Age and Ageing 1999; **28:** 275–281

Goal attainment scaling as a measure of clinically important change in nursing-home patients

JANET E. GORDON, COLIN POWELL, KENNETH ROCKWOOD

Centre for Health Care of the Elderly, Queen Elizabeth II Health Sciences Centre, 5955 Jubilee Road, Halifax, Nova Scotia, B3H 2E1, Canada

Address correspondence to: K. Rockwood. Fax: (+1) 902 473 7133

Abstract

Objectives: to assess the feasibility, validity and responsiveness of an individualized measure—goal attainment scaling—in long-term care.
Design: prospective descriptive study.
Setting: one academic and three community-based long-term care facilities.
Subjects: 53 nursing-home patients seen in consultation between July 1996 and June 1997.
Intervention: specialized geriatric medicine consultation.
Main outcome measures: effect size and relative efficiency of the Barthel index, hierarchical assessment of balance and mobility, global deterioration scale, axis 8 (behaviour) of the brief cognitive rating scale, cumulative illness rating scale and the goal attainment scale.
Results: mean goal attainment scale at follow-up was 46 ± 7. The goal attainment scale was the most responsive measure, with an effect size of 1.29 and a relative efficiency of 53.7. The goal attainment scale did not correlate well with the other measures (-0.22 to 0.17).
Conclusion: goal attainment scaling is a feasible and responsive measure in long-term care. Although fewer problems in nursing-home patients than elderly inpatients are susceptible to intervention, clinically important goals can be achieved in this population.

Keywords: *goal attainment scaling, long-term care, nursing homes, outcome measurement*

Introduction

Evaluating the benefit of a geriatric consultation service in long-term care poses theoretical and practical challenges. The patient population is heterogeneous, medical interventions vary, and functional decline and death are expected. Standard measurement tools, such as the Barthel index [1], Katz index of activities of daily living [2] and the Mini-Mental State Exam (MMSE) [3], which have been used to evaluate geriatric consultations in acute care [4–6], are often irrelevant in long-term care. Improving mobility after stroke may be a goal in one patient; in another, successful palliation of bone pain may be accomplished without increasing mobility. Cognition may improve with treatment of the underlying cause of a delirium, whereas a dementing patient's cognition is expected to decline, whatever other goals might usefully be accomplished.

Goal attainment scaling is an individualized measurement developed in the 1960s to overcome similar limitations of traditional measures used in community mental health services. It has since been used in such diverse areas as drug abuse, marriage and family counselling, medical resident teaching and traumatic brain injury rehabilitation [7–10]. Latterly, goal attainment scaling has been applied to geriatric inpatients. In geriatric wards, it has been shown to be feasible, valid and responsive [11, 12]. Given the problems with standard measures in long-term care, we undertook a study to determine if the advantages of goal attainment scaling in acute care would apply in the nursing home.

Subjects and methods

Subjects

A prospective descriptive study of goal attainment scaling was conducted in patients seen by one of us

(J.E.G.) over 1 year (July 1996–June 1997) at one teaching (a veterans' facility) and three community-based long-term care facilities, ranging from 100 to 600 beds. In all cases, the populations served are elderly and disabled, with a high prevalence of dementia. Consultation was at the request of the patient's attending primary care physician. At two sites, these were conducted during regular fortnightly visits, while at the other two, consultations were on request. Follow-up monitoring of patients was undertaken during regular visits by the geriatricians and otherwise by ward nurses.

Sample and power

In earlier investigations, we found goal attainment scaling to be very responsive—with effect sizes ranging from 0.61 in a study of goal attainment scaling in an anti-dementia drug trial [13] to 4.60 in a study of goal attainment scaling in patients on acute and rehabilitation geriatric wards [12]. The effect size is calculated as the mean difference in pre- and post scores divided by the baseline standard deviation. It can be considered as a 'signal to noise' ratio, in that the absolute difference between the pre- and post-test measure is adjusted for the variability in the measures.

At their most basic, all statistics are effect sizes, themselves adjusted for sample size, so that when sample sizes are very large, even small effects become statistically significant. As compared with statistical significance, there are no clear criteria for when a given effect size becomes clinically important but, as a rule of thumb, Cohen has proposed that any effect size over 0.20 is clinically detectable. An effect size of 0.20 is said to be 'small', one of about 0.50 is 'medium' and one greater than 0.80 is 'large' [14]. Based on earlier studies, we sampled for an effect size of at least 0.80. According to the method of Lipsey [15] with $\alpha = 0.05$ and $\beta = 0.10$, we calculated that a sample size of 29 would be required to demonstrate responsiveness. To allow for refusals and incomplete follow-up and to maintain comparability with earlier studies, we set out to study 50 subjects, which would allow us to detect an effect size as small as 0.72, with $\beta = 0.10$.

