429.
26. Bolzani A, Ramsenthaler C, Hodiamont F, et al. Monitoring of palliative care symptoms and concerns in specialized palliative home care using an electronic version of the Integrated Palliative care Outcome Scale (Palli-MONITOR):

protocol for a mixed-methods study. *BMJ Open* 2021; 11: e042266.
27. Dwan K, Li T, Altman DG, et al. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* 2019; 366: l4378.
28. Muehlhausen W, Byrom B, Skerritt B, et al. Standards for instrument migration when implementing paper patient-reported outcome instruments electronically: recommendations from a qualitative synthesis of cognitive interview and usability studies. *Value Health* 2018; 21: 41–48.
29. Zbrozek A, Hebert J, Gogates G, et al. Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. *Value Health* 2013; 16: 480–489.
30. No authors. *enuvo*, http://www.umfrageonline.de (2022, accessed 8 September 2018).
31. IBM. IBM SPSS version 27. Armonk, NY: IBM, 2020.
32. R Core Team. *R. A language and environment for statistical computing*, https://www.R-project.org/ (2022, accessed 15 June 2021).
33. de Vet HCW, Terwee CB, Mokkink LB, et al. *Measurement in medicine*. Cambridge: Cambridge University Press, 2011.
34. Bishop FL, Lewis G, Harris S, et al. A within-subjects trial to test the equivalence of online and paper outcome measures: the Roland Morris disability questionnaire. *BMC Musculoskelet Disord* 2010; 11: 113.
35. Cook AJ, Roberts DA, Henderson MD, et al. Electronic pain questionnaires: a randomized, crossover comparison with paper questionnaires for chronic pain assessment. *Pain* 2004; 110: 310–317.
36. Velikova G, Wright EP, Smith AB, et al. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *J Clin Oncol* 1999; 17: 998–1007.
37. Koo TK and Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
38. Liljequist D, Elfving B and Skavberg Roaldsen K. Intraclass correlation – a discussion and demonstration of basic features. *PLoS One* 2019; 14: e0219854.
39. Byrt T, Bishop J and Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993; 46: 423–429.
40. Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
41. Ashley L, Keding A, Brown J, et al. Score equivalence of electronic and paper versions of the Social Difficulties Inventory (SDI-21): a randomised crossover trial in cancer patients. *Qual Life Res* 2013; 22: 1435–1440.
42. Chang Y-J, Chang C-H, Peng C-L, et al. Measurement equivalence and feasibility of the EORTC QLQ-PR25: paper-and-pencil versus touch-screen administration. *Health Qual Life Outcomes* 2014; 12: 23.
43. Knoerl R, Gray E, Stricker C, et al. Electronic versus paper-pencil methods for assessing chemotherapy-induced peripheral neuropathy. *Support Care Cancer* 2017; 25: 3437–3446.
44. Lee E-H. Touch-screen computerized quality-of-life assessment for patients with cancer. *Asian Nurs Res* 2009; 3: 41–48.
45. Campbell N, Ali F, Finlay AY, et al. Equivalence of electronic and paper-based patient-reported outcome measures. *Qual Life Res* 2015; 24: 1949–1961.
46. Stukenborg GJ, Blackhall L, Harrison J, et al. Cancer patient-reported outcomes assessment using wireless touch screen tablet computers. *Qual Life Res* 2014; 23: 1603–1607.
47. Byrom B, Gwaltney C, Slagle A, et al. Measurement equivalence of patient-reported outcome measures migrated to electronic formats: a review of evidence and recommendations for clinical trials and bring your own device. *Ther Innov Regul Sci* 2019; 53: 426–430.
48. Mowlem FD, Sanderson B, Platko JV, et al. Optimizing electronic capture of patient-reported outcome measures in oncology clinical trials: lessons learned from a qualitative study. *J Comp Eff Res* 2020; 9: 1195–1204.
49. Bradshaw A, Santarelli M, Mulderrig M, et al. Implementing person-centred outcome measures in palliative care: an exploratory qualitative study using normalisation process theory to understand processes and context. *Palliat Med* 2021; 35: 397–407.
50. Bush RA, Pérez A, Baum T, et al. A systematic review of the use of the electronic health record for patient identification, communication, and clinical support in palliative care. *JAMIA Open* 2018; 1: 294–303.
51. Willis L, Demiris G and Oliver DP. Internet use by hospice families and providers: a review. *J Med Syst* 2007; 31: 97–101.
52. Kamal AH, Kavalieratos D, Bull J, et al. Usability and acceptability of the QDACT-PC, an electronic point-of-care system for standardized quality monitoring in palliative care. *J Pain Symptom Manag* 2015; 50: 615–621.
53. Kotronoulas G. Benefits, challenges, and opportunities of integrating patient-reported outcome measures in geriatric oncology to advance patient screening for functional fitness for treatment. *Semin Oncol Nurs* 2021; 37: 151230.
54. Anderson NE, McMullan C, Calvert M, et al. Using patient-reported outcome measures during the management of patients with end-stage kidney disease requiring treatment with haemodialysis (PROM-HD): a qualitative study. *BMJ Open* 2021; 11: e052629.
55. Antunes B, Harding R and Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: a systematic review of facilitators and barriers. *Palliat Med* 2014; 28: 158–175.
56. Burton CZ, Twamley EW, Lee LC, et al. Undetected cognitive impairment and decision-making capacity in patients receiving hospice care. *Am J Geriatr Psychiatry* 2012; 20: 306–316.
57. Gerlach C, Taylor K, Ferner M, et al. Challenges in the cultural adaptation of the German Myeloma Patient Outcome Scale (MyPOS): an outcome measure to support routine symptom assessment in myeloma care. *BMC Cancer* 2020; 20: 245.