

COMMENTARY

GOAL ATTAINMENT SCALING AND ITS RELATIONSHIP WITH STANDARDIZED OUTCOME MEASURES: A COMMENTARY

Two articles in this issue address goal attainment scaling (GAS):

- Ertzgaard et al. (1) provide a descriptive review of the available literature for GAS as an outcome measure in patients undergoing rehabilitation, particularly following acquired brain injury. They discuss the now extensive literature to support the use of GAS as a sensitive and reliable measure of clinically meaningful change. Their overall conclusions are favourable, although there are significant methodological challenges in its application, for which they make some practical suggestions.
- Bovendeert et al. (2) report a study of agreement and reliability of GAS in the context of a randomized controlled trial (RCT) of motor imagery in neurorehabilitation for a small group of 29 patients with various neurological disorders. They found poor agreement in goal scoring between 2 different scoring procedures, undertaken by: (i) the patient's therapists and (ii) an independent assessor unfamiliar with the patient; and therefore raise a note of caution before GAS is used as an outcome measure in blinded RCTs.

The two articles highlight a number of important issues in relation to GAS. Firstly, it must be remembered that GAS is not a measure of outcome *per se*, but a measure of the achievement of expectation (3). It does not replace standardized measures, but may be used alongside them to assist interpretation. This is particularly important in the context of rehabilitation, where many patients will have significant ongoing disability. For example, the treating team may anticipate that an individual who is unable to walk at the start of the programme (level "1" on the "Walking" item of the Functional Independence Measure (FIM) (4)) may be expected to walk short distances with contact guarding (level 4), but not to achieve full independence (level 7) by the end of the programme. In this case it is pertinent to record both the starting and the *expected* level, in order to determine whether the intended outcome was met and, for this reason, the UK FIM+FAM¹ recommends the recording of goal scores for all patients (5). GAS may be used, in this context, as an aid to negotiate realistic expectations of outcome.

Secondly, the practice of goal-setting is now well-established as a central part of rehabilitation (6), as it supports coordination of effort and because patients are more likely to engage actively in the programme if they perceive the treatment goals to be relevant (7). Fundamental to this approach, however, is the *collaborative* involvement of both the patient and the treating team in the goal-setting process. This supports the

development of a working partnership and a shared understanding of the agreed goals. By the same token, the evaluation of goal attainment should be undertaken collaboratively, the perspectives of *both patient and clinical team* having equal value. The involvement of patients with acquired brain injury presents some particular challenges for GAS, as cognitive and communicative problems may limit their ability to remember and articulate goals. Tight *a priori* definition of the agreed goals is therefore critical.

The study by Bovendeert et al. (2) has a number of design limitations, recognized by the authors, which illustrate the difficulties of applying GAS as part of blinded assessment. The treating therapists (who worked with the patients several times a week and were familiar with their actual abilities and performance) could allocate a goal score without much trouble, on the basis of observation and interaction over the preceding days. However the independent assessors did not have that advantage. Not only did they not know the patients, they were not allowed to consult either the treating team or *any of the clinical staff*. They had just one session in which to extract all the information to score a diverse set of goals, based on a combination of direct assessment and patient self-report. Some goals could not be assessed directly due to safety considerations, lack of the appropriate equipment, or because the goal related to a certain situation, which could not be reproduced within the assessment session. Under these circumstances they had to rely on the patient's verbal report, the accuracy of which will have been limited by cognitive and/or communicative deficits, at least in some patients. In addition, although the team attempted to record "SMART" goal statements, these may have been interpreted differently by the independent assessors, who relied purely on the written text and lacked the other general information about the patient that would inevitably be retained by the treating team. Therefore, as the authors rightly point out, the best information was not available to the blinded assessor; thus, it is not at all surprising that the two scores did not tally with one another; this was a comparison between "apples and pears".

