

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/14137690>

Ten-Language Translation and Harmonization of the International Prostate Symptom Score: Developing a Methodology for Multinational Clinical Trials

Article in *European Urology* · February 1997

DOI: 10.1159/000474438 · Source: PubMed

CITATIONS

125

READS

394

3 authors, including:



M. Garcia-Losa

IMS Health

15 PUBLICATIONS **307** CITATIONS

[SEE PROFILE](#)



Rafael Dal-Re

Universidad Autónoma de Madrid

232 PUBLICATIONS **2,576** CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Written informed consent in low-risk pragmatic RCT [View project](#)

X. Badia^a
M. García-Losa^b
R. Dal-Ré^b

^a Institute of Public Health of Catalonia,
University of Barcelona, and

^b Medical Department, SmithKline Beecham
Pharmaceuticals, Madrid, Spain

Ten-Language Translation and Harmonization of the International Prostate Symptom Score: Developing a Methodology for Multinational Clinical Trials

Key Words

International Prostate Symptom Score
Benign prostatic hyperplasia, symptoms
Questionnaire translation and harmonization
Clinical trials, multinational

Abstract

Objective: Production of cross-culturally adapted and harmonized translations of the International Prostate Symptom Score (I-PSS) into 10 languages (Canadian French, Danish, Dutch, Flemish, French, German, Italian, Norwegian, Spanish, Swedish) by means of an appropriate methodology. **Methods:** A method based on state-of-the-art guidelines about translation of questionnaires was developed which consisted of forward-translation, rating for clarity, common language and cultural adequacy, group discussion and consensus, testing for comprehension in lay subjects, back-translation and rating for conceptual equivalence. **Results:** The study lasted 3 months. Items 1-7 scored high for clarity, common language and cultural adequacy (mean range 5.3-9.8) as also did item 8 (5.1-9.7). Only 10% of items were rated less than 5.0. Comprehension tests in lay subjects produced comments in 10% of items, most frequently in items 2, 7, and 8, and response categories for items 1-6. Comparisons between the back-translations and the source version showed that 84% of all items, 88% of response categories for items 1-6, 100% of response categories for item 7, and 75% of response categories from item 8 were conceptually equivalent. **Conclusion:** The present set of 10 translations plus the original I-PSS questionnaire can be used in multi-country studies in which benign prostatic hyperplasia symptoms are to be assessed using the I-PSS and the data collected from different countries can be legitimately aggregated.

This study was supported by a grant from SmithKline Beecham Pharmaceuticals, Harlow, UK.

KARGER

E-Mail karger@karger.ch
Fax +41 61 306 12 34
<http://www.karger.ch>

© 1997 S. Karger AG, Basel
0302-2838/97/0312-0129\$12.00/0

This article is also accessible online at:
<http://BioMedNet.com/karger>

Manuel García-Losa, MD
Medical Department, SmithKline Beecham Pharmaceuticals
Valle de la Fuenfria, 3
E-28034 Madrid (Spain)
Tel. +34 1 734 6565, Fax +34 1 372 1490

Introduction

Benign prostatic hyperplasia (BPH) causes morbidity through the urinary symptoms associated with this condition [1]. Relief of those symptoms is the single most relevant outcome to the patient [2]. Although a weak correlation between symptoms and other commonly used clinical markers of BPH has been found, symptom severity correlated better than other measures with overall health status [3, 4]. The development of symptom indexes to capture symptom severity in BPH patients has been pursued since the late 1970s [5], and the US Agency for Health Care Policy and Research recommends that patient symptoms be quantified [6]. However, it was not until the Measurement Committee of the American Urological Association (AUA) developed the AUA BPH symptom index (AUA-7) that a well-validated instrument became available [7]. Furthermore, the Correlations Committee of the World Health Organization (WHO)-sponsored First International Consultation on BPH recommended that the AUA-7, plus one additional question to assess the global impact of BPH on quality of life (QoL), should be adopted as the global standard for BPH symptom scoring [8]. It was then renamed the International Prostate Symptom Score (I-PSS). The WHO Committee on Symptoms, Quality of Life and Sexuality has also accepted the I-PSS as an appropriate symptom-quantifying instrument among its clinical research criteria and recommended retaining it as the basic questionnaire to be used when making comparisons between different studies [9], thereby responding to a previously unmet need [3]. This WHO Committee has raised concerns about problems in the translation of the I-PSS from English into other languages and has therefore included in its recommendations advice about the use of an appropriate translation methodology. It has also stated that either the American-English or the French versions included in the proceedings of the 1993 meeting can be used as the source versions for translations [9].

The use of the I-PSS in multinational clinical trials makes it necessary to use translated versions of the questionnaire in various languages [10]. It must be possible to pool symptom scores, as well as other study end-points, from all participating centers and the measuring instrument (I-PSS) must be 'calibrated' in a uniform fashion at every study site [11, 12]. This means that all translations have to be produced by following a standard methodology which ensures that the same symptom severity is given the same score regardless of the language used [13]. The concept of harmonization of a set of translations can thus

be defined as the process by which all translated versions are made homogenous in terms of the questionnaire's instructions and format, content of translated items, and comparisons with the original.

The objectives of the present study were: (1) to develop a methodology for translation, cross-cultural adaptation and harmonization of this disease-specific symptom questionnaire (I-PSS); (2) to produce translations of the original American-English I-PSS questionnaire into 10 languages (Canadian French, Danish, Dutch, Flemish, French, German, Italian, Norwegian, Spanish, Swedish) and to adapt them cross-culturally, and (3) to harmonize them with each other and with the original American-English version.

Methods

For the purposes of this study, the original, validated English AUA-7 questionnaire as developed by Barry et al. [6] was used as the source version (TSO). It was decided to adopt this version because there are slight differences between it and the version published by the BPH Consultation in 1991 and 1993 [8, 9]. The quality of life item (No. 8) is also the one validated and included in the same publication [6].

