1   **Validation of the French Version of the Integrated Palliative care Outcome Scale (IPOS)**

2   Anca-Cristina Sterie, PhD, Gian Domenico Borasio, Prof., Mathieu Bernard, PhD, and the French

3   IPOS Consortium*

4   *Palliative and supportive care service, Lausanne University Hospital and University of Lausanne,*

5   *Switzerland*

6

7   *Michel Beauverd, MD, Christian Bernet, MD, Boris Cantin, MD, Nathalie Dieudonné Rahm, MD,

8   Monica Escher, MD, Sylvie Jeanneret-Brand, MD, Floriana Lurati, MD, Gérard Pralong, MD, Josianne

9   Pralong, MD, Tony Tai, MD, Emmanuel Tamchès, MD, Anne Vacanti, MD, and Gilbert Zulian, MD,

10   *Palliative and supportive care service (M.B., E.T.), Lausanne University Hospital, Lausanne,*

11   *Switzerland; Palliative Care Unit (B.C.), Fribourg Hospital, Fribourg, Switzerland; Pain and Palliative*

12   *Care Consultation (M.E.), Division of Pharmacology and Toxicology, University Hospital of Geneva*

13   *and Geneva University, Geneva, Switzerland; Palliative Medicine Service (N.D.R., G.Z.), Bellerive*

14   *Hospital, University Hospital of Geneva and Geneva University, Geneva, Switzerland; Palliative Care*

15   *Center "La Chrysalide" (S.J.-B., C.B.), Neuchâtel Hospital, Neuchâtel, Switzerland; Palliative Mobile*

16   *Team (F.L.), Haut Léman Health Network, Blonay, Switzerland; Inter-Hospital Palliative Mobile Team*

17   *(F.L.), Riviera-Chablais Hospital, Blonay, Switzerland; Palliative Care Unit (G.P.), Lavaux Hospital,*

18   *Cully, Switzerland; Palliative Care Unit (J.P.), Rive-Neuve Foundation, Blonay, Switzerland; Palliative*

19   *Mobile Team (T. T.), North Broye Health Network, Orbe Hospital, Orbe, Switzerland; Palliative Care*

20   *Unit (A.V.), Valais Hospital, Martigny, Switzerland.*

21

23

24   *Address correspondence to*: Mathieu Bernard, PhD, Palliative and Supportive Care Service,

25   Lausanne University Hospital, Avenue Pierre-Decker 5; 1011 Lausanne, Switzerland. E-mail:

26   mathieu.bernard@chuv.ch

27 **Abstract**

28 **Context.** The Integrated Palliative care Outcome Scale (IPOS) is a widely used tool for assessing
29 patient needs in palliative care.

30 **Objectives.** The aim of this study is to provide a validated version of the patient and staff IPOS for
31 French-speaking Switzerland (IPOS-Fr) and assess its psychometric properties.

32 **Methods.** The validation took place in 12 palliative care units and mobile teams. At baseline (T1) and
33 three days later (T2), patients' general health status, palliative care needs (IPOS-Fr) and quality of life
34 (McGill Quality of Life scale Revised-MQOL-R) were assessed by patients and staff.

35 **Results.** We included 173 patients (mean age: 68.8; 92 women; 85% oncologic disease). IPOS internal
36 consistency was high for the total score (.69 and .71). Staff-patient inter-rater agreement was good to
37 moderate for 13 items (intra-class correlations >.516). Results indicated strong correlations between
38 IPOS-Fr and MQOL-R for the total score (-.623 at T1) and the psychological domain (item 11:-.601 at
39 T1; item 13: -.633 at T2). Regarding sensitivity to change, there was a significant difference between
40 T1 and T2 for patients with an improved health condition (z=-2.326; p=.020).

41 **Conclusion.** IPOS-Fr has fair to good validity, especially with regard to inter-rater agreement and
42 construct validity, is sensitive to positive change, and has good interpretability and acceptability for
43 patients and staff. IPOS-Fr is not optimal in terms of internal consistency and structure when using
44 subscale scores, except for the emotional subscale.

