



# Paper and electronic versions of HM-PRO, a novel patient-reported outcome measure for hematology: an equivalence study

Journal of **Comparative Effectiveness Research**

Pushpendra Goswami<sup>1</sup> , Esther N Oliva<sup>2</sup>, Tatyana Ionova<sup>3</sup>, Roger Else<sup>4</sup>, Jonathan Kell<sup>5</sup>, Adele K Fielding<sup>6</sup>, Daniel M Jennings<sup>7</sup>, Marina Karakantza<sup>8</sup>, Saad Al-Ismaïl<sup>9</sup>, Jack Lyness<sup>9</sup>, Graham P Collins<sup>10</sup>, Stewart McConnell<sup>8</sup>, Catherine Langton<sup>8</sup>, Magda Jabbar Al-Obaidi<sup>11</sup>, Metod Oblak<sup>11</sup> & Sam Salek<sup>\*,1</sup>

<sup>1</sup>School of Life & Medical Sciences, University of Hertfordshire, Hatfield, UK

<sup>2</sup>Haematology Unit, Grande Ospedale Metropolitano, Reggio Calabria, Italy

<sup>3</sup>St Petersburg State University Medical Centre & Multinational Centre for Quality of Life Research, St Petersburg, Russia

<sup>4</sup>Patient Research Partner, Milton Keynes, UK

<sup>5</sup>Haematology, Cardiff & Vale University Health Board, Cardiff, UK

<sup>6</sup>Haematology, University College London Cancer Institute, London, UK

<sup>7</sup>Haematology, Royal Surrey County Hospital NHS Foundation Trust, Guildford, Surrey, UK

<sup>8</sup>Haematology, Leeds Teaching Hospitals NHS Trust, Leeds, UK

<sup>9</sup>Haematology, Singleton Hospital, ABM University Health Board, Swansea, UK

<sup>10</sup>Haematology, Oxford University Hospitals NHS Trust, Oxford, UK

<sup>11</sup>Haematology, West Middlesex University Hospital, Isleworth, UK

\*Author for correspondence: [m.s.salek@herts.ac.uk](mailto:m.s.salek@herts.ac.uk)

**Aim:** To determine measurement equivalence of paper and electronic application of the hematological malignancy-patient-reported outcome (HM-PRO), a specific measure for the evaluation of patient-reported outcomes in HMs. **Patients & methods:** Following International Society of Pharmacoeconomics and Outcomes Research ePRO Good Research Practice Task Force guidelines, a total of 193 adult patients with different HMs were recruited into a multicenter prospective study. The paper and the electronic version of the instrument were completed in the outpatient clinics in a randomized crossover design with a 30-min time interval to minimize the learning effect. Those who completed the paper version first, completed the electronic version after 30 min and vice versa. Instrument version and order effects were tested on total score of the two parts of the HM-PRO (Part A: quality of life and Part B: signs & symptoms) in a two-way ANOVA with patients as random effects. Intraclass correlation coefficients (95% CI) and Spearman's rank correlation coefficients were used to evaluate test-retest reliability and reproducibility. The effects of instrument version and order were tested on total score of the two parts of HM-PRO. **Results:** The questionnaire version and administration order effects were not significant at the 5% level. There were no interactions found between these two factors for HM-PRO (Part A [quality of life];  $p = 0.95$ ); and (part B [signs and symptoms];  $p = 0.72$ ). Spearman's rank correlation coefficients were greater than 0.9, and intraclass correlation coefficients ranged from 0.94 to 0.98; furthermore, the scores were not statistically different between the two versions, showing acceptable reliability indexes. Noteworthy, the difference between the completion time for both paper (mean = 6:38 min) and electronic version (mean = 7:29 min) was not statistically significant ( $n = 100$ ;  $p = 0.11$ ). Patients did not report any difficulty in completing the electronic version during cognitive interviews and were able to understand and respond spontaneously. **Conclusion:** Measurement equivalence has been demonstrated for the paper and electronic application of the HM-PRO.

First draft submitted: 9 October 2018; Accepted for publication: 12 February 2019; Published online: 30 April 2019

**Keywords:** hematological malignancy • HM-PRO • paper/electronic PRO

Patient-reported outcome (PRO) measures are widely used in clinical research, and there is a general consensus that the incorporation of PROs in daily clinical practice will play an important role in the delivery of tailored

Future  
Medicine

care, which includes understanding health outcomes, monitoring performance and informing clinical and policy decisions [1,2]. In addition, understanding signs and symptoms early in the course of the disease is important to identify disease state and severity in hematological malignancies (HMs). Furthermore, such symptoms can impact different aspects of a patient's life, such as physical, social, emotional and functional well-being, and often go undetected or are underestimated during the clinical consultation [3–5]. Although PROs in clinical practice can be very useful to overcome such issues, there are several barriers for their implementation, such as accessibility, userfriendliness, processing the collected information and finally the interpretation. Integration of an electronic version of PROs could offer a potential solution to the problem [6] by increasing accessibility to patients through mobile smart devices or a dedicated webpage. The information collected in such a manner would be easy to process and record, would provide immediate interpretation of the outcome to the clinicians and, in turn, back to the patients. In addition, the electronic versions also help in reducing the 'missing items' in the responses [7].