For organizational purposes two levels were established: a central level (CL) located in Spain, and a local level (LL) in each country. The CL included the authors and a linguist, whilst the LLs included one SmithKline Beecham local coordinator (LoCo) and the translators used in each country. A methodology and standardized procedure were developed at CL and included in a protocol. Study record forms in English were provided to all participants. The process design was based on current guidelines regarding cross-cultural adaptations of translations [13, 14], and focussed especially on obtaining quantitative information about the quality of the translations and on the development of a simple, standardized, generalizable and cost-effective procedure.

The overall framework comprised the following activities: (1) production of forward translations; (2) quantification (rating) of their quality; (3) testing in target subjects; (4) back-translations into English, and (5) ratings for conceptual equivalence between the original and back-translations. A pilot run of this initial methodology was performed in one country (Spain) in order to test its feasibility and to eventually refine any design issue. All participating countries used the same study record forms and followed a common protocol with the following steps (fig. 1).

Step 1: Production of Forward Translations

Two or three independently produced translations (TV1, TV2 and TV3) were needed at LLs. TV1 was an already existing translation which could either be a published version or one which was commonly used in urological practice in a particular country. TV2 and TV3 were new, independent translations of the source version carried out by bilingual male translators whose mother tongue was the target language. Male translators were used because it was found that some of the expressions relating to urinary symptoms were not easily understood by female translators. It was specified that at least one of

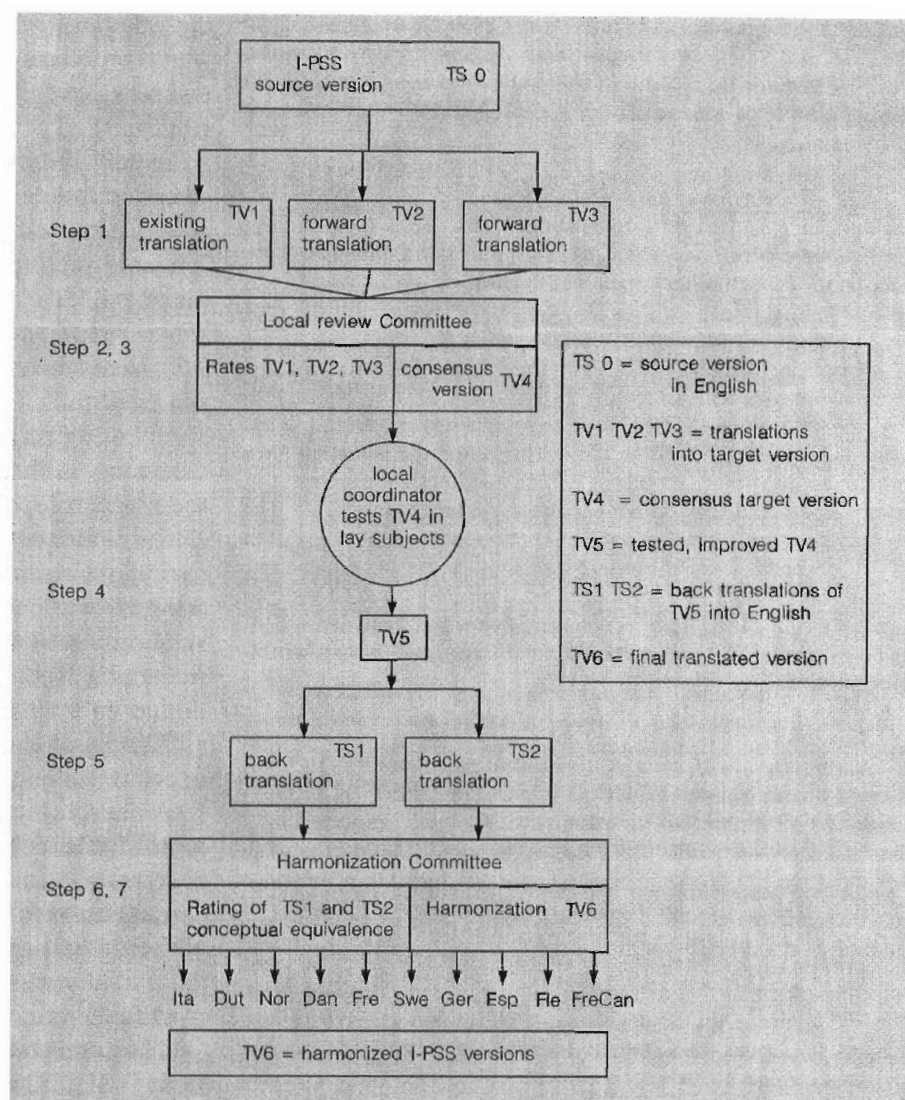


Fig. 1. I-PSS translation and harmonization process flow.

the translators at LL should be a professional translator. Standard instructions and record forms to be used by all translators at LL were provided by CL.

Step 2: Rating of Forward-Translations

Each LoCo identified 2 additional bilingual males (whose mother tongue was the target one) who were asked to independently rate each item with respect to *clarity*, *common language* and *cultural adequacy*. The raters were not asked to rate the response categories, but just to choose the one which they judged to be the best. The two or three translated versions of each item were grouped on a single page with a 10-cm horizontal visual analogue scale next to each item. All raters at all LLs were provided with the same standard instructions which were worded as follows: '... by comparing these three items rate the *clarity*, i.e. the item must express a single idea, should not contain ambiguous terms, should be easy to read and understand ...'; '... by comparing these three items rate if the item was written in *common language*, i.e. the item is expressed with the language used by the

general population ...', and '... by comparing these three items rate the *cultural adequacy*, i.e. if the item is appropriate and relevant for the culture that is being adapted ...'.