Measures

All patients underwent a comprehensive assessment. Data were collected using standard scales for cognition, (MMSE and global deterioration scale [16]); behaviour (axis 8 of the brief cognitive rating scale [17]); co-morbidity (cumulative illness rating scale [18]); mobility and balance (hierarchical assessment of balance and mobility; HABAM [19]); and functional capacity (Barthel index). For the MMSE, HABAM and Barthel index, a higher score represents a better performance; for the others, a lower score indicates a better performance. Goal attainment scaling was conducted as follows:

1. Problems were identified and goals were determined for those areas in which intervention was planned. This always included the issues raised by the family physician's consultation request. Goals were set with the consensus of at least two geriatricians and usually with contributions from nurses.
2. The current description of the problem was recorded. If a clinically relevant deterioration were plausible, the current description was scored as '−1' with the worsened state scored as '−2'. If the problem was at its worst, the current level of functioning was scored as '−2'.
3. The expected outcome achievable with effective intervention was recorded in the '0' category.
4. The other scale levels were completed, setting outcomes somewhat more than expected at '+1' and much more than expected at '+2'. The aim was to make the goals as practical and verifiable as possible (Table 1).

Patients were followed until the goal was reached, was determined unachievable or the patient died. At the end of follow-up, the patient was scored at the level they had reached, and the overall goal attainment scaling score calculated. When goals are weighted equally, as in our study, the standard formula [7] can be simplified to: goal attainment score $= 50 + C\ (\Sigma x_i)$, where x_i is the score of the individual goal and C is a constant that varies with the number of goals set for that particular patient. For example, if one goal is set $C = 10$, if two are set $C = 6.2$, etc. This constant adjusts for the fact that different patients have different numbers of goals and that some of the goals are interrelated. For example, in the case illustrated in Table 1, improving the pain alone may have some beneficial effect on aggression. If all goals are achieved in a particular patient, each x_i will be 0 and therefore Σx_i will equal 0. As any term multiplied by 0 equals 0, the patient score will be $50 + 0 = 50$. For other scores, it not necessary to use the formula each time to calculate the goal attainment scaling score, as it can be obtained from a table (Appendix 1).

At follow-up, the global deterioration scale, axis 8 of the brief cognitive rating scale, cumulative illness rating scale, HABAM and Barthel index were also repeated, unless the patient had died.

Analysis

We assessed the validity of goal attainment scaling in the following ways, using a 'trinitarian approach' of content, construct and criterion validation [20]. The content validity was inferred from its use in other geriatric settings [11, 12]. There is no criterion reference, so criterion validity could not be tested.

Table 1. Sample goal attainment scaling in an 84-year-old man with advanced Alzheimer's disease and metastatic prostate cancer whose family wants 'everything done'

		Goal area		
	Score	Back pain	Level of care	Aggression
Much less than expected	−2	Cries out in pain constantly and/or does not respond to voice and light touch	Full cardiopulmonary resuscitation, transfer to intensive care, family's position is 'everything must be done'[a]	Injures a staff member
Somewhat less than expected	−1	Screams with pain with any movement, cries out ≥ 2 times per night[a]	No cardiopulmonary resuscitation, no intensive care unit but otherwise aggressive care and/or conflict within family about decision	Physically aggressive with care, nurses able to complete care only 50% of the time[a]
Programme goal	0	Sleeping through the night, minimal pain on transfers. Sleeping more through the day, but rouses easily[b]	No transfer to hospital, no investigations, family in acceptance[b]	Nurses able to complete care 75% of the time, verbally but not physically aggressive[b]
Somewhat better than expected	+1	No pain	Palliative care, no antibiotics for new infections, family in acceptance	Able to complete care 100% of the time
Much better than expected	+2	No pain, no sedation	—	No verbal aggression

[a]Level set at initial assessment.
[b]Level achieved.

The construct validity was assessed chiefly by correlation with other measures. To test convergent validity, Spearman correlations were calculated. Responsiveness, which is also a test of the construct validity of any measure purported to detect change, was determined in two ways. The effect size was calculated as the difference in post and pretest scores divided by the standard deviation of the change score. While we have previously used the baseline standard deviation as the variance term in calculating effect size [12, 13, 21], the inherently small standard deviation of the pre-test goal attainment scaling score means that the effect size of goal attainment scaling is over-estimated. In consequence, this method provides a more conservative estimate of the responsiveness of goal attainment scaling. In addition, we estimated responsiveness using the relative efficiency statistic, in which the *t*-test statistic of any measure is expressed as the ratio of the *t*-test statistic of an index or comparative measure [22]. As in an earlier study [12], we chose the Barthel index as the comparative measure. A score of 1.00 indicates the same efficiency as the standard, greater than 1.00 is more efficient and less than 1.00 is less efficient than the standard.