A systematic review conducted by Gwaltney *et al.* concludes that the electronic and the paper versions of PROs are equivalent and have the same measurement properties [8]. However, US FDA and other regulatory authorities require evidence to support such equivalence. Hence, adapting the paper version of the instrument to its electronic version requires demonstration of equivalent measurement properties so that both versions can be accepted and used interchangeably. The International Society of Pharmacoeconomics and Outcomes Research (ISPOR) provides recommendations on evidence needed to prove the equivalence between electronic and paper version of PROs [9].

The newly developed HM-PRO is a specific instrument for patients with HM in routine clinical practice to evaluate benefits–risks of treatment, capture unmet patient's needs and facilitate physician–patient communication, which is currently undergoing an assessment to meet the minimum standards set out by the US FDA. The use of an electronic application of a PRO measure not only provides a quick and most reliable assessment of quality of life (QoL), but would also enhance the integrity and accuracy of data and is also incited by regulatory authorities. The aim of this study was to evaluate the psychometric properties of the electronic version of the HM-PRO, in particular, reliability and face validity, in order to demonstrate equivalence between paper and electronic version of the instrument.

## Patients & methods

### Ethics

Multicenter ethics approval was obtained from the NRES South West Bristol, UK (ref 14/SW/0033) followed by individual R&D approvals from all the participating centers. Relevant research passport was obtained to carry out patient interviews at the hospital sites. Signed informed consent was obtained from all the study participants.

### Recruitment

This mixed method study was conducted in seven secondary care hospitals in the UK, following the ISPOR ePRO Good Research Practice Task Force Guidelines on evidence needed to support measurement equivalence between electronic and paper-based applications of a PRO instrument. The inclusion criteria were ability to give informed written consent; diagnosed with HM as per the most recent WHO classification [10]; state of the disease, defined as stable, progressing or remission; any stage of the treatment (due to start the treatment, on treatment and finished treatment); and able to read and write in English. A purposive sampling methodology was adopted to recruit patients with different HMs, disease state and treatment, both in- and out-patients. A randomized crossover design between the paper and the electronic versions of the HM-PRO was adopted. Patients were approached by the member of the clinical practice staff or by a member of the research team. Those who agreed to participate gave written informed consent. Anonymity of each patients was observed by allocating them a unique identification number.

### Instrument

The HM-PRO is a composite measure consisting of two scales: Part A (impact); and Part B (signs and symptoms). Part A measures the impact of the HM and its treatment on a patient's health-related QoL, and Part B captures the severity of different disease symptoms and treatment side effects. Based on factor analysis as well as RASCH modeling, Part A has a total of 24 items in four domains: physical behavior (7); social behavior (3); emotional behavior (11); and eating and drinking habits (3). Patients' responses are recorded on a three-point Likert scale (0 = not at all to two = a lot) and not applicable as a separate response option. Part B consists of 18 items in a single domain and the responses are captured on a three-point severity Likert scale (0 = not at all to two = severe) [11,12].

The third item of the ‘eating and drinking habits’ domain in Part A, in other words, ‘my drinking habits have changed’, and ninth item of Part B related to ‘skin problems’ are not included in the scoring system but collected for additional information. Although a misfit in the RASCH modeling, these two items were kept based on their high prevalence in the qualitative phase as reported important by the patients. The HM-PRO has shown good reliability with Cronbach’s  $\alpha$  and ICC greater than 0.8 for all four domains of Part A and for Part B. Using the anchor-based approach, individual score bands for Part A (0–7, 8–25, 26–41, 42–74, 75–100, representing ‘no impact’ to ‘very large impact’ on patients’ life) and Part B (0–3, 4–16, 17–29, 30–65, 66–100, representing ‘no effect’ to ‘very large effect’ of signs and symptoms on patients’ life) were established for translating the science of PROs in clinical practice. The bands are valid across gender and different age groups [13]. Furthermore, HM-PRO is capable of detecting small but clinically important change in patients’ HRQoL over time [14]. The minimum clinically importance difference (MCID) of ‘6’ was proposed based on standard error of the mean for both Part A and Part B of the HM-PRO [14].

The electronic version of HM-PRO was very similar to the paper version, except that each domain of Part A was displayed on a separate page and the whole of Part B on a single page. Patients are unable to skip any response and move to the next page. The general formatting, font type, font size, colour, instructions were similar in both the paper and the electronic version of the instrument (Figure 1).

### Data collection procedures & HM-PRO administration

All the patients received the study ‘patient information sheet’ and a flow chart of the study procedure prior to their participation. The study participants attending the outpatient clinics were also informed about the length of time they were required to stay in the clinic to complete the study. All patients gave written informed consent in person and completed the demographic information sheet. The paper and the electronic versions of the HM-PRO were then completed in a randomized crossover design with a 30 min time interval to minimize the learning effect. The random numbers were generated using RANDOM.ORG webpage before starting the patient recruitment in each center. Those who completed the paper version first, completed the electronic version after 30 min and vice versa. Ipads installed with the electronic application of the HM-PRO were provided to the patients. All the patients completed both the paper and electronic version of the HM-PRO on the same day.

In addition, a separate online survey was conducted in order to better understand the patients’ general impression and preference following completion of the paper-based and electronic application of the HM-PRO. The responses to the survey were automatically recorded electronically. Moreover, ten patients were randomly selected for cognitive interviewing. The interviews were audio recorded and transcribed verbatim.

