Clinical Endocrinology (2003) **58**, 436–445

Evaluation of two health status measures in adults with growth hormone deficiency

C. V. McMillan*, C. Bradley*, J. Gibney†, D. L. Russell-Jones‡ and P. H. Sönksen†

**Royal Holloway, University of London, Egham, Surrey, †St Thomas' Hospital, London and ‡Royal Surrey County Hospital, Guildford, UK*

(Received 6 June 2002; returned for revision 15 July 2002; finally revised 4 October 2002; accepted 8 November 2002)

Summary

OBJECTIVE **To evaluate the psychometric properties of two health status measures for adults with growth hormone deficiency (GHD): Nottingham Health Profile (NHP) and Short-Form Health Survey (SF-36).**

DESIGN **(1) A cross-sectional survey of adults with treated or untreated GHD to assess reliability and validity of the questionnaires. (2) A randomized, placebo-controlled study of 3 months' GH withdrawal from GH-treated adults to assess the sensitivity of the questionnaires to change.**

PATIENTS **(1) A cross-sectional survey of 157 patients with severe GHD (peak GH < 10 mU/l on provocative testing), mean age 48·9 years (range 23–70 years), who had either received GH replacement therapy for at least 6 months immediately prior to the study or had not received GH treatment in the previous 6 months. (2) GH treatment was withdrawn from 12 of 21 GH-treated adults, all with severe GHD (peak GH < 7·7 mU/l on provocative testing), mean age 44·9 years (range 25–68 years).**

MEASUREMENTS **The NHP and SF-36 were used once in the cross-sectional survey, but twice in the GH-withdrawal study, at baseline and end-point (after 3 months).**

RESULTS **(1) Cross-sectional survey. Both questionnaires had high internal consistency reliability with subscale Cronbach's alphas of > 0·73 (NHP) and > 0·78 (SF-36). Calculation of an NHP Total Score, occasionally reported in the literature, was shown to be inadvisable. Overall, patients with GHD were found to have significantly worse perceived functioning than the UK general population in SF-36 subscales of General Health, Bodily Pain, Social Functioning, Physical Functioning, Role-Emotional, Role-Physical, and Vitality. Although neither questionnaire found significant differences between GH-treated and non-GH-treated patients, there were correlations with duration of GH treatment ($P < 0{\cdot}01$) for GH-treated patients in SF-36 Mental Health ($r = 0{\cdot}29$, $N = 87$) and SF-36 Vitality ($r = 0{\cdot}33$, $N = 88$), indicating improvement with increasing treatment duration. The SF-36 was also more sensitive than the NHP to sex differences: men had significantly better health status compared with women ($P < 0{\cdot}05$) in all SF-36 subscales but Mental Health, but only in one NHP subscale (Physical Mobility). (2) GH-withdrawal study. Significant between-group differences in change were found in SF-36 General Health [$t(17) = 2{\cdot}76$, $P = 0{\cdot}013$, two-tailed] and SF-36 Mental Health [$t(17) = 2{\cdot}41$, $P = 0{\cdot}027$, two-tailed]: patients withdrawn from GH reported reduced general health and mental health at end-point. The NHP found no significant change.**

CONCLUSIONS **The SF-36 is a better measure than the NHP of health status of people with GH deficiency because of its greater discriminatory power, with ability to detect lesser degrees of disability. It also has superior sensitivity to some subgroup differences and superior sensitivity to change compared with the NHP. The SF-36 is highly acceptable to respondents, and has very good internal consistency reliability. The SF-36 is recommended to measure the health status of adults with GH deficiency.**

The physical symptoms of adult growth hormone deficiency (GHD) include abnormal body composition with reduced lean body mass and increased central adiposity; reduced muscle strength and exercise performance (Carroll *et al*., 1998). However, psychological symptoms may be as important as physiological (Powrie *et al*., 1995) and it has been recommended that psychological variables should be considered when assessing patients for treatment (Bengtsson *et al*., 2000). Psychological symptoms reported by adults with untreated GHD include low energy, tiredness, sleepiness, poor concentration, poor memory, irritability (Hunt *et al*., 1993), anxiety, depression and mood swings (Wallymahmed *et al*., 1996).