Results

Fifty-three patients (mean age 81 ± 8 years; 33 women) were evaluated using Comprehensive Geriatric Assessment, after which goals were set for all patients. They were followed for a mean of 44 ± 44 days. Seventeen percent of patients died during this period. Most patients (77%) had dementia, all were disabled (mean Barthel index 47 ± 33) and the average length of residence in the nursing home was 2.5 ± 2.4 years.

Table 2 reports the mean pre- and post-test results, with standard deviations, of each measure. Table 3 is a correlation matrix of the various measures. Goal attainment scaling did not correlate well with any other measure (−0.22 to 0.17).

The assessment process identified 463 problems, of which goals were set in 89. The most common problems identified for intervention were medical problems (47), behavioural problems (25) and ethical problems (17). Examples of problems which were identified, but for which no goals were set, included dementia, stable medical problems and incontinence or immobility with no clinically apparent potential for improvement. The mean number of goals per patient was 1.7, with a range of 1–5. Once the assessment was completed, the goals were usually set in less than 10 min. The mean goal attainment scaling score at admission was 37 ± 3.5 and at follow-up 46 ± 6.9 ($t = -9.13$; $P < 0.0001$).

Table 4 displays the responsiveness of the various measures. Goal attainment scaling was the most

Table 2. Mean scores of selected assessment instruments, pre- and post-consultation

Measure	Value indicating		Mean score ± SD	
	Worst factor	Best factor	Pre-consultation (n = 53)	Post-consultation (n = 44)[b]
Barthel index	0	100	46.7 ± 32.9	51.2 ± 33.7
HABAM	0	25	13.4 ± 8.7	14.6 ± 8.9
Global deterioration scale	7	1	5.4 ± 1.5	5.5 ± 1.4
Cumulative Illness Rating Scale	56	0	7.6 ± 3.3	6.8 ± 2.8
Axis 8 (behaviour) of the BCRS	7	1	5.5 ± 1.4	5.2 ± 1.6
Goal attainment scaling[a]	20	80	37.3 ± 3.5	45.7 ± 6.9[c]

[a]In goal attainment scaling the best/worst score was calculated as the score for +2/−2 on each of a theoretical maximum of five goals per patient. A score of 50 would indicate the achievement of all goals.
[b]Nine patients died during follow-up.
[c]We calculated goal attainment scale scores for patients who died; when death was unexpected, it was scored in each domain as −2 (much worse than expected). When death occurred some months after, goals were scored as achieved. Where the goals were for palliative care, these were scored as appropriate, if achieved.
BCRS, brief cognitive rating scale; HABAM, hierarchical assessment of balance and mobility.

responsive measure, with the highest effect size (1.29) and the highest relative efficiency (53.7).

Discussion

We investigated goal attainment scaling as a method of detecting clinically important change in nursing-home patients. These patients had multiple, often untreatable functional and medical problems. The mean MMSE score (12 ± 11) and the mean Barthel index (47 ± 33) were lower than in the acutely ill geriatric medicine inpatients (22 ± 7 and 57 ± 24 respectively) whom we studied earlier [12]. In that population, 5.8 goals were set per patient; in this group, in spite of greater frailty [23], we decided we could successfully meet 1.7 goals per patient. As in earlier studies, we chose the Barthel as the reference measure, thus allowing the data to be compared across studies. In contrast to our earlier study [12], however, we did not expect the Barthel index to change. Indeed, it was the lack of responsiveness in long-term care patients, in whom successful treatment can nevertheless be undertaken, that gave the impetus for this study.

One limitation of our study is the absence of a criterion reference to confirm the clinical relevance of a change in goal attainment scaling score. This remains a clinical judgement and one could argue that other doctors would set different goals. We sought to safeguard against arbitrariness by always determining goals with at least two geriatricians and, usually, the nursing staff. Clinical importance can also be inferred from the family physician's consultation request always being addressed as a goal area.

In other settings, patients have been involved in setting their own goals, both to ensure relevance and as a therapeutic tool. However, in our study the high prevalence of dementia did not make this feasible.