Step 3

A local review meeting was then held at each LL between the LoCo and the raters to develop a consensus translated version (TV4) consisting of item versions picked out from TV1, TV2 and TV3 or newly generated. The choice could be made based on the ratings obtained or on discussion and agreement among the participants in the local review meeting. The LoCo collected comments on the topics discussed and produced a report describing the problems found and the way they were solved.

Step 4

A comprehension test was carried out by administering the consensus version (TV4) to 3 lay non-patient male subjects, 1 from each of 3 different age groups: <45, 45–64, and >65. The LoCo read out

each item to the subject and then asked the subject what he thought the item meant. The wording of items could be improved at the LoCo's discretion if problems in comprehension were found. Again, the LoCo collected comments and produced a report. The translation after this step was designated TV5.

Step 5: Production of Back-Translations

Two independent back-translations (TS1, TS2) of TV5 into English were performed at LL by two professional bilingual male translators whose mother tongue was English.

Step 6: Conceptual Equivalence Assessment

This phase took place at CL. Two English linguists rated all items from all versions with respect to their conceptual equivalence with the original version (TSO). There were four options on the sheets provided for rating: (1) 'Both the wordings and the meanings have nothing to do with each other'; (2) 'The wordings are equivalent in some respects but they do not capture the same meaning'; (3) 'The wordings are not equivalent but they capture the same meaning', and (4) 'The wordings are equivalent and capture the same meaning'. Those items rated 1 or 2 by both raters were considered as nonconceptually equivalent. Items which were nonconceptually equivalent in both TS1 and TS2 were scheduled for discussion at the harmonization meeting.

Step 7

Harmonization Meeting. A harmonization committee (HC) was formed which consisted of an expert on QoL measurement, 2 physicians and a pharmacist with experience of clinical research, and 2 English linguists from the University of Barcelona. The HC first of all reviewed the full process, focussing on steps with the potential to produce quantitative information or qualitative modifications of adapted items, i.e. rating of forward-translations (step 1), and the comprehension test (step 4). The HC then discussed all those items which had been rated as nonconceptually equivalent (step 6). Finally, the HC analyzed the homogeneity of all translated versions with respect to format and content (number of items and response options), as well as some expressions which had been found to lead to confusion.

Queries. As a result of the harmonization meeting, queries were forwarded to the LoCos asking for alternative wordings for non-equivalent items.

Final Versions. Once all queries had been answered, the HC produced the final translated versions which were checked by LoCos for spelling errors.

Results

The pilot run of the initial methodology led to some changes. Firstly, it was realized that the raters of the forward-translations (step 2) did not need to be provided with the original English version; in fact the rating could be performed by non-bilinguals. As mentioned earlier, it was also found that expressions relating to urinary symptoms were not easily understood by female linguists and therefore the final protocol specified males to carry out

these tasks. Finally, the pilot run helped to anticipate where problems were most likely to occur later (item 8, response categories) and to refine the logistics of the process.

Overall duration of the study was 3 months. A previously translated version of the I-PSS (TV1) was found in all languages except Canadian French and Swedish. All forward-translators were professionals. Mean scores of items 1–7 (AUA-7) were high for *clarity* (range 6.4–9.2), *common language* (5.3–8.4) and *cultural adequacy* (6.4–9.8). Item 8 of the I-PSS also scored high with respect to the same three features (range of the means 5.1–9.1 for clarity and common language, and 6.1–9.7 for cultural adequacy; table 1). Excluding the response categories, there were 672 possible item ratings, of which 10% (19 for clarity, 34 for common language, and 14 for cultural adequacy) were rated below 5.0. Most (86%) of the items selected by the local committees to be included in TV4 versions were those with the highest scores, whereas the remaining ones were generally covered by the TV1 option, even when it was given a lower score.

The comprehension test in lay non-patient males produced comments from the subjects for 10% of items and response categories. Of all item comments, 11% related to item 7 (nocturia), 10% to item 8 (QoL), and 7% to item 2 (frequency). Surprisingly, most of the comments concerned response categories for items 1–6 (89% of the overall comments on response categories). However, only 3.6% of all items and response categories from TV4 were modified in some way when producing TV5.

The *conceptual equivalence* between the back-translations (TS1, TS2) and the source version (TSO) was scored independently by 2 bilingual English linguists, who found that 84% of all items, 88% of response categories for items 1–6, 100% of response categories for item 7, and 75% of response categories from item 8 were conceptually equivalent.

At the harmonization meeting, two kinds of problems were identified: some were specific to a particular country whilst others were general and affected all the countries. The LCs were accordingly asked for clarifications and/or requested to produce changes in wordings of particular items and response categories. Table 2 lists the general queries and some examples of specific ones. Eight item versions (10%) and 8 of the response options from item 8 (11%) produced specific queries. All LoCos satisfactorily solved all queries and the process ended with a spelling check by LoCos of all final, translated, cross-culturally adapted, harmonized versions (TV6; table 3).

Discussion

Translations of the I-PSS into 10 languages were linguistically harmonized for use in clinical trials. The method followed was that proposed by Brislin [15], though modified and tested specifically for this study. The methodology was found to be feasible, practical, cost-effective and was carried out in a short time (3 months). It was also standardized across all participating countries, the objective being to ensure that final versions maintained content validity. One of the elements involved in standardizing the methodology was the rating of forward-translations and allowing participants to use their discretion when deciding whether to base their choices on rating scores. This seems to have been effective since, in most cases (86%), the highest rated option was used in the pilot version. The efficacy of this stage was shown by the fact that 84% of items were found to be conceptually equivalent after back-translation. Conceptual equivalence testing did, however, reveal several major differences between back-translated versions and the source version, a fact which highlights the importance of including this stage in the translation process.

The translation of questionnaires has some limitations, which were recently addressed and summarized as follows [16]: (1) deficiencies in the original are carried over to translations; (2) there might be items which do not translate well, do not translate at all, or do not make sense within the new cultural context, and (3) items which were important for the original target population can be of trivial importance in the cultures of the new target populations.