The fact that poor reliability was seen between these two entirely different methods of GAS, therefore, should not be taken to mean that it is an unreliable measure. On the contrary, inter-rater reliability is shown to be good across a range of different settings, when GAS is applied by the same method (8). The demonstration of poor reliability here underlines the fact that, by its very nature, GAS requires the collaborative involvement of both the patient and their treating team, and the exclusion of one of these elements does not deliver the same results.

Does this mean that GAS can never be used in blinded RCTs? Not entirely. Where the intervention of interest is a blindable intervention (e.g. a drug), it is easy enough to blind the patient,

¹The UK FIM+FAM is the UK version of the Functional Assessment Measure with a further 12 items added to the FIM primarily addressing psychosocial function. It is designed primarily for use in acquired brain injury.

assessor and treating team to the nature of the treatment. In this case it may be perfectly acceptable for the treating team to carry out GAS rating, so that the critical scoring partnership between patient and team can be maintained. However, many physical interventions can never be fully concealed from the patient or team, in which case application by a blinded independent assessor offers the only real chance for reducing bias. We may have to accept that GAS could not be the *primary* outcome in such studies, but then, as Ertzgaard et al. (1) emphasize, it does not replace standardized outcome measures. By recording GAS (as evaluated by the patient and treating team) alongside other measures that are applied by a blinded independent assessor, it may still make a valid contribution in a supporting role. For example, McCrory et al. (9) used GAS as a secondary measure in an double-blind RCT for spasticity. GAS correlated strongly with reduction in spasticity (measured by the Modified Ashworth Scale) and both measures showed significant treatment effects between the active and placebo group. In this context, the standard measure demonstrated effectiveness of the intervention at the level of impairment, and GAS provided important confirmation (both quantitative and qualitative) of the functional benefits conferred by the active treatment (10).

But could GAS actually be applied through an independent assessor, as attempted in the Bovendeerd study? Learning from their unsuccessful approach, we could perhaps improve on that method. The principal problem appears to have been the complete exclusion of the treating team from the evaluation. Perhaps there are ways to include the perspective of the treating team at some level. There is a balance to be found between the risk of un-blinding and providing the assessor with enough of the relevant information to make a proper judgement, instead of giving them so little information that their evaluation amounts to little more than guessing. This is particularly important in the context of acquired brain injury, where the client group are expected to be poor witnesses by the very nature of their injury.

The incorporation of standardized measures into goal definitions may assist the process of independent GAS evaluation. As we become more experienced in using GAS in different areas of clinical practice, more limited “goal banks” are starting to emerge. Goals are still tailored to the individual, based on their current and expected level, but instead of recording entirely “free-flowing” individualized goals (which are often subjective and time-consuming to define), goal definitions are increasingly based on standard scales (such as a self-report scale of 0–10 for recording “pain” or “ease of care”). This not only supports clear objective goal-setting, but also speeds up the process of GAS application. For example, where pain reduction is a goal, a range of tools may be used to record pain levels (e.g. verbal, visual analogue, numbered graphic scales, “pain thermometer” Scale of Pain Intensity (11) etc.) according to the patient’s level of ability to report their symptoms (12). A common feature of all these scales, however, is that they provide a rating of 0–10 against which the various GAS levels of “–2” to “+2” may be defined, depending on the individual’s starting level. Providing all 5 goal levels are clearly identified *a priori* in a “follow-up guide” (as recommended by the originators, Kiresuk et al. (13))

and the method of assessment is clearly identified, it should then be relatively easy for an independent assessor to derive the GAS score from these more standard tools. In this way, a GAS T-score may be used to assimilate an overall estimate of achievement of the expected outcome across a range of different standardized measures (5), thus making it a more robust tool for the purposes of research.

In summary, the two articles presented in this issue provide important information about the use of GAS in neurological rehabilitation. Each, in its own way, takes us a step further in understanding what does and does not work in the application of GAS within clinical research. Clearly GAS cannot stand alone as a primary outcome measure, but both articles affirm its conceptual usefulness as a sensitive measure of relevant change in evaluation of complex interventions. Critically, it provides a person-centred perspective, as well as vital information to support interpretation of standardized outcomes in terms of what might reasonably be expected. Further exploration is now required to define parameters for its use in clinical trials, so that the full benefits of its inclusion can be retained, without compromise of its conceptual integrity as a measure of the achievement of expectation, applied through collaboration between the patient and the treating team.