### Statistical analysis

#### Statistical tests

All statistical analyses of the data collected were performed using IBM SPSS statistics software, version 23. The descriptive statistical analysis was performed on the demographic information. The scores of the paper and the electronic versions of the HM-PRO were calculated for the scales and the individual domains. The scores between the two versions of the HM-PRO were compared using the nonparametric Wilcoxon test for paired samples. A two-way ANOVA with patients as a random effect was used to test the ‘version effect’ of the HM-PRO (paper vs electronic), and ‘administration order effect’ (i.e., paper first vs electronic first). Spearman’s rank correlation coefficient (Spearman’s rho) and ICC were used to assess reliability between the two versions of the HM-PRO. Descriptive statistics were performed for the survey responses and content analysis of the transcribed interviews was performed using NVivo 11, a qualitative analysis software. In addition, analysis of ‘completion time’ was carried out between the paper and the electronic versions of the HM-PRO.

#### Study power

A sample size of 65 individuals is sufficient to verify if the ICC is greater than 0.75 with a 95% CI, considering a theoretical reliability of 0.8 [15].

## Results

### Demographic characteristics of the study participants

A total of 193 adult patients (62% male; 38% female) with the mean age of 63.4 years (SD = 15.18; range: 18–89 years) were recruited into the study (Table 1). The majority of the study participants consisted of out-

<

HM-PRO

Section A: The following statements describe your **physical behaviour**....

Please select the option which best describes how your haematological condition affects you and **how you feel about these things today**

I have difficulty with walking.

Not at all

A little

A lot

Not applicable

I have difficulty with self-care (e.g. dressing, bathing, etc.).

Not at all

A little

A lot

Not applicable

I have difficulty with physical activity/sports.

Not at all

A little

A lot

Not applicable

I have difficulty travelling (e.g. bus, train, flight and car).

Not at all

A little

A lot

Not applicable

I have difficulty leaving the house.

Not at all

A little

A lot

Not applicable

I have difficulty with work (or studies).

Not at all

A little

A lot

Not applicable

I have difficulty going on holidays.

Not at all

A little

A lot

Not applicable

>>

**Figure 1. Physical behavior domain of Part A of electronic HM-PRO.**  
HM-PRO: Electronic hematological malignancy-patient-reported outcome.

patients (93%). The ethnic origin of the patients was: 176 (91.2%) White; 14 (7.3%) Asian or Asian British; and three (1.6%) Black British. The diagnoses were: acute leukaemias – lymphoblastic (n = 12) and myeloid (n = 28); chronic leukaemias – lymphoid (n = 17) and myeloid (n = 13); multiple myeloma (n = 33); non Hodgkin lymphomas – indolent (n = 16) and aggressive (n = 22); Hodgkin lymphoma (n = 14); myelodysplastic syndromes (n = 15); and myeloproliferative neoplasms (n = 23) (Table 1). The mean time since diagnosis was 3.8 years (SD = 4.9; median = 1.7; range = 0.002–26 years). The majority of the patients received treatment:

**Table 1. Sociodemographics characteristics of the study participants (n-193).**

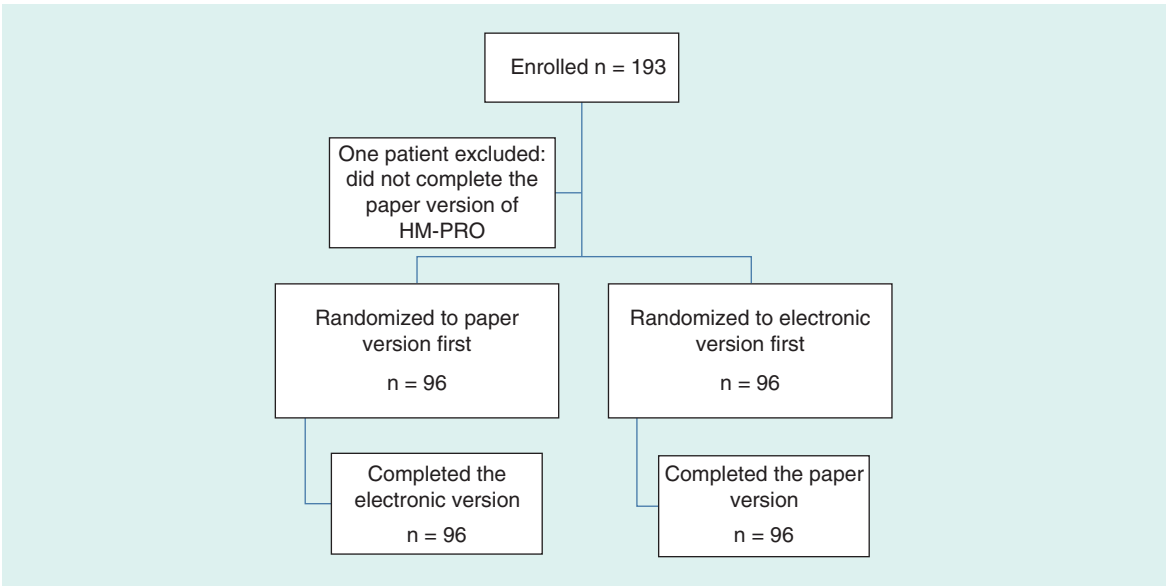
| n = 193                      |                        | Median | Range      |
|------------------------------|------------------------|--------|------------|
| Age (years)                  |                        | 66.5   | 17.9–89.1  |
| Time since diagnosis (years) |                        | 1.7    | 0.002–25.8 |
|                              |                        | n      | %          |
| Gender                       | Male                   | 119    | 61.7       |
|                              | Female                 | 74     | 38.3       |
| Inpatient/outpatient         | Inpatient              | 13     | 6.7        |
|                              | Outpatient             | 180    | 93.3       |
| Ethnic origin                | White                  | 176    | 91.2       |
|                              | Asian or Asian British | 14     | 7.3        |
|                              | Black British          | 3      | 1.6        |
| Disease type                 | ALL                    | 12     | 6.2        |
|                              | AML                    | 28     | 14.5       |
|                              | CLL                    | 17     | 8.8        |
|                              | CML                    | 13     | 6.7        |
|                              | MM                     | 33     | 17.1       |
|                              | INHL                   | 16     | 8.3        |
|                              | ANHL                   | 22     | 11.4       |
|                              | HL                     | 14     | 7.3        |
|                              | MDS                    | 15     | 7.8        |
|                              | MPN                    | 23     | 11.9       |
| State of disease             | Stable                 | 71     | 36.8       |
|                              | Remission              | 64     | 33.2       |
|                              | Progressing            | 58     | 30.1       |
| Employment status            | Employed full-time     | 61     | 31.6       |
|                              | Unemployed             | 7      | 3.6        |
|                              | Self employed          | 1      | 0.5        |
|                              | Homemaker              | 6      | 3.1        |
|                              | Retired                | 105    | 54.4       |
|                              | Unknown                | 10     | 5.2        |
|                              | Student                | 3      | 1.6        |