With respect to the first-mentioned limitation (carry-over of deficiencies), it can be argued that the process used in this study allows for the detection of such deficiencies and their improvement in the new versions. For instance, in this study it was found that instructions for self-administration of the questionnaire were lacking; it was also found that changing the way the time-frame was expressed from 'the past month' to 'the last 30 days' made this aspect clearer, particularly when the questionnaire was to be used in the middle of the month; in item 7, 'on average' worked better than 'on a typical night', and a twofold improvement was proposed for item 8, 'urinary symptoms' instead of 'prostate symptoms', and using a scalable attribute, such as satisfaction, instead of different attributes on a continuum. Response options to items 1-6 were not understood by some of the individuals who completed the comprehension test. These response options could be improved by using numbers (i.e. between 30 and

Table 1. Mean ratings of the forward-translations for clarity, common language, and cultural adequacy in each language

	Clarity		Common language		Cultural adequacy	
	AUA-7	item 8	AUA-7	item 8	AUA-7	item 8
Flemish	7.4	5.1	6.3	7.0	9.8	8.7
Danish	7.2	8.7	8.1	9.5	8.3	9.7
Dutch	7.1	8.1	7.9	7.4	7.7	7.9
French	7.4	6.9	8.7	8.2	8.8	7.2
French/Canada	6.9	6.3	6.5	8.3	6.7	8.5
German	6.4	5.9	5.5	6.3	6.4	7.0
Italian	7.3	7.7	6.7	8.2	6.7	7.0
Norwegian	7.2	6.4	7.4	6.4	8.3	8.9
Spanish	9.2	9.1	8.4	9.1	6.6	6.1
Swedish	6.9	6.5	5.3	5.1	8.0	8.3

Three translations were available in each language except Canadian French and Swedish with only two.

Table 2. General queries and examples of specific queries made to local coordinators in each language

General queries

'The past month' should be substituted by 'the past 30 days' because some individuals understood it as the previous calendar month.

'Items 1-6 of the original version use the expression "how often" (measures frequency), while in item 7 "how many times" (measures number of times) is used. Both expressions are associated with the response category options. Since some of the back-translations do not consistently show the mentioned wording, please check ...'

The original English version mentions 'prostate symptoms' instead of 'urinary symptoms', and should be included in the version.

Specific queries (examples)

Item 5: The word 'size' is present in both back-translations but is missing from the original item. Please check.

Item 7: The original item tries to capture the idea of 'how many times this fact occurred in a typical night'. This idea is not present in both back-translations. Please, review the version because it seems that it is asking for the overall number of times on a month instead of the number of times in a typical night.

'The two back-translations of "terrible" were "terrified/very frightened", which are not conceptually equivalent ...'

50% of the time, etc.) instead of the actual wording (i.e. less than half of the time). We do not believe that these changes influence comparability or the measurement properties of the scale, but we do believe that they improve comprehension of the questionnaire.

Table 3. Final translated, harmonized versions in Canadian French, Danish, Dutch, Flemish, French, German, Italian, Norwegian, Spanish, and Swedish

Canadian French		Jamais	Moins de 1 fois sur 5	Moins de 1 fois sur 2	Environ 1 fois sur 2	Plus de 1 fois sur 2	Presque toujours
1	Au cours des derniers 30 jours, combien de fois avez-vous eu l'impression, après avoir uriné, que vous n'aviez pas vidé complètement votre vessie?	0	1	2	3	4	5
2	Au cours des derniers 30 jours, combien de fois avez-vous été obligé d'uriner à nouveau moins de deux heures après avoir uriné?	0	1	2	3	4	5
3	Au cours des derniers 30 jours, sur le nombre de fois que vous êtes allé uriner, combien de fois avez-vous été obligé d'arrêter et de recommencer pour vous soulager complètement?	0	1	2	3	4	5
4	Au cours des derniers 30 jours, lorsque vous avez en envie d'uriner, combien de fois avez-vous eu de la difficulté à vous retenir?	0	1	2	3	4	5
5	Au cours des derniers 30 jours, combien de fois est-ce que la force de votre jet d'urine a été faible?	0	1	2	3	4	5
6	Au cours des derniers 30 jours, combien de fois avez-vous été obligé de pousser ou de forcer pour arriver à uriner?	0	1	2	3	4	5
		jamais	1 fois	2 fois	3 fois	4 fois	5 fois ou plus
7	Au cours des derniers 30 jours, combien de fois par nuit avez-vous été obligé d'aller uriner après vous être couché?	0	1	2	3	4	5
		Très satisfait	Assez satisfait	Satisfait	Indifférent (ni satisfait ni insatisfait)	Assez insatisfait	Très insatisfait
8	Comment vous sentiriez-vous si vous deviez vivre avec les symptômes que vous venez de décrire pour le reste de votre vie?	0	1	2	3	4	5
Danish		Slet ikke	Mindre end 1 ud af 5 gange	Mindre end halvdelen af gangene	Omkring halvdelen af gangene	Mere end halvdelen af gangene	Næsten hver gang
1	Hvor mange gange gennem de sidste 30 dage, har De haft fornemmelsen af, at blæren ikke blev tømt ordentligt efter endt vandladning?	0	1	2	3	4	5
2	Hvor ofte har De indenfor de sidste 30 dage måttet lade vandet med mindre end 2 timers mellemrum?	0	1	2	3	4	5
3	Hvor mange gange har De gennem de sidste 30 dage oplevet, at urinstrålen afbrydes og starter igen flere gange under vandladningen?	0	1	2	3	4	5
4	Hvor mange gange har De gennem de sidste 30 dage oplevet, at det er svært at udsætte en vandladning?	0	1	2	3	4	5
5	Hvor mange gange har De gennem de sidste 30 dage bemærket, at kraften på urinstrålen var svag?	0	1	2	3	4	5
6	Hvor ofte har De gennem de sidste 30 dage måttet presse eller anstrenge Dem for at få vandladningen igang?	0	1	2	3	4	5
		ingen	1 gang	2 gange	3 gange	4 gange	5 gange eller mere
7	Gennem den sidste 30 dage, hvor mange har De gennemsnitligt måttet op om natten for at lade vandet?	0	1	2	3	4	5
		Fortræffeligt	Tilfreds	Overvejende tilfreds	Blandet (omtrent lig så tilfreds som utilfreds)	Overvejende utilfreds	Elendigt
8	Hvordan ville De have det med at skulle leve resten af livet med de prostata symptomer, De har i øjeblikket?	0	1	2	3	4	5
							Rædselsfuldt