ACKNOWLEDGEMENTS

Funding. Financial support for the preparation of this paper was kindly provided by the Luff Foundation and the Dunhill Medical Trust.

Competing interests. Outcome measurement is a specific research interest of Professor Turner-Stokes – in particular goal attainment scaling, for which her centre acts as a source of training and advice. However, no profits are made from its dissemination and she has no personal financial interests in the work undertaken or any findings reported.

REFERENCES

1. Ertzgaard P, Ward AB, Wissel J, Borg J. Practical considerations for goal attainment scaling during rehabilitation following acquired brain injury. *J Rehabil Med* 2011; 43: 8–14.
2. Bovend'Eert T, Dawes H, Hooshang I, Wade DT. Agreement between two different scoring procedures for goal attainment scaling is low. *J Rehabil Med* 2011; 43: 46–49.
3. Turner-Stokes L. Goal attainment scaling in rehabilitation; a practical guide. *Clin Rehabil* 2009; 23: 362–370.
4. Hamilton BB, Granger CV, Sherwin FS, Zielezny M, Tashman JS. A uniform national data system for medical rehabilitation. In: Fuhrer JM, editor. *Rehabilitation outcomes: analysis and measurement*. Baltimore: Brookes; 1987, p. 137–147.
5. Turner-Stokes L, Williams HJJ. Goal attainment scaling: does it provide added value as a person-centred measure for evaluation outcome in neurerehabilitation following acquired brain injury? *J Rehabil Med* 2009; 41: 528–535.
6. Wade DT. Evidence relating to goal-planning in rehabilitation. *Clin Rehabil* 1998; 12: 273–275.
7. Bovendeert TJ, Botell RE, Wade DT. Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide. *Clin Rehabil* 2009; 23: 352–361.
8. Hurn J, Kneebone I, Cropley M. Goal setting as an outcome measure: a systematic review. *Clin Rehabil* 2006; 20: 756–772.
9. McCrory P, Turner-Stokes L, Baguley IJ, De Graaff S, Katrak P, Sandanam J, et al. Botulinum toxin A for treatment of upper limb

- spasticity following stroke: a multi-centre randomised placebo-controlled study of the effects on quality of life and other person-centred outcomes. *J Rehabil Med* 2009; 41: 536–544.
10. Turner-Stokes L, Baguley I, De Graff S, McCrory P, Katrak P, Davies L, et al. Goal attainment scaling in the evaluation of treatment of upper limb spasticity with botulinum toxin: a secondary analysis from a double-blind placebo-controlled randomised clinical trial. *J Rehabil Med* 2010; 42: 81–89.
 11. Turner-Stokes L, Disler R, Shaw A, Williams H. Screening for pain in patients with cognitive and communication difficulties: evaluation of the SPIN-Screen. *Clin Med* 2008; 8: 393–398.
 12. Turner-Stokes L, Rusconi S. Screening for ability to complete a questionnaire: a preliminary evaluation of the AbilityQ and ShoulderQ for assessing shoulder pain in stroke patients. *Clin Rehabil* 2003; 17: 150–157.
 13. Cardillo JE, Choate RO. Illustrations of goal setting. In: Kiresuk T, Smith A, Cardillo J, editors. *Goal attainment scaling: application, theory and measurement*. New York: Lawrence Erlbaum Associates; 1994, p. 15–37.

Submitted November 15, 2010; accepted November 17, 2010

Lynne Turner-Stokes, MD, FRCP

King's College London School of Medicine,

Department of Palliative Care, Policy and Rehabilitation,

Cicely Saunders Institute, Bessemer Road,

London SE5 9PJ, UK.

E-mail: lynne.turner-stokes@dial.pipex.com