n: Number of patients recruited; SD: Standard deviation; AML: Acute myeloid leukaemia; ALL: Acute lymphoid leukaemia; CLL: Chronic lymphoid leukaemia; MM: Multiple myeloma; ANHL: Aggressive non-Hodgkin lymphoma; INHL: Indolent non-Hodgkin lymphoma; CML: Chronic myeloid leukaemia; MPN: Myeloproliferative neoplasm; MDS: Myelodysplastic syndromes; HL: Hodgkin lymphoma.

ten (5.2%) stem cell transplant; nine (4.7%) hydroxycarbamide; seven (3.6%) blood transfusions; seven (3.6%) R-CHOP chemotherapy; and 90 (46.7%) received other chemotherapy regimens. One (0.5%) patient was on wait and watch and one (0.5%) patient was due to start treatment. Furthermore, 68 (35.2%) were not on any treatment. The order of administration for the two versions of the HM-PRO was randomized using a standardized randomization procedure (Figure 2). One patient did not complete the paper version of the HM-PRO and therefore was excluded from the study.

### Health-related quality of life scores

A total of 192 pairs of completed questionnaires (electronic and paper version) were available for analysis. The summary statistics of scores for the four domains and both scales (Part A: impact, and Part B: signs and symptoms) are presented in Table 2. The mean difference between the scores of both versions was lowest for social behavior (0.02) and eating and drinking habits (0.02), whereas the highest difference was observed for the overall score of Part A (0.28). However, the difference in the mean scores of all the domains and the two scales was not statistically significant according to Wilcoxon signed rank tests for paired samples, even without Bonferroni correction for multiple testing (Table 2). This suggests that the scores are comparable for both the electronic and the paper versions of the HM-PRO in the sample. The Spearman's rank correlation coefficients were greater than 0.8 for all



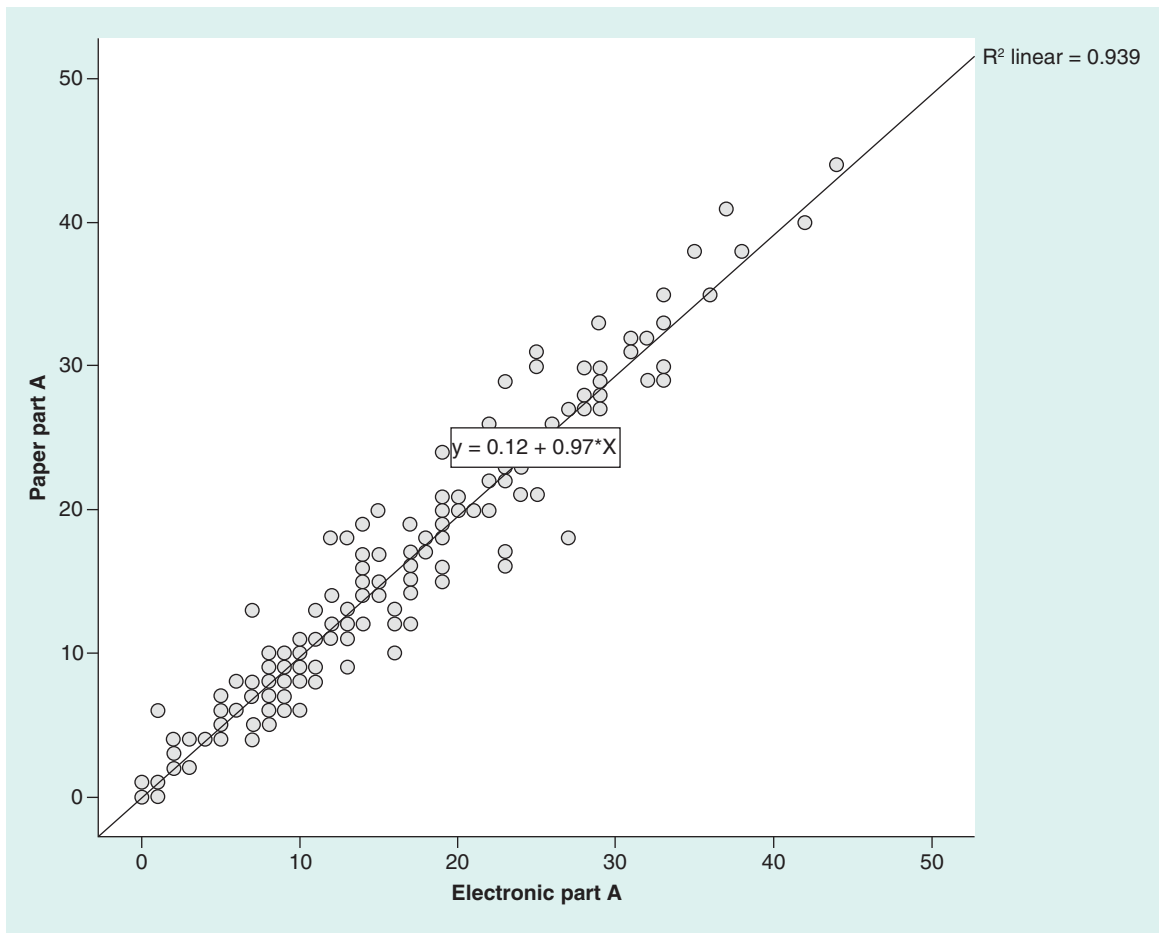
**Figure 2. The study participant flow chart.**  
HM-PRO: Hematological malignancy-patient-reported outcome.