Table 3 (continued)

Dutch		Niet één keer	Minder dan eens in de vijf keer	Minder dan de helft van het aantal keren	Ongeveer de helft van het aantal keren	Meer dan de helft van het aantal keren	Vrijwel iedere keer	
1	Hoe vaak heeft u de afgelopen 30 dagen het gevoel gehad dat uw blaas niet helemaal leeg was nadat u klaar was met plassen?	0	1	2	3	4	5	
2	Hoe vaak moest u de afgelopen 30 dagen na het plassen binnen twee uur opnieuw plassen?	0	1	2	3	4	5	
3	Hoe vaak kwam het de afgelopen 30 dagen voor dat u tijdens het plassen verschillende keren ophield en opnieuw begon, dus 'met schokjes' plaste?	0	1	2	3	4	5	
4	Hoe vaak kwam het de afgelopen 30 dagen voor dat u moeilijk uw plas kon ophouden?	0	1	2	3	4	5	
5	Hoe vaak produceerde u de afgelopen 30 dagen tijdens het plassen slechts een mager straaltje?	0	1	2	3	4	5	
6	Hoe vaak moest u de afgelopen 30 dagen persen of drukken om te kunnen beginnen met plassen?	0	1	2	3	4	5	
		niet één keer	1 keer	2 keer	3 keer	4 keer	ten minste 5 keer	
7	Hoeveel keer moest u de afgelopen 30 dagen gemiddeld per nacht uw bed uit om te plassen?	0	1	2	3	4	5	
		Uitstekend	Goed	Over het algemeen tevreden	Gemengde gevoelens (ongeveer even tevreden als ontevreden)	Over het algemeen ontevreden	Ongelukkig	Verschrikkelijk
8	Hoe zou u zich voelen als u voor de rest van uw leven last zou houden van uw huidige prostaatklasten?	0	1	2	3	4	5	6
Flemish		Nooit	Minder dan 1 keer op 5	Minder dan de helft van de keren	Ongeveer de helft van de keren	Meer dan de helft van de keren	Bijna elke keer	
1	Hoe dikwijls hebt U de voorbije 30 dagen het gevoel gehad dat Uw blaas niet volledig leeg was, nadat U gedaan had met wateren?	0	1	2	3	4	5	
2	Hoe dikwijls hebt U de voorbije 30 dagen opnieuw moeten gaan wateren minder dan twee uur, nadat U gedaan had met wateren?	0	1	2	3	4	5	
3	Hoe dikwijls hebt U de voorbije 30 dagen bij het wateren opgemerkt dat de straal meerdere keren stopte en dan weer opnieuw begon?	0	1	2	3	4	5	
4	Hoe dikwijls hebt U het de voorbije 30 dagen moeilijk gevonden om het wateren uit te stellen?	0	1	2	3	4	5	
5	Hoe dikwijls hebt U de voorbije 30 dagen met een zwakke straal gewaterd?	0	1	2	3	4	5	
6	Hoe dikwijls hebt U de voorbije 30 dagen moeten duwen of persen om te beginnen wateren?	0	1	2	3	4	5	
		nooit	1 keer	2 keer	3 keer	4 keer	5 keer of meer	
7	Hoeveel keren per nacht moest U gewoonlijk opstaan om te gaan wateren, de voorbije 30 dagen?	0	1	2	3	4	5	
		Schitterend	Tevreden	Eerder tevreden	Met gemengde gevoelens (even voldaan als onvoldaan)	Eerder ontevreden	Ongelukkig	Vreselijk
8	Hoe zou U zich voelen indien U de rest van Uw leven prostaatklasten zou hebben zoals nu?	0	1	2	3	4	5	6

Table 3 (continued)