Some readers might worry that we set goals at too easy a level, so as to ensure success. There are safeguards against this inherent in the process. If goals are consistently set too low, consistently high scores (i.e. > 50) will result. In addition, the detailed, explicit

Table 3. Spearman rank correlations of outcome measure change scores pre- and post-consultation with the goal attainment scaling (GAS) follow-up score

	Correlation			
	GAS	Barthel index	HABAM score	CIRS
Barthel index	0.1484			
HABAM score	0.1690	0.6900		
CIRS	−0.1649	−0.3786	−0.3210	
Axis 8 of BCRS	−0.2158	−0.1903	−0.0391	0.1850

CIRS, cumulative illness rating scale; BCRS, brief cognitive rating scale; HABAM, hierarchical assessment of balance and mobility.

Table 4. Effect size and relative efficiency of outcome measures

	Effect size	Relative efficiency
Barthel index	0.025	1
HABAM score	0.14	5.3
Global deterioration scale	0	–
CIRS	−0.25	−9.7
Axis 8 (behaviour) of BCRS	−0.41	−15.9
Goal attainment scaling	1.29	53.7

CIRS, cumulative illness rating scale; BCRS, brief cognitive rating scale; HABAM, hierarchical assessment of balance and mobility.

nature of the goals readily exposes them for peer review. We did not test the inter-rater reliability of goal attainment scaling in this study. However, in earlier studies of elderly patients, it had been excellent ($\kappa >$ 0.80) [11, 12].

Goal attainment scaling did not correlate well with the other standard measures (Table 3). In contrast, in inpatients on a geriatric unit, goal attainment scaling correlated moderately well with functional scales (0.59 for the Barthel index, 0.49 for the Katz index of activities of daily living) [12]. Many patients on rehabilitation units are recovering from surgery for a hip fracture and are expected to make significant gains in mobility and activities of daily living. On acute geriatric units, most patients have acute medical illnesses from which they are expected to recover. In nursing-home patients with multiple chronic medical and functional problems, it is not surprising that no single global measure is responsive to change. We believe goal attainment scaling to be capturing different information. For example, in one of our patients, who was wheelchair-bound, establishing end-of-life decisions, behaviour modification and even controlling pain did not improve independence in activities of daily living, even though all the goals achieved were clinically important.

If evaluation is limited to global measures, which are standardized but irrelevant to the needs of the patients, successful intervention is likely to be unrecognized. In contrast, when a clinically relevant but individualized measure is used, the achievement of important goals can be demonstrated when treatment is successful.

These data support studies which suggest that the multiple needs of elderly patients, which require a multi-disciplinary approach, equally require a multidimensional evaluation. Cross-validation of the approach in other institutions would be of interest.

Acknowledgements

J.E.G. is the recipient of the Dalhousie University Joseph P. Shannon Bursary in Geriatric Medicine. K.R. is the recipient of a National Health Scholar Award from the National Health Research and Development Program, Health Canada.

Key points

- Global measures such as the Barthel index are often not relevant for nursing-home patients being assessed in geriatric consultation.
- Goal attainment scaling is an individualized measure which has been used in community mental health and geriatric inpatient units; this study reports its use in nursing-home patients.
- In 53 consultations in long-term care, an average of 1.7 goals per patient were set, covering medical, behavioural and ethical issues.
- A mean follow-up goal attainment score of 46 demonstrates that most of these goals were met.
- Goal attainment scaling is more responsive to change than any of the other measures examined.

References

1. Granger CV, Albrecht GL, Hamilton BB. Outcome of comprehensive medical rehabilitation. Arch Phys Med Rehabil 1979; 60: 145–54.

2. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA 1963; 185: 914–9.

3. Folstein MF, Folstein SE, McHugh PR. Mini-mental state—a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12: 189–98.

4. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. Lancet 1993; 342: 1032–6.

5. Winograd CH, Gerety MB, Lai NA. A negative trial of inpatient geriatric consultation: lessons learned and recommendations for future research. Arch Intern Med 1993; 153: 2017–23.

6. Reuben DB, Borok GM, Wolde-Tsadik G *et al.* A randomized trial of comprehensive geriatric assessment in the care of hospitalized patients. N Engl J Med 1995; 332: 1345–50.

7. Kiresuk TJ, Smith A, Cardillo JE. Goal Attainment Scaling: applications, theory and measurement. Hillsdale, NJ: Lawrence Erlbaum, 1994.

8. Malec JF, Smigielski JS, DePompolo RW. Goal attainment scaling and outcome measurement in postacute brain injury rehabilitation. Arch Phys Med Rehabil 1991; 72: 138–43.