Correspondence: Dr C. V. McMillan, Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.
Tel.: + 44 1784 443915; Fax: + 44 1784 434347;
E-mail: c.mcmillan@rhul.ac.uk

Health status has been one of the key psychological outcomes in adult GHD research, and the Nottingham Health Profile (NHP) (Hunt *et al*., 1985) the most frequently used health status questionnaire with this patient group. The NHP has six subscales to measure Emotional Reactions, Energy, Pain, Physical Mobility, Sleep and Social Isolation. After 6 months of GH treatment in placebo-controlled studies, improvements have been found in Energy (McGauley *et al*., 1990; Mardh *et al*., 1994; Carroll *et al*., 1997; Wallymahmed *et al*., 1997), Emotional Reactions (Carroll *et al*., 1997), Physical Mobility and Social Isolation (Attanasio *et al*., 1997; Carroll *et al*., 1997). One study only obtained significant improvement in Emotional Reactions and Energy when patients with floor baseline scores (good health status) were excluded from analysis (Burman *et al*., 1995). Others only found improvements in some NHP variables in the open label stages of the trial after more than 12 months of GH treatment (Mardh *et al*., 1994; Cuneo *et al*., 1998) or after 20–50 months (Wiren *et al*., 1998). Some studies found no significant effects using the NHP (Whitehead *et al*., 1992; Baum *et al*., 1998). Thus energy has generally been found to improve when patients with GHD are given GH replacement therapy, but results for other aspects of health status have not always been consistent. The Short-Form Health Survey (SF-36), a more recent measure of health status (Ware & Sherbourne, 1992), has eight subscales to measure Bodily Pain, General Health, Mental Health, Physical Functioning, Role-Emotional, Role-Physical, Social Functioning and Vitality. It has been used less frequently than the NHP in research into adult GHD, generally in comparison studies.

Both the NHP and the SF-36 measure how people feel about their health, but they have often been inappropriately referred to as measures of quality of life (QoL) in the research literature, where the terms 'health status' and 'quality of life' are frequently misused and treated as if they were interchangeable (Smith *et al*., 1999; Bradley, 2001). Even if people feel that their health is poor, they may or may not also feel that their QoL is impaired, and vice versa. Treatments that improve health may conversely damage QoL, and may lead to poor adherence to therapy. 'When a health-status measure is used to assess quality of life, the conclusions can be misleading' (Bradley, 2001).

GH replacement therapy in adult GHD is not the norm in the UK, as there is controversy about whether the benefits outweigh the high financial costs and a treatment regimen requiring daily injections. Measures of psychological outcomes are required in addition to clinical outcomes to assess the effects of GH treatment. The frequently used NHP does not appear to have been previously validated psychometrically for use in adult GHD, and the opportunity was taken in the two studies described in the following, not only to validate the NHP but also to compare its properties with those of the SF-36, which has been shown to be more sensitive than the NHP in the general population (Brazier *et al*., 1992). The first study was a cross-sectional survey of 157 adults with severe GHD, both GH-treated and non-GH-treated, to investigate reliability, factor structure and construct validity of the questionnaires. Sensitivity to change was investigated in a randomized placebo-controlled study of 3 months' withdrawal of GH treatment from 12 of 21 GH-treated adults, where nine continued with GH. The Guy's and St Thomas' Hospital Trust Ethics Committee gave approval for both studies.

The questionnaires

NHP

The NHP has 38 dichotomous negatively worded items. 'No' responses (indicating good functioning) score zero, 'yes' responses can be weighted as described in the manual (Hunt & McKenna, 1989). Subscale scores range from 0 to 100 (lower scores indicating better health status).

SF-36

This 36-item health survey has eight subscales with Likert scales of two to six response options and an additional item on perceived change in health over the previous year. SF-36 subscale scores range from 0 to 100 (higher scores indicating better functