

Linguistic Validation of the Geriatric Depression Scale (Yesavage et al 1983) for the Welsh Language

Background

This report offers a brief account of the linguistic validation of the Geriatric Depression Scale' (Yesavage et al 1983) for the Welsh language. The work was undertaken between August and November 2010 in preparation for its use as a secondary outcome measure for the OTCH study, A Cluster Randomised Controlled Trial of an Occupational Therapy Intervention for Residents with Stroke living in UK Care-homes. This four year study, led by Birmingham University, sets out to evaluate the effects of a targeted course of occupational therapy for people with stroke living in a care home, with particular emphasis on independence in self care activities of daily living and mobility. The setting for the study includes care homes in areas of bilingual Gwynedd where a substantial proportion of the population are Welsh speaking (NAW 2003); and whose language needs / preferences in research studies are facilitated through legislative processes (Welsh Language Act 1993) and research governance (WAG 2009).

Translation and Validation of Patient Reported Outcome Measures (PROMs)

Owing to the universality of the English language in science and medicine, the majority of PROMs have been compiled in the English language to be used with the English speaking population (Acquadro et al 2008). However, this may not be the preferred language of a significant proportion of service users or research respondents. Moreover, there is evidence to suggest that language barriers may compromise the validity and reliability of such measures (Fitzpatrick et al 1998). Thus, in order to capture and compare the health outcomes of diverse populations accurately, PROMs must be translated from the source language into the target languages most commonly used within study locations. Establishing equivalence between source and target versions poses significant challenges for researchers since direct translation alone may fail to capture the reality of the experience for respondents who have very different cultural and language orientations. Streiner & Norman (2003) propose that at least five key dimensions should be considered in establishing equivalence across research instruments, as follows:

- *Conceptual equivalence* where theoretical constructs are comparable in each culture
- *Semantic equivalence* where words have the same meaning in both language versions
- *Item equivalence* where items are relevant and acceptable in the target language
- *Operational equivalence* where the format of the scale and mode of administration are relevant and acceptable in the target language
- *Measurement equivalence* where the psychometric properties of the scale are comparable in both language versions

Adopting a systematic and comprehensive approach to translation can help ensure that the adapted measure is conceptually equivalent to the source version; culturally

relevant and acceptable to the target population; and psychometrically comparable (Wang et al 2006).

Geriatric Depression Scale

The Geriatric Depression Scale (GDS) was first developed in the US in 1982 by Yesavage and others (Yesavage et al 1983). It was specifically designed for rating depression amongst older people and is now commonly used as part of the comprehensive assessment of the older person. The scale is made up of 30 items with a 'yes' / 'no' response. This simplicity enables the scale to be used with ill or moderately cognitively impaired individuals. One point is assigned to each answer and the cumulative score is rated on a scoring grid. The grid sets a range of 0-9 as "normal", 10-19 as "mildly depressed", and 20-30 as "severely depressed". The measure is sensitive; reliable; has concurrent and predictive validity; and a high internal consistency (Yesavage et al 1983). It has been translated into a number of languages, and has proved to be psychometrically robust and appropriate for use in countries across the world e.g. Japan (Niino et al 1991); Portugal (Zilenovski, 1991); Italy (Ferrario et al 1990); and Spain (Izal & Montorio, 1993). There is, however, currently no Welsh language version of the GDS.

This report offers an evaluation of the Welsh translation and adaptation process; the challenges encountered: and the efforts adopted to overcome discrepancies.

Methodology

Following permission granted by Professor Jerome Yesavage, Stanford University School of Medicine, California on 6th July 2010, the team set out to develop a Welsh adaptation of the GDS in line with the ISPOR Guidelines (Wild et al 2005) (see Figure 1). This is the translation approach currently adopted by LLAIS (www.llais.org) in our advisory role within NISCHR CRC.

In line with the guidelines, the following key personnel were employed to the project team and a timetable was established:

- Project manager (GR) - bilingual (Welsh/English)
- Project co-ordinator (HO) - bilingual (Welsh/English)
- Terminologist (DP) - bilingual (Welsh/English)
- 3 Independent translators - bilingual (Welsh/English)
- Lay respondents (n = 8) - bilingual (Welsh/English)

Work commenced in August 2010 whereby two independent translators set to work on the Welsh translation of the GDS. The translators were asked to focus on conceptual rather than literal equivalence; neutral wording and phraseology, in terms of dialect; and compatibility with a reading level of age 14 years. Reconciliation of the two forward translations into a single merged document was carried out and this consensus version was then translated back into English by the third independent translator as a quality control step (Brislin 1970). Comparison of the back translation with the original measure highlighted any discrepancies in the reconciled translation which were then revised through discussion amongst the main project team. Eight

Welsh-speaking lay respondents (four males and four females) aged between 50 and 80+ years were then recruited through Involving People and invited to complete the new Welsh GDS and participate in a semi-structured focus group interview to check their comprehension and interpretation of the translation as well as its cultural relevance. Comparing these interpretations with the original version of the measure revealed any remaining discrepancies and enabled us to agree on the final Welsh translation.

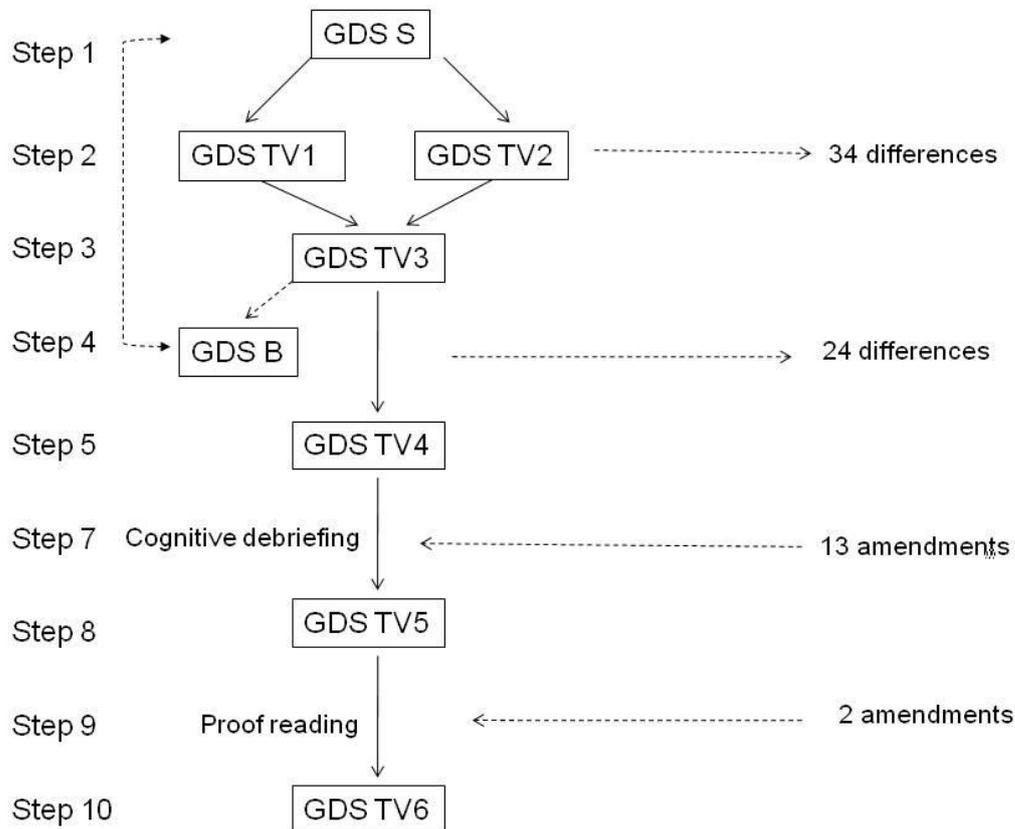
Figure 1 Summary of good practice guidelines for translation (adapted from Wild et al 2005)

<i>No.</i>	<i>Step</i>	<i>Brief description</i>
(1)	Preparation	Initial work carried out before the translation work begins, e.g. gaining permission, deploying staff and establishing timetable
(2)	Forward translation	Translation of the original language, also called source, version of the instrument into another language, often called the target language.
(3)	Reconciliation	Comparing and merging more than one forward translation into a single forward translation.
(4)	Back translation	Translation of the new language version back into the original language.
(5)	Back translation review	Comparison of the back-translated version of the instrument with the original to highlight and investigate discrepancies between the original and the reconciled translation, which is then revised in the process of resolving the issues.
(6)	Harmonization	Comparison of back translations of multiple language versions with each other and the original instrument to highlight discrepancies between the original and its derivative translations, as well as to achieve a consistent approach to translation problems.
(7)	Cognitive debriefing	Testing the instrument on a small group of relevant patients or lay people in order to test alternative wording and to check understanding, interpretation and cultural relevance of the translation
(8)	Review of cognitive debriefing results and finalization	Comparison of the patients' or lay persons' interpretation of the translation with the original version to highlight and amend discrepancies.
(9)	Proofreading	Final review of the translation to highlight and correct any typographic, grammatical or other errors
(10)	Final report	Report written at the end of the process documenting the development of each translation.

Results

A process diagram of the linguistic validation process is outlined in Figure 2.

Figure 2 Linguistic validation of the Welsh language version of the Geriatric Depression Index (GDS)



Key

S = source language; T = target language; B = back translation; V = version

In line with the ISPOR guidelines (Wild et al 2005), all 30 items of the GDS, along with the instructions for completing the measure, were translated into Welsh initially by two independent translators. A detailed comparison of these two versions at Step Two revealed a total of 34 differences (excluding duplications) and these offered a basis for discussions at the reconciliation meeting at Step Three. Most of these differences were attributed to the choice of different synonyms of the same concept, that is, a simple substitution of a word or phrase with another with the same meaning. For example, 'ofni' (to fear/be fearful) and 'pryderu' (to worry/be anxious) were offered as translations of 'afraid' but 'ofni' was favoured as it was felt to be a more accurate term in this context.

A few words and phrases proved particularly challenging in reaching agreement at the reconciliation stage and this was mainly because of their lack of direct Welsh equivalent or lack of use in everyday, informal Welsh health discourse. For example, whilst in the English language, 'blue' may refer to the feeling of sadness; the colour

does not carry the same connotation in Welsh. Thus, 'yn y felan' (lit: melancholy) was chosen as the most appropriate translation of 'feeling blue'.

As soon as consensus was reached at Step Three, the third version of the GDS was translated back into English and then reviewed against the source version. This revealed a total of 24 differences (excluding duplications). Close examination showed that the majority of lexical items in the English back translation, whilst not identical to the original English version, were in fact close conceptual equivalents or synonyms, such as 'basically' and 'generally'; 'satisfied' and 'content'; and 'spirits' and 'mood'.

Whilst Step Five was considered a valuable process to identify discrepancies in translation, Step Six was omitted because the measure was only being translated into one language. Reports of the ensuing cognitive debriefing confirmed the respondents' overall comprehension of the Welsh measure and its clarity for completion within five to ten minutes.

Step Seven was conducted in order to test alternative wording and to check the comprehension, interpretation, and cultural relevance of the translation. The new instrument was tested on a group of 8 Welsh speaking volunteers aged between 50 and 80+ years of age. In light of this process, 13 minor amendments were made which helped to enrich and clarify the items. In some cases a word/term was changed for consistency or clarification whilst, in other cases, alternative expressions were sought that were deemed to be more natural. For example, 'geriatrig' (lit: geriatric) was changed to 'pobl hŷn' (lit: older people); 'yn y felan' (lit: melancholy) to 'isel' (lit: low); and 'sefyllfaoedd cymdeithasol' (lit: social situations) to 'mynd i ganol pobl' (lit: going amongst people).

Conclusions

The process of commissioning two independent translations, a back translation, and then creating a synthesis translation for testing, was a useful exercise in achieving a clear and accurate rendering of the original GDS in Welsh. Minor refinements were made following respondents' feedback, and the measure may be further refined in due course in light of psychometric testing with end-users.

References

Acquadro C, Conway K, Hareendran A & Aaronson N (2008) Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health* 11, 3, 509-521.

Brislin R (1970) Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology* 1, 185-216.

Ferrario E, Cappa G, Bertone O, Poli L & Fabris F (1990) Geriatric Depression Scale and assessment of cognitive disturbances in the elderly: A preliminary report on an Italian sample. *Clinical Gerontologist*, 10, 67-74.

Fitzpatrick R., Davey C., Buxton M. & Jones D. (1998) Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 2, 14.

Izal M & Montorio I (1993) Adaptation of the GDS in Spain: A preliminary study. *Clinical Gerontologist*, 13, 2, 83-91.

National Assembly for Wales, 2003. *National Statistics*. National Assembly for Wales, Cardiff.

Niino N, Imaizumi T, & Kawakami N (1991) A Japanese translation of the Geriatric Depression Scale. *Clinical Gerontologist*, 10, 85-87.

Streiner D & Norman G (2003) *Health Measurement Scales*. (3rd edition) Oxford University Press, Oxford.

Wang W., Lee H. & Fetzer S. (2006) Challenges and strategies of instrument translation. *Western Journal of Nursing Research* 28, 3, 310-321.

Welsh Assembly Government (2009) *National Research Governance Framework for Health and Social Care*. (2nd edition) Welsh Assembly Government, Cardiff.

Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A & Erikson P (2005) Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health* 8 (2), 94-104.

Yesavage J, Brink T, Rose T, Lum O, Huang V, Adey M & Leirer V (1982) Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research* 17, 1, 37-49.

Zilenovski A (1991) Geriatric Depression Scale: A Portuguese translation. *Clinical Gerontologist*, 10, 84-86.

Gwerfyl Roberts
Heledd Owen
Delyth Prys

December 2010