French		Jamais	Moins de 1 fois sur 5	Moins de 1 fois sur 2	Environ 1 fois sur 2	Plus de 1 fois sur 2	Presque toujours	
1	Au cours des derniers 30 jours, combien de fois avez-vous eu l'impression que votre vessie n'était pas complètement vidée après avoir uriné?	0	1	2	3	4	5	
2	Au cours des derniers 30 jours, combien de fois avez-vous eu besoin d'uriner à nouveau moins de 2 heures après avoir fini d'uriner?	0	1	2	3	4	5	
3	Au cours des derniers 30 jours, combien de fois avez-vous eu des interruptions répétées du jet d'urine, c'est-à-dire démarrage du jet puis arrêt, puis redémarrage?	0	1	2	3	4	5	
4	Au cours des derniers 30 jours, combien de fois avez-vous eu des difficultés à vous retenir d'uriner?	0	1	2	3	4	5	
5	Au cours des derniers 30 jours, combien de fois avez-vous eu une diminution de la taille ou de la force du jet d'urine?	0	1	2	3	4	5	
6	Au cours des derniers 30 jours, combien de fois avez-vous du pousser ou faire un effort pour commencer à uriner?	0	1	2	3	4	5	
		jamais	1 fois	2 fois	3 fois	4 fois	5 fois ou plus	
7	Au cours des derniers 30 jours, combien de fois par nuit, en moyenne, vous êtes vous levé pour uriner, entre le moment de votre coucher et celui de votre lever définitif?	0	1	2	3	4	5	
		Très satisfait	Satisfait	Plutôt satisfait	Mitigé (ni satisfait, ni ennuyé)	Plutôt ennuyé	Ennuyé	Très ennuyé
8	Si vous deviez passer le restant de votre vie avec votre manière d'uriner actuelle, diriez-vous que vous en seriez?	0	1	2	3	4	5	6
German		Überhaupt nicht	Weniger als jedes fünfte Mal	Weniger als jedes zweite Mal	Etwa jedes zweite Mal	Mehr als jedes zweite Mal	Fast immer	
1	Wie oft hatten Sie das Gefühl, dass Ihre Blase nach dem Wasserlassen nicht ganz entleert war (während der letzten 30 Tage)?	0	1	2	3	4	5	
2	Wie oft mussten Sie innerhalb von 2 Stunden erneut Wasser lassen (während der letzten 30 Tage)?	0	1	2	3	4	5	
3	Wie oft mussten Sie beim Wasserlassen mehrmals aufhören und wieder neu beginnen (während der letzten 30 Tage)?	0	1	2	3	4	5	
4	Wie oft hatten Sie Schwierigkeiten, das Wasserlassen hinauszuzögern (während der letzten 30 Tage)?	0	1	2	3	4	5	
5	Wie oft hatten Sie einen schwachen Strahl beim Wasserlassen (in den letzten 30 Tagen)?	0	1	2	3	4	5	
6	Wie oft mussten Sie pressen oder sich anstrengen, um mit dem Wasserlassen zu beginnen (in den letzten 30 Tagen)?	0	1	2	3	4	5	
		nie	1mal	2mal	3mal	4mal	5mal oder öfter	
7	Wie oft sind Sie im Durchschnitt nachts aufgestanden, um Wasser zu lassen (vom Zubettgehen bis zum morgendlichen Aufstehen; Zeitraum: letzte 30 Tage)?	0	1	2	3	4	5	
		Sehr zufrieden	Zufrieden	Überwiegend zufrieden	Gemischte Gefühle (etwa zu gleichen Teilen zufrieden und unzufrieden)	Überwiegend unzufrieden	Unzufrieden	Äußerst unzufrieden
8	Wie würden Sie sich fühlen, wenn Sie den Rest Ihres Lebens mit den jetzigen Prostatasymptomen verbringen müssten?	0	1	2	3	4	5	6

Table 3 (continued)

Italian		Mai	Meno di una volta su cinque (a volte)	Meno della metà delle volte	Circa la metà delle volte	Più della metà delle volte	Quasi sempre	
1	Negli ultimi 30 giorni, quante volte ha avuto la sensazione di non aver svuotato completamente la vescica dopo aver urinato?	0	1	2	3	4	5	
2	Negli ultimi 30 giorni, quante volte ha avuto la necessità di urinare nuovamente a meno di due ore di distanza dalla volta precedente?	0	1	2	3	4	5	
3	Negli ultimi 30 giorni, quante volte si è accorto di urinare in modo intermittente, iniziando e smettendo più volte mentre urinava?	0	1	2	3	4	5	
4	Negli ultimi 30 giorni, quante volte ha trovato difficoltà a trattenere l'urina?	0	1	2	3	4	5	
5	Negli ultimi 30 giorni, quante volte ha avuto un debole getto d'urina?	0	1	2	3	4	5	
6	Negli ultimi 30 giorni, quante volte ha dovuto spingere o sforzarsi per iniziare ad urinare?	0	1	2	3	4	5	
		mai	una volta	due volte	tre volte	quattro volte	almeno cinque volte	
7	Negli ultimi 30 giorni, in media quante volte per notte è stato costretto ad alzarsi per urinare?	0	1	2	3	4	5	
		Felice	Contenta	Abbastanza soddisfatto	Indifferente (né soddisfatto né contrariato)	Dispiaciuto	Infelice	Disperato
8	Come si sentirebbe se dovesse passare il resto della sua vita con i disturbi che le dà adesso la sua prostata?	0	1	2	3	4	5	6
Norwegian		Aldri	Mindre enn 1 av 5 ganger	Mindre enn halvparten av gangene	Omtrent halvparten av gangene	Mer enn halvparten av gangene	Nesten alltid	
1	I løpet av den siste 30 dager, hvor ofte har du hatt følelsen av at urinblæren ikke er helt tømt etter at du har latt vannet?	0	1	2	3	4	5	
2	I løpet av den siste 30 dager, hvor ofte har du måttet late vannet igjen mindre enn to timer etter forrige gang?	0	1	2	3	4	5	
3	I løpet av den siste 30 dager, hvor ofte har du merket at du stoppet og startet igjen flere ganger under vannlatingen?	0	1	2	3	4	5	
4	I løpet av den siste 30 dager, hvor ofte har du hatt vanskeligheter med å utsette vannlatingen?	0	1	2	3	4	5	
5	I løpet av den siste 30 dager, hvor ofte har du hatt svak urinstråle?	0	1	2	3	4	5	
6	I løpet av den siste 30 dager, hvor ofte har du måttet trykke på eller anstrenge deg for å starte vannlatingen?	0	1	2	3	4	5	
		ingen	en gang	to ganger	tre ganger	fire ganger	fem eller flere ganger	
7	I løpet av den siste 30 dager, hvor mange ganger har du vanligvis måttet stå opp for å late vannet fra du legger deg om kvelden til du står opp om morgenen?	0	1	2	3	4	5	
		Godt fornøyd	Fornøyd	Stort sett fornøyd	Blandet (omtrent like fornøyd som misfornøyd)	Stort sett misfornøyd	Trist	Forferdelig
8	Hvis du måtte leve resten av livet med dagens prostata-problemer, hvordan ville du da føle deg?	0	1	2	3	4	5	6

Table 3 (continued)