9. Joyce BM, Rockwood K, Mate-Cole C. Use of goal attainment scaling in brain injury in a rehabilitation hospital. Am J Phys Med Rehabil 1994; 73: 10–4.

10. Rockwood K, Joyce B, Stolee P. Use of goal attainment scaling in measuring clinically important change in cognitive rehabilitation patients. J Clin Epidemiol 1997; 50: 581–8.

11. Stolee P, Rockwood K, Fox RA, Streiner DL. The use of goal attainment scaling in a geriatric care setting. J Am Geriatr Soc 1992; 40: 574–78.

12. Rockwood K, Stolee P, Fox RA. Use of goal attainment scaling in measuring clinically important change in the frail elderly. J Clin Epidemiol 1993; 46: 1113–8.

13. Rockwood K, Stolee P, Howard K, Mallery L. Use of goal attainment scaling to measure treatment effects in an anti-dementia drug trial. Neuroepidemiology 1996; 15: 330–8.

14. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum, 1988: 15–27.

15. Lipsey MW. Design Sensitivity: statistical power for experimental research. Newbury Park, CA: Sage, 1990; 91.

16. Reisberg B, Ferris SH, de Leon MJ, Crook T. The global deterioration scale for assessment of primary degenerative dementia. Am J Psychiatry 1982; 139: 1136–9.

17. Reisberg B, Ferris S. The brief cognitive rating scale. Psychopharmacol Bull 1988; 24: 629.

18. Parmelee PA, Thuras PD, Katz IR, Lawton MP. Validation of the cumulative illness rating scale in a geriatric residential population. J Am Geriatr Soc 1995; 43: 130–7.

19. MacKnight C, Rockwood K. A hierarchical assessment of balance and mobility. Age Ageing 1995; 24: 126–30.

20. Streiner DL, Norman GR. Health Measurement Scales, 2nd edition. Oxford: Oxford University Press, 1996.

21. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27: S178–89.

22. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985; 28: 542–7.

23. Rockwood K, Fox RA, Stolee P, Robertson D, Beattie BL. Frailty in elderly people: an evolving concept. Can Med Assoc J 1994; 150: 489–95.

Received 20 August 1997; accepted in revised form 11 June 1998

Appendix 1. Calculated goal attainment scores

	No. of goals							
Sum score	1	2	3	4	5	6	7	8
−16								18
−15								20
−14							18	22
−13							21	24
−12						19	23	26
−11						22	25	28
−10					20	24	27	30
−9					23	27	30	32
−8				21	26	29	32	34
−-7				25	29	32	34	36
−6			23	28	32	35	36	38
−5			27	32	35	37	39	40
−4		25	32	35	38	40	41	42
−-3		31	36	39	41	42	43	44
−2	30	38	41	43	44	45	45	46
−1	40	44	45	46	47	47	48	48
0	50	50	50	50	50	50	50	50
1	60	56	55	54	53	53	52	52
2	70	62	59	57	56	55	55	54
3		69	64	61	59	58	57	56
4		75	68	65	62	60	59	58
5			73	68	65	63	61	60
6			77	72	68	65	64	62
7				76	71	68	66	64
8				79	74	71	68	66
9					77	73	70	68
10					80	76	73	70
11						78	75	72
12						81	77	74
13							79	76
14							82	78
15								80
16								82

Appendix 2. Specific goals set in 53 patients

Medical goals (47)
Treatment of:
- Polypharmacy (11)
- Depression (8)
- Pain control (6)
- Reflux (3)
- Constipation (2)
- Diarrhoea (1)
- Gastrointestinal bleed (1)
- Fever (1)
- Hypotension (1)
- Dyspnoea (2)
- Diabetes (1)
- Cellulitis (1)
- Parkinson's disease (2)
- Huntington's disease (1)
- Immobility (2)
- Pruritus (2)
- Thrombocytopaenia (1)
- Hyponatraemia (1)

Behavioural goals (25)
Management of:
- Aggression towards other residents (5)
- Aggression during personal care (11)
- Screaming (2)
- Pacing, restlessness (3)
- Inappropriate sexual behaviour (2)
- Delusions (1)
- Hallucinations (1)

Ethical goals (17)
Questions about:
- Cardiopulmonary resuscitation (2)
- Transfer to hospital (3)
- Aggressiveness of investigations (3)
- Pursuing active treatment (5)
- Poor feeding (2)
- Competence regarding place of residence (1)
- Suitability of placement (1)

'Slow Bull's Wife' One of thousands of photographs by Edward S. Curtis for his multivolume work *The North American Indian* (1907–1930). Photography reproduced courtesy of University of Exeter Library.