45

46

Introduction

The Palliative care Outcome Scale (POS) was designed for evaluating essential outcomes in palliative care, and has demonstrated validity(1). The Integrated Palliative care Outcome Scale (IPOS)(2), an advancement of POS(3), is composed of 10 questions and exists in patient and staff versions, to be completed within a 3 or 7-day recall period. IPOS embraces a holistic perspective by evaluating patients' physical, emotional, spiritual, and communicational needs. A Rasch analysis of IPOS supported its use as a clinical and research measure(4).

IPOS's 17 items are scored with a Likert scale (from "0", not affected, to "4", extremely affected). For items 14-16, the Likert scale options were reversed and data was re-scaled. According to the POS development team, items can be considered independently, as subscales (physical symptoms, items 1-10; emotional symptoms, items 11-14; problems and communication, items 15-17), or summed to yield a total score (range 0-68). Open comments about additional symptoms, a question assessing how the patient filled the questionnaire, and the staff Likert option "cannot assess" are not considered for score calculation.

IPOS already has several translations(2, 5-6). French is ranked the fifth most widely spoken language in the world(7). Having already performed its cross-cultural adaptation to French(7,8), our aim was to provide a psychometrically validated version of IPOS in French (IPOS-Fr).

Methods

Participants and procedure

The study was performed between January 2017 and February 2018 in seven palliative care units (PCUs) and five mobile palliative care teams in French-speaking Switzerland. Inclusion criteria were patients ≥18 years old, good comprehension of French, stable condition over the past day. Exclusion criteria were impaired mental capacity according to the clinical judgement of the referring physician or existing diagnosis and evidence of psychiatric disease affecting decision-making capacity.

Eligible patients provided informed consent and filled IPOS-Fr three or more days after admission for palliative treatment (T1). In parallel, staff IPOS-Fr was completed by a referring palliative specialist (physician, nurse, psychologist, or specialized nursing auxiliary). If possible, a second assessment was performed three days after (T2).

Missing data strategies

Psychometric analysis was performed according to seven scenarios for dealing with missing data (MD), and estimated that the best strategy for calculating subscale and total score was the subscale median imputation for up to one MD per subscale (see table 1 supplementary material). Admitting more MD would require too much interpretation. This strategy allowed to include most participants (169/160 valid

82   cases at T1 for patients/staff, 108/102 at T2) and corresponded to the non-normal distribution of the

83   dataset (after performing the Shapiro-Kolmogorov test).

84   *[Insert table 1 supplementary material-Scenarios for dealing with MD]*

85   Reliability

86   For reliability measures we considered only values at T1, given that at T2 patients might have been

87   biased by prior knowledge of the items. *The internal structure* of patient and staff IPOS-Fr was tested

88   with a factorial analysis using varimax rotation. The *internal consistency* of patient and staff IPOS-Fr

89   versions was measured by calculating Cronbach's alpha for the total scales at T1 and for the factors

90   revealed in the factorial analysis. Cronbach alpha was recalculated by excluding each item one at a

91   time, in order to evaluate the precise influence of each item on the identified subscale. Acceptable

92   values range from 0.7 to 0.95(9). We then compared these results with the Cronbach's alphas

93   calculated from the subscales proposed by the original version.

94   For *inter-rater agreement*, we calculated intraclass correlations (ICC) between IPOS-Fr staff and patient

95   scores at T1 on individual and subscale scores. Using the averaged reliability of different raters, we

96   considered values <0.5, between 0.5 and 0.75, between 0.75 and 0.9, and >0.9 as indicative of poor,

97   moderate, good, and excellent reliability(10).

98   Construct validity

99   Construct validity was tested through Pearson correlations between IPOS and the McGill Quality of Life

100   scale-Revised version (MQOL-R). It contains 14 items forming four subscales: physical, psychological,

101   existential, and relationship. We checked correlations between IPOS individual and total scores, and

102   MQOL-R subscale and total scores. We considered values r>.50 as indicator of strong to exceptional

103   association; .40<r<.50 as indicator of medium association; and r<.40 as indicator of poor to inexistent

104   association(11). We expected negative correlations since IPOS-Fr displays need for palliative care and

105   MQOL-R displays patients' quality of life.

106   Sensitivity to change

107   We compared the consistency of patient and staff IPOS-Fr scores at T1 and T2 with the consistency of

108   their evaluation of the patients' condition ("How do you evaluate your general health state?") using

109   Wilcoxon's non-parametric test. This allowed categorizing patients in a "stability", "improvement", or

110   "deterioration" group. The hypothesis was that IPOS-Fr score would not change for patients of the

111   "stability" group, but would for the others.