| Table 2. Comparisons of the hematological malignancy-patient-reported outcome scores, reproducibility and reliability for the two versions. |                    |        |                         |       |                    |                    |                     |      |         |
|---------------------------------------------------------------------------------------------------------------------------------------------|--------------------|--------|-------------------------|-------|--------------------|--------------------|---------------------|------|---------|
| Domain/scale<br>(number of<br>items)                                                                                                        | Paper mean<br>(SD) | Range  | Electronic<br>mean (SD) | Range | Mean<br>difference | r(r <sub>s</sub> ) | ICC (95% CI)        | α    | p-value |
| Physical<br>behavior (7)                                                                                                                    | 4.8                | 2–8    | 4.94                    | 2–8   | 0.14               | 0.95 (0.95)        | 0.97<br>(0.97–0.98) | 0.97 | 0.17    |
| Social<br>behavior (3)                                                                                                                      | 1.5                | 0–2    | 1.50                    | 0–3   | 0.02               | 0.88 (0.88)        | 0.94<br>(0.92–0.96) | 0.94 | 0.45    |
| Emotional<br>behavior (11)                                                                                                                  | 8.1                | 4–11   | 8.19                    | 4–12  | 0.10               | 0.95 (0.95)        | 0.98<br>(0.97–0.98) | 0.98 | 0.40    |
| Eating and<br>drinking (2)                                                                                                                  | 1.4                | 0–2    | 1.41                    | 0–2   | 0.02               | 0.92 (0.93)        | 0.96<br>(0.95–0.97) | 0.96 | 0.52    |
| Part A<br>(impact) (23)                                                                                                                     | 15.6               | 7.5–23 | 15.89                   | 8–23  | 0.28               | 0.97 (0.97)        | 0.98<br>(0.98–0.99) | 0.98 | 0.07    |
| Part B (signs<br>and<br>symptoms)<br>(17)                                                                                                   | 7.5                | 4–10   | 7.64                    | 4–10  | 0.15               | 0.96 (0.96)        | 0.98<br>(0.97–0.98) | 0.98 | 0.22    |

ICC: Intraclass correlation coefficient; SD: Standard deviation; IQR: Inter quartile range; r: Pearson's correlation coefficient; r<sub>s</sub>: Spearman's correlation coefficient; α-Cronbach's α.

the domains and the two scales, ranging from 0.87 (social behavior) to 0.96 (Part A: impact). The ICC was greater than 0.9 for all the four domains and the two scales, showing a strong reliability (Table 2).

Being a randomized crossover design, a total of 96 patients completed the electronic version first and the remaining 96 completed paper version first. Hence, to assess the effects of the type of the questionnaire (electronic or paper) and the order of administration, using a random-effect two-way ANOVA, a total of 96 combination of cases were available. There were no interactions between the type of questionnaire and the order of the administration for both Part A ( $F_{1,96} = 0.002$ ;  $p = 0.95$ ) and Part B ( $F_{1,196} = 0.13$ ;  $p = 0.72$ ), respectively. Furthermore, subgroup analysis according to gender and age group also showed no interactions between the type of questionnaire and the order of the administration for both Part A (gender: females -  $F_{1,27} = 0.003$ ;  $p = 0.95$ ; males -  $F_{1,54} = 0.015$ ;  $p = 0.90$ , age group: 18–40 years -  $F_{1,6} = 0.04$ ;  $p = 0.84$ ; 40–60 years -  $F_{1,20} = 0.00$ ;  $p = 0.98$ ; 60–80 years -  $F_{1,45} = 0.00$ ;  $p = 0.93$ ; above 80 years -  $F_{1,10} = 0.02$ ;  $p = 0.87$ ) and Part B (gender: females -  $F_{1,27} = 0.05$ ;  $p = 0.82$ ; males -  $F_{1,54} = 0.14$ ;



**Figure 3.** Ordinary least square regression of individual scores of paper and the electronic version for Part A (impact).