Spanish	Ninguna	Menos de 1 vez cada 5	Menos de la mitad de las veces	Aproximadamente la mitad de las veces	Más de la mitad de las veces	Casi siempre	
1 Durante más o menos los últimos 30 días, ¿cuántas veces ha tenido la sensación de no vaciar completamente la vejiga al terminar de orinar?	0	1	2	3	4	5	
2 Durante más o menos los últimos 30 días, ¿cuántas veces ha tenido que volver a orinar en las dos horas siguientes después de haber orinado?	0	1	2	3	4	5	
3 Durante más o menos los últimos 30 días, ¿cuántas veces ha notado que, al orinar, paraba y comenzaba de nuevo varias veces?	0	1	2	3	4	5	
4 Durante más o menos los últimos 30 días, ¿cuántas veces ha tenido dificultad para aguantarse las ganas de orinar?	0	1	2	3	4	5	
5 Durante más o menos los últimos 30 días, ¿cuántas veces ha observado que el chorro de orina es poco fuerte?	0	1	2	3	4	5	
6 Durante más o menos los últimos 30 días, ¿cuántas veces ha tenido que apretar o hacer fuerza para comenzar a orinar?	0	1	2	3	4	5	
	ninguna	1 vez	2 veces	3 veces	4 veces	5 veces o más	
7 Durante más o menos los últimos 30 días, ¿cuántas veces suele tener que levantarse para orinar desde que se va a la cama por lo noche hasta que se levanta por lo mañana?	0	1	2	3	4	5	
	Encantado	Muy satisfecho	Más bien satisfecho	Tan satisfecho como insatisfecho	Más bien insatisfecho	Muy insatisfecho	Fatal
8 ¿Cómo se sentiría si tuviera que pasar el resto de la vida con los síntomas prostáticos tal y como los tiene ahora?	0	1	2	3	4	5	6
Swedish	Aldrig	Mindre än en gång av fem	Mer sällan än varannan gång	Ungefär varannan gång	Oftare än varannan gång	Nästan alltid	
1 Hur ofta under de senaste 30 dagarna har du haft en känsla av att inte ha tömt blåsan helt när du kastat vatten?	0	1	2	3	4	5	
2 Hur ofta under de senaste 30 dagarna har du varit tvungen att kasta vatten igen mindre än två timmar efter det att du senast kastat vatten?	0	1	2	3	4	5	
3 Hur ofta under de senaste 30 dagarna har du märkt att du slutar och börjar om flera gånger när du kastar vatten?	0	1	2	3	4	5	
4 Hur ofta under de senaste 30 dagarna har du märkt att det varit svårt att vänta en stund med att kasta vatten?	0	1	2	3	4	5	
5 Hur ofta under de senaste 30 dagarna har det endast kommit en svag stråle, när du kastat vatten?	0	1	2	3	4	5	
6 Hur ofta under de senaste 30 dagarna har du varit tvungen att pressa på eller krysta för att börja kasta vatten?	0	1	2	3	4	5	
	ingen gång	1 gång	2 gånger	3 gånger	4 gånger	5 eller fler gånger	
7 Hur många gånger under de senaste 30 dagarna måste du normalt gå upp på natten för att kasta vatten?	0	1	2	3	4	5	
	Mycket nöjd	Nöjd	Ganska nöjd	Blandade känslor (ungefär lika mycket nöjd som missnöjd)	Ganska missnöjd	Olycklig	Fruktansvärt
8 Hur skulle du känna dig om du visste att du måste leva resten av ditt liv med dina nuvarande prostatabesvär?	0	1	2	3	4	5	6

We did not anticipate that the second limitation (items which are difficult or impossible to translate) would be a problem in this case (except for item 8), since items 1–7 refer to symptoms which could be expected to be similar across the different countries participating in the study. Item 8, however, contained response options related to concepts such as unhappiness and delight, which often have different meanings in different cultures [17]. The problems which arose were solved by dropping conflictive expressions out of the scale and replacing them with alternative translations which maintained conceptual equivalence with the original when graded on a Likert scale. For instance, in some versions, the response options for item 8 were based around the root concept of satisfaction, with possible responses ranging from 'extremely satisfied' to 'extremely dissatisfied'. It can therefore be expected that clinimetric questionnaires will present fewer problems than QoL questionnaires, since the latter contain words, concepts and situations which are very often culture-sensitive.

Good cultural adequacy was found in all translations from this study, as reflected by the fact that no item received a mean rating of less than 5.0. As expected then, the third previously mentioned limitation (different cultural adequacy) did not arise in this case.

Although one of the apparent strengths of the present methodology was quantification, i.e. rating, of some properties of the translated items, some aspects of the way it was done might be improved. Firstly, the three properties selected to be rated (clarity, common language and cultural adequacy) would appear to cover the main features to be explored, but in fact other aspects to be quantified could also be proposed, such as the cultural relevance of the item (adequacy and relevance would be different things to test). The decision to make 5.0 the lower limit of adequacy is rather arbitrary, the continuum of clarity, common language and cultural adequacy could be defined more precisely, and the raters' reliability might warrant investigation, something which could be done by simply replicating the exercise after a 2-week interval.

It is recommended that comprehension tests be performed on large samples of individuals in optimal models (more than 50) [18]. However, there are few reports which analyze the effects on the final version of increasing the sample size of the respondents [14]. Comprehension testing means a greater additional effort in multicountry translations; in the present study, translations were tested on a sample of 3 non-patients for each language. During development of the original questionnaire [3], the same test performed on a much larger population (150 patients)

produced results which were similar to those obtained here in the comprehension test.