112   Interpretability and acceptability

113   These two aspects were assessed through analysis of the free text in IPOS-Fr comments and through

114   measure of required time to complete IPOS-Fr.

115 Ethics

116 The study was approved by the Research Ethics Committee of the canton of Vaud, Switzerland, with
117 patients' written agreement.

118 **Results**

119 *Descriptive results*

120 173 patients and 169 staff completed IPOS-Fr at T1, and 108 patients and 102 staff at T2. The
121 difference in numbers between T1 and T2 is due to worsening state or departure (see table 2
122 supplementary material). Recruitment and participation was higher in PCUs.

123 *[Insert table 2 supplementary material-Participants' characteristics]*

124 At baseline, mean item scores ranged from 0.4 for item 5 to 2.3 for item 12 for patients, and from 0.3 to
125 2.5 for the same items for staff (see table 1).

126 *[Insert table 1-Mean symptom intensity and scores]*

127 *Missing data*

128 At T1, 78% of patient and 69% of staff had no MD; at T2, 60% and 72% respectively (see table 3
129 supplementary material).

130 *[Insert table 3 supplementary material-MD at T1 and T2]*

131 Items 12, 15, and 17 had most MD; the first two were highlighted during the cross-cultural adaptation
132 as potentially difficult to understand(8) (see table 4 supplementary material).

133 *[Insert table 4 supplementary material-Frequency of MD]*

134 *Reliability*

135 Internal structure

136 The factorial analysis with varimax rotation revealed six factors with an eigenvalue ≥1 explaining 61%
137 of the total variance for patient IPOS-Fr, and five such factors explaining 59% for staff IPOS-Fr (see
138 table 5 and 6 supplementary material). The three-subscale pattern of IPOS was not confirmed.
139 However, for patients factor 4 is identical to the problems and communication subscale and for staff,
140 factor 1 to the emotional subscale.

141 *[Insert table 5 supplementary material-Factorial analysis]*

142 *[Insert table 6 supplementary material-Correlations between IPOS-Fr items and factors]*

143   Internal consistency

144   Cronbach's alpha was .69 and .71 for total scores. Cronbach's alpha for factors 4, 5, and 6 for patients
145   and 4 and 5 for staff were lower than .70. No single item was essential to guarantee the subscales'
146   consistency (see table 7 supplementary material).

147   *[Insert table 7 supplementary material-Cronbach's alpha for factors]*

148   For the subscales, Cronbach's alpha varied between .34 and .81 (see table 8 supplementary material).

149   *[Table 8 supplementary material-Cronbach's alpha for subscales]*

150   Inter-rater consistency

151   ICC coefficients indicated good reliability for item 2, moderate for items 1, 3-7, 9-11, 13, 14, 17, and for
152   the three subscales, and poor for items 8, 12, 15, 16 (see table 9 supplementary material).

153   *[Insert table 9 supplementary material-Intra-class correlations]*

154   ***Construct validity***

155   At T1, our results indicate strong correlations between MQOL-R and patient IPOS-Fr for the total score,
156   the psychological domain (IPOS-Fr item 11 and 13), and the social subscale (item 15). At T2,
157   correlations were medium to weak for the physical subscale, the existential domain (IPOS-Fr item 14)
158   and the social domain (IPOS-Fr item 15) (see table 10 supplementary material).

159   *[Insert table 10 supplementary material-Pearson's correlations]*

160   ***Sensitivity to change***

161   The data show a significant difference between T1 and T2 for the "improvement" group, but not for the
162   "stability" and "deterioration" groups (see table 11 supplementary material).

163   *[Insert table 11 supplementary material-Sensitivity to change]*

164   ***Mobile team vs PCU setting***

165   Regarding Cronbach's alpha, in the mobile team setting patient IPOS-Fr systematically scored lower
166   than in PCU; for staff, it was the opposite. Stronger ICC correlations were found for the emotional
167   subscale (PCUs) and the problems and communication subscale (mobile team). Correlations between
168   IPOS-Fr and MQOL-R did not change for PCUs while for the mobile setting the only significant
169   correlations were between MQOL-R psychological subscale and item 13 (T1 and T2) and 11 (at T1).

170   ***Interpretability and acceptability***

171 Patients completed IPOS-Fr in one day, mostly in one time (97% at T1), in less than 20 minutes (68%),
172 aided by staff (56% at T1).

173 At T1 and T2, 45 patients made overall comments regarding IPOS-Fr: 23 noted its usefulness and
174 clarity, while 33 made precisions concerning the assessment of symptoms.

175 At T1 and T2, 20 staff members reported comments about IPOS-Fr. Three noted that questions are
176 useful and interesting. Four considered IPOS-Fr too long or inadequate for patients, three found the
177 Likert scale imprecise, seven noted the difficulty in evaluating items, three suggested more attention to
178 goals of care.

179 **Discussion**

180 Our study reports results on IPOS-Fr's psychometric validation based on a large sample of patients
181 representative of the French-speaking palliative care context.

182 Regarding IPOS's internal reliability, the three-subscale structure of the original IPOS was not,
183 originally, backed by a psychometric validity and was not confirmed by a Rasch analysis that highlighted
184 the existence of several "super-items"(3). The factorial analysis that we performed on patient and staff
185 IPOS-Fr revealed six and five main factors, respectively, and therefore did not confirm the three
186 subscale structure of IPOS, even though factor 4 for patients corresponded to the problems and
187 communication subscale (items 15-17) and factor 1 for staff corresponded to the emotional subscale
188 (items 11-14). While some items could be removed in order to create new subscales, this is impossible
189 due to their clinical importance but also because as a translated version, IPOS-Fr cannot significantly
190 differ in items from the original version. Regarding our factors, additional elements do not speak in favor
191 of their validity: (i) the fact that the reduction of the information is not very important (from 17 items to 6
192 and 5 factors respectively, leaving approximately 40% of the variance unexplained), (ii) the
193 heterogeneity of the items' number per factor, (iii) the fact that a common point between items is
194 sometimes difficult to highlight, and (iv) finally the fact that several factors clearly focus on the same
195 aspect (three factors concern the physical area in both IPOS patient and staff).

196 In addition, when looking at the internal consistency of our factors, half of them showed insufficient
197 values (below .45) from the patient IPOS-Fr, which is also a reason not to recommend the use of our
198 subscales. Similar results were obtained with the staff IPOS-Fr. Results were better when considering
199 the internal consistency calculated from the original three-subscale structure but, once again, this
200 structure was not confirmed by our factorial analysis. We therefore conclude that the use of any
201 subscale is not advisable for IPOS-Fr and we recommend the use of either the total score or individual
202 items.

203 In terms of inter-rater agreement, our results showed that staff and patient views on symptoms and
204 outcomes are globally similar, except for item 8 ("sore or dry mouth"), two items involving the relatives
205 (items 12 "anxiety of close ones" and 15 sharing of feelings"), and interestingly, the item 16 assessing

206 the satisfaction with the transmitted information. Differences in staff and patient interpretation were
207 revealed during the adaptation phase(8).

208 In terms of construct validity, similar to the POS(1) and its translations(12; 13), IPOS showed good
209 patient-staff agreements. Patient IPOS-Fr showed strong correlations(11) with the MQOL-R for the total
210 score and the psychological domain. Weaker but still moderate correlations were found for the physical
211 domain, the existential domain, and the social domain at T2. These lower correlations may be explained
212 by the fact that IPOS-Fr does not allow for a complete and in-depth evaluation of these dimensions,
213 except perhaps for the emotional dimension which evaluates both depression and anxiety outcomes,
214 the most frequent psychiatric manifestations in the palliative care context(14).

215 Results showed that patient and staff IPOS-Fr are sensitive enough to detect improvement of patient's
216 condition. This is encouraging knowing that the formed groups have a relatively similar profile in terms
217 of palliative care needs and that there is overall little evolution of their health state. As reflected through
218 the difference in patient population at T1 and T2, it remains challenging to assess sensitivity to change
219 in this context.

220 Results in terms of interpretability and acceptability are rather encouraging within this francophone
221 population and their staff. Nevertheless, its clinical applicability might be affected by the fact that some
222 specific items showed more missing data than others (items 12, 15, 17) and that most patients required
223 the aid of a member of the staff. Moreover, the clinical applicability might also be affected by the context,
224 as mobile teams reported more difficulty than PCUs in recruiting patients (only 18% of patients were
225 recruited through mobile teams) and in ensuring that the questionnaire was filled in on the same day by
226 patient and staff. A possible cause of this disparity resides on the fact that mobile teams are smaller,
227 and therefore had less opportunities for ensuring that, during an intervention, one professional can aid
228 the patient to complete the IPOS patient and another one can fill in the IPOS staff. In addition, most of
229 the time, mobile teams intervene in critical moments, so fewer of their patients met the "stability"
230 inclusion criteria.

231 This study has several limitations. First, we had to employ a missing data strategy, which requires a
232 degree of interpretation. Tolerating one MD per subscale meant that the total score was calculated with
233 up to three MD, which is not optimal because it means that we have accepted up to 17.5% of MD (3
234 items on 17 in total). Second, we could only include patients who had been in a stable condition over
235 the past day, generating a selection bias and floor effect in a pool of relatively well-faring patients. Scant
236 data for the mobile context and lack of inclusion of other settings limit the generalizability of the results.

237 **Conclusions**

238 IPOS-Fr demonstrated fair to good inter-rater agreement and construct validity, is sensitive to positive
239 change, and has good interpretability and acceptability. IPOS-Fr is not optimal in terms of internal
240 consistency and structure when using subscale scores. We recommend the use of total or single item

241 scores in both research and clinical settings. Health care professionals should be familiar with this tool,
242 but also aware of its limitations.

References

248 (1) Hearn J, Higginson IJ. Development and validation of a core outcome measure for palliative care:
249 the palliative care outcome scale. Palliative Care Core Audit Project Advisory Group. Qual Health
250 Care 1999;8:219-27.
251 (2) https://pos-pal.org/maix/ipos_in_english.php
252 (3) Schildmann EK, Groeneveld EI, Denzel J, Brown A, Bernhardt F, et al. Discovering the hidden
253 benefits of cognitive interviewing in two languages: The first phase of a validation study of the
254 Integrated Palliative care Outcome Scale. Palliat Med 2016;30:599-610.
255 (4) Sandham MH, Medvedev O, Hedgecock E, Higginson IJ, Siegert RJ. A Rasch Analysis of the
256 Integrated Palliative Care Outcome Scale (IPOS). J Pain Symptom Manage 2018;5: 290-296.
257 (5) Sakurai H, Miyashita M, Imai K, Miyamoto S, Otani H, et al. Validation of the Integrated Palliative
258 care Outcome Scale (IPOS) – Japanese Version. Jpn J Clin Oncol 2019;49:257-262.
259 (6) Antunes B, Ferreira PL. Integrated palliative care outcome scale: Protocol Validation for the
260 Portuguese population. Revista Cuidados Paliativos 2016;4:65-102.
261 (7) Organisation internationale de la francophonie. La langue francaise dans le monde, Paris, 2014,
262 (8) Sterie AC, Bernard M. Stakes in questionnaire translation: the cross-cultural adaptation of the
263 Integrated Palliative Care outcome Scale. BMC Palliat Care (in press).
264 (9) Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011;2:53-5
265 (10) Koo TK, Li MY. 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients
266 for Reliability Research. J Chiropr Med 2016;15:155-63
267 (11) Gendre F. L'analyse statistique multivariée. Genève, Droz, 1976.
268 (12) Eisenchlas JH, Harding R, Daud ML, Perez M, De Simone GG, et al. Use of the palliative outcome
269 scale in Argentina: a cross-cultural adaptation and validation study. J Pain Symptom Manage
270 2008;35:188-202
271 (13) Pukrittayakamee P, Sapinum L, Suwan P, Harding R. Validity, Reliability, and Responsiveness of
272 the Thai Palliative Care Outcome Scale Staff and Patient Versions Among Cancer Patients. J Pain
273 Symptom Manage 2018;56:414-20
274 (14) Mitchell AJ, Chan M, Bhatti H, Halton M, Grassi L, et al. Prevalence of depression, anxiety, and
275 adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94
276 interview-based studies. The Lancet Oncol 2011;12:160-74.