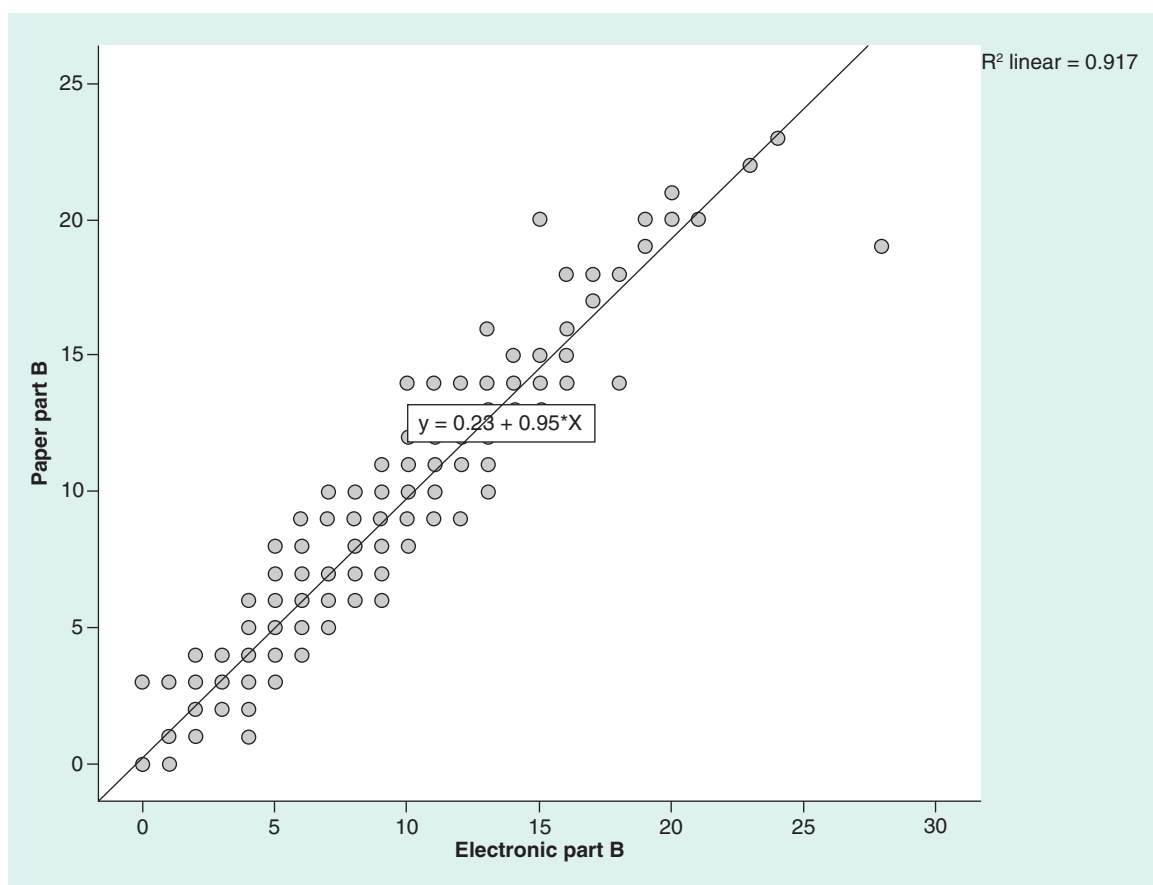
$p = 0.71$ , age group: 18–40 years -  $F_{1,6} = 0.005$ ;  $p = 0.94$ ; 40–60 years -  $F_{1,20} = 0.09$ ;  $p = 0.76$ ; 60–80 years -  $F_{1,45} = 0.00$ ;  $p = 0.99$ ; above 80 years -  $F_{1,10} = 0.21$ ;  $p = 0.64$ ).

Likewise, there was no effect of the type of questionnaire for Part A ( $F_{1,96} = 0.09$ ,  $p = 0.76$ ) and Part B ( $F_{1,96} = 0.80$ ,  $p = 0.78$ ). Furthermore, no effect of administration order was found for Part A ( $F_{1,96} = 0.099$ ;  $p = 0.752$ ) and Part B ( $F_{1,96} = 3.62$ ;  $p = 0.058$ ). This confirms that scores between the two versions of the HM-PRO were not statistically different, irrespective of their order of administration.

The distribution of scores for Part A and Part B obtained from both the paper and the electronic versions of the HM-PRO was analyzed using linear regression. The distribution of individual scores of Part A was:  $R^2 \text{ linear} = 0.94$  (Figure 3); and  $R^2 \text{ linear} = 0.92$  for Part B (Figure 4), accounting for more than 90% of the variance. This means that the coefficient of determination for the relationship between the scores of HM-PRO for the electronic and the paper versions is above 0.9 and only less than 10% of the variance in the scores might be due to other factors such as type of the instrument (electronic or paper) or change in the environment while completing the HM-PRO.