The I-PSS is a symptom-based (clinimetric) questionnaire and no great problems arose in the process of translation and adaptation. It has also been shown that all the versions are conceptually equivalent. An important question in such a case is whether or not it is necessary to retest the psychometric properties (validity, reliability, responsiveness) of the translated versions. Intuitively, we might expect that, in the case of symptom-based questionnaires, if there were no great problems in translation and adaptation, the psychometric properties of the translated versions would be highly similar or identical to those of the original questionnaire. Experience also suggests that this is the case [19, 20]. A full discussion of these issues is beyond the scope of this article; however, they do merit further investigation, especially as retesting of psychometric properties entails further time and expense, something which might be avoided if it could be shown that conceptually equivalent versions of clinimetric questionnaires always have the same psychometric properties as the original. At present, there is not sufficient evidence to say that this is definitely the case; in some of the countries validation studies are currently underway. For the purposes of this study, however, we have assumed that translations did not alter the psychometric properties of the questionnaire.

Thus we can conclude that the present set of 10 translations plus the original I-PSS questionnaire can be used in multi-country studies in which BPH symptoms are to be assessed using this WHO-recommended questionnaire and that the data collected from different countries can be legitimately aggregated.

Acknowledgements

The authors wish to thank F. Jiménez-Cruz, MD, E. Rappaport, MD, and M. Tabona, MD, for their critical review of this manuscript; Mr. R. Rycroft (Department of Linguistics, University of Barcelona) for his advice and contribution, and Ms. M. Moharra and Ms. L. Martín for secretarial assistance. We are also indebted to the following SmithKline Beecham Pharmaceuticals personnel (local coordinators): Paola Barbieri (Italy), Jan De Wit (Belgium), Elena García (Spain), Kjell Hernes (Norway), Mai Thu Hoang (Canada), Rémi Le Coent (France), Erik Lyngholm (Denmark), Caroline Melman (Holland), Harriet Ohlsson (Sweden), and Detlev Prüehs (Germany).

References

- 1 Department of Veterans Affairs Cooperative Study of Transurethral Resection for Benign Prostatic Hyperplasia: A comparison of quality of life with patient reported symptoms and objective findings in men with benign prostatic hyperplasia. *J Urol* 1993;150:1696-1700.
- 2 O'Leary MP, Barry MJ, Floyd J, Fowler JR: Hard measures of subjective outcomes: Validating symptom indexes in urology. *J Urol* 1992;148:1546-1548.
- 3 Barry MJ, Cockett ATK, Holtgrewe L, McConnell JD, Silhelnik SA, Winfield HN: Relationship of symptoms of prostatism to commonly used physiological and anatomical measures of the severity of benign prostatic hyperplasia. *J Urol* 1993;150:351-358.
- 4 Yalla SV, Sullivan MP, Lecamwasam HS, Dubeau CE, Vickers MA, Cravalho EG: Correlation of American Urological Association symptom index with obstructive and nonobstructive prostatism. *J Urol* 1995;153:674-680.
- 5 Barry MJ, Fowler FJ, O'Leary MP, Bruskewitz RC, Holtgrewe L, Mebust WK: Correlation of the American Urological Association symptom index with self-administered versions of the Madsen-Iversen, Boyarsky and Maine medical assessment program symptom indexes. *J Urol* 1992;148:1558-1563.
- 6 McConnell JD, Barry MJ, Bruskewitz RC, et al: Benign Prostatic Hyperplasia: Diagnosis and Treatment. Clinical Practice Guideline Number 8. AHCPR Publication No. 94-0582. Rockville, Agency for Healthy Care Policy and Research, Public Health Service, US Department of Health and Human Services, 1994.
- 7 Barry MJ, Fowler FJ, O'Leary MP, Bruskewitz RC, Holtgrewe L, Mebust WK, Cockett ATK: The American Urological Association symptom index for benign prostatic hyperplasia. *J Urol* 1992;148:1549-1557.
- 8 Cockett ATK, Aso Y, Chatelain C, Denis L, Griffiths K, Khoury S, Murphy G (eds): Proceedings of the International Consultation on Benign Prostatic Hyperplasia, Paris, June 26-27, 1991.
- 9 Cockett ATK, Khoury S, Aso Y, Chatelain C, Denis L, Griffiths K, Murphy G (eds): Proceedings of the 2nd International Consultation on Benign Prostatic Hyperplasia, Paris, June 27-30, 1993.
- 10 Cella DF, Wiklund I, Shumaker SA, Aaronson NK: Integrating health-related quality of life into cross-national clinical trials. *Qual Life Res* 1994;2:433-440.
- 11 Patrick DL, Wild DJ, Johnson ES, Wagner TH, Martin MA: Cross-cultural validation of quality of life measures; in Orley A, Kuyken W (eds): *Quality of Life Assessment: International Perspectives*. Berlin, Springer, 1995.
- 12 Badia X, Alonso J: Re-scaling the Spanish version of the Sickness Impact Profile: An opportunity to assess cross-cultural equivalence. *J Clin Epidemiol* 1995;48:949-957.
- 13 Guillemin F, Bombardier C, Beaton D: Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417-1432.
- 14 Bullinger M, Anderson R, Cella DF, Aaronson N: Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Qual Life Res* 1994;2:451-459.
- 15 Brislin RW: Back-translation for cross-cultural research. *J Cross-Cultural Psychol* 1970;1:185-216.
- 16 Guyatt GH: The philosophy of health-related quality of life translation. *Qual Life Res* 1994;2:461-465.
- 17 Hunt SM: Cross-cultural comparability of quality of life measures. *Drug Inform J* 1993;27:395-400.
- 18 Guyatt GH, Bombardier C, Tugwell PX: Measuring disease-specific quality of life in clinical trials. *Can Med Assoc J* 1985;134:889-895.
- 19 Badia X, Alonso J: Validity and reproducibility of the Spanish version of the Sickness Impact Profile. *J Clin Epidemiol* 1996;49:359-365.
- 20 Alonso J, Prieto L, Antó JM: The Spanish version of the Nottingham Health Profile: A review of adaptation and instrument characteristics. *Qual Life Res* 1994;3:385-393.