### The HM-PRO practicality & applicability

A total of 165 patients responded to the applicability and practicality questions, and the results are shown in (Table 3). The majority of the patients were able to read the statements easily, able to respond spontaneously and found the display on the screen of the electronic version understandable and easy to follow. Most 160 (97%) patients found the font size appropriate. Overall 143 (86.7%) showed preference toward completing the electronic version of the HM-PRO compared with its paper version. Only 20 (12.1%) preferred the paper version. Furthermore, the cognitive debriefing interviews confirmed that the patients were able to understand the items as intended and found responding to them straightforward.



**Figure 4.** Ordinary least square regression of individual scores of paper and the electronic version for Part B (signs and symptoms).

**Table 3:** Practicality and applicability of hematological malignancy-patient-reported outcome (n = 165).

| Question (n = 165)                                                           | Yes          | No          | M        |
|------------------------------------------------------------------------------|--------------|-------------|----------|
|                                                                              | n (%)        | n (%)       | n (%)    |
| Were the statements easy to read?                                            | 164 (99.4%)  | 0 (0%)      | 1 (0.6%) |
| Were you able to respond to the statements spontaneously?                    | 164 (99.4%)  | 0 (0%)      | 1 (0.6%) |
| Did you find the display on the screen understandable and easy to follow?    | 164 (99.4%)  | 0 (0%)      | 1 (0.6%) |
| Did the font size look appropriate?                                          | 160 (97.0%)  | 3 (1.8%)    | 2 (1.2%) |
| Did you face any difficulty while selecting a response option for each item? | 27 (16.4%)   | 136 (82.4%) | 2 (1.2%) |
| Would you prefer to fill out the electronic version or the paper version?    | 143 (86.7 %) | 20 (12.1%)  | 2 (1.2%) |

n = Number of patients; M = Missing response

### The HM-PRO completion time

A total of 100 patients recorded completion time for the paper version of the instrument, whereas the completion time for the electronic version was recorded automatically for all 193 patients. Therefore, 100 patients whose completion time was available for both versions were included in the analysis. The median time to complete the paper version of the HM-PRO was 5 min (mean = 6.4 min, range = 2.0–25.0 min, interquartile range = 4.2 min). For the electronic application of the HM-PRO, the median completion time was 6.5 min (mean = 7.3 min, range = 1.27–16.17 min, interquartile range = 3.6). The difference between the completion time for paper and electronic version of the HM-PRO was not statistically significant ( $p = 0.114$ ). Although statistically insignificant, the completion time for the electronic version was slightly higher than the paper version because each domain is

displayed on a separate page, which requires additional number of screen touch, and filling in the comment box electronically consumes more time than writing on the paper version.

## Discussion

According to the ISPOR ePRO Good Research Practice Task Force report [9], adapted from Shields *et al.* [16], the rationale of a minor change is that ‘the modification can be justified on the basis of logic and/or existing literature. No change in content or meaning’. The electronic version of the HM-PRO is no different with respect to its content (including the instruction for completion) or meaning to the paper-based version of the instrument. The only difference in the electronic version is that the items are presented by one domain per a page, which is under the definition of the ‘minor changes’ according to the ePRO task force report. The level of evidence recommended for such minor changes are ‘cognitive debriefing and usability testing’ [9]. This study goes beyond the standard recommendations and assesses the measurement equivalence by calculating the ICC, comparison of mean scores, Cronbach’s  $\alpha$  and test–retest reliability. The design of this study was chosen as randomized crossover design based on task force recommendations so to have greater statistical power even with a smaller sample. The demographics of the recruited patients showed a good mix of patients with different diagnoses and different disease states which confirms that the study sample is representative of the target population, in other words, patients with HMs; hence, the heterogeneity in the sample should be considered as a strong point. Furthermore, the age range of 18–89 years shows that the HM-PRO has been validated in a vast age distribution.

The cognitive debriefing interviews are performed to assess that the items are perceived in the same way irrespective of the version of the instrument. The cognitive debriefing of ten patients confirmed that the items are understood and responded as intended. No patient reported any difference between the versions of the instrument.

The findings suggest that the electronic version of the HM-PRO fulfills the requirements of measurement equivalence with the paper version of the instrument. In fact, the intraclass correlation was greater than 0.9 and the correlation coefficients were greater than 0.8 for all domains of PART A and for both scales, Part A and Part B, confirming good test–retest reliability for all the domains and the scales. The difference in mean score between the paper and the electronic version was insignificant even before Bonferroni corrections for multiple testing, further confirming reliability. The practicality testing of the eHM-PRO confirmed that it is easy to follow, understand and respond spontaneously. The general outlook of the electronic version was also found to be userfriendly. Interestingly, 95% of the study participants preferred the electronic version of the HM-PRO compared with the paper version.

The usability testing of the electronic version of the HM-PRO confirmed that it is easy to follow, understand and respond spontaneously. The general outlook of the electronic version was also found to be userfriendly. Interestingly, 87% of the study participants preferred the electronic version of the HM-PRO compared with the paper version.

The electronic version of the HM-PRO can be of great significance not only in daily clinical practice but also in clinical research. Patients’ change in scores can be incorporated in the healthcare records and monitored over time very easily. The electronic version allows the immediate calculation and interpretation of scores; hence, its use would be very useful at the time of consultation to facilitate the communication between the doctor and the patients. The individual signs and symptoms in Part B can be discussed based on the severity reported by the patients. Further, patients as well as clinicians could more readily use the electronic version of the HM-PRO to keep a record as part of the general electronic health record to monitor changes in their QoL over time in relation to the treatment they are receiving at any given time. The additional benefit of the electronic version of HM-PRO for all the stakeholders would be its straightforward incorporation into the patient’s health records which could be accessed at the point of care during the consultation. The patients can complete it in the comfort of their home prior to attending the outpatient clinic. Such an approach would also facilitate the periodic review of such patients’ records which is considered of paramount importance in regulating the treatment and other needs by the care team.

One bias of the study may be the difficulty to recruit patients who were not comfortable with using an electronic device, though the number of such patients was very small. Furthermore, in the paper version patients could respond to the items randomly compared with the electronic version where the responses were to be given in the presented order. However, random response in the paper version could not be detected and was not observed even while calculating the differences in the mean scores.

## Conclusion

The electronic application of the HM-PRO is equivalent to the paper-based version for its measurement properties, with minor changes as per the ISPOR ePRO good research practice task force report. The electronic version has

good reliability, face validity and comparable scores to the paper version of the HM-PRO. The interest of patients in completing the electronic version of the instrument over the paper version is an interesting finding and would help the implementation of the electronic version of HM-PRO both in daily clinical practice and clinical research.

### Summary points

- The electronic version of the hematological malignancy-patient-reported outcome (HM-PRO) is equivalent to the paper version for its measurement properties, with minor changes as per the The International Society of Pharmacoeconomics and Outcomes Research electronic PRO good research practice task force report.
- The electronic version has good reliability, face validity and has comparable reproducibility of scores.
- The Spearman's correlation coefficient was greater than 0.8 between the score of the two versions of the instrument suggesting that they are highly related.
- Administration order and type of questionnaire effects were insignificant.
- The cognitive debriefing of results confirmed that the items are understood and responded as intended.
- The patients' preference of the electronic version of HM-PRO over the paper version is an interesting finding and it is hoped to facilitate the use of eHM-PRO in daily clinical practice and clinical research.

### Acknowledgments

The authors would like to thank the input of all the study participants and participating centers for their support throughout the study.

### Financial & competing interests disclosure

The study was funded by the European Haematology Association Scientific Working Group 'Quality of life and Symptoms' through unrestricted grants from Novartis, Bristol Myers Squibb and Sanofi. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

### Ethical conduct of research

Multicentre ethics approval was obtained from the NRES South West Bristol, UK (ref 14/SW/0033) followed by individual R&D approvals from all the participating centers. Signed informed consent was obtained from all the study participants.

### Open access

This work is licensed under the Creative Commons Attribution 4.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

### References

1. Dobrozi S, Panepinto J. Patient-reported outcomes in clinical practice. *Hematology Am. Soc. Hematol. Educ. Program* 2015, 501–506 (2015). [www.ncbi.nlm.nih.gov/pubmed/26637765](http://www.ncbi.nlm.nih.gov/pubmed/26637765)
1. Black N, Burke L, Forrest CB *et al.* Patient-reported outcomes: pathways to better health, better services, and better societies. *Qual. Life Res.* 25(5), 1103–1112 (2016).
2. Reiser SJ. The clinical record in medicine. Part 1: learning from cases. *Ann. Intern. Med.* 114(10), 902–907 (1991).
3. Basch E. The missing voice of patients in drug-safety reporting. *N. Engl. J. Med.* 362(10), 865–869 (2010).
4. Basch E, Abernethy AP. Supporting clinical practice decisions with real-time patient-reported outcomes. *J. Clin. Oncol.* 29(8), 954–956 (2011).
5. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. *Am. J. Manag. Care* 14(8), 530–539 (2008).
6. Wu AW, Kharrazi H, Boulware LE, Snyder CF. Measure once, cut twice—adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. *J. Clin. Epidemiol.* 66(8 Suppl.), S12–S20 (2013).
7. Bernhard J, Cella DF, Coates AS *et al.* Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat. Med.* 17(5–7), 517–532 (1998).
8. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health* 11(2), 322–333 (2008).

9. 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* 12(4), 419–429 (2009).
10. Swerdlow SH, Campo E, Harris NL *et al.* (Eds). Summary Table. In: *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. (4th Ed.), Volume 2, IARC Publications, Lyon, France (2017).
11. Goswami P, Salek S, Fielding AK *et al.* Development of a novel patient-reported outcome measure in haematological malignancy for use in routine clinical practice: item generation. *Blood* 128(22), 5985–5985 (2016).
12. Goswami P, Salek S, Ionova T *et al.* HM-PRO: a novel patient-reported outcome measure in hematological malignancy for use in clinical practice. *Blood* 130(Suppl. 1), 2176–2176 (2017).
13. Goswami P, Oliva EN, Ionova T, Salek S. Translating the science of patient reported outcomes into practice: meaningfulness of HM-PRO scores in patients with hematological malignancies. *Blood* 132(Suppl. 1), 4860–4860 (2018).
14. Goswami P, Oliva EN, Ionova T, Salek S. Responsiveness and the minimal clinically important difference for HM-PRO in patients with hematological malignancies. *Blood* 132(Suppl. 1), 2294–2294 (2018).
15. Streiner DL, Norman GR. Reliability. *Health Measurement Scales a Practical Guide to their Development and Use*. (3rd Ed.) Oxford University Press, Oxford, UK, 12–28 (2003).
16. Shields A, Gwaltney C, Tiplady B, Paty J, Shiffman S. Grasping the FDA's PRO guidance. *App. Clin. Trials* 15, 69–72 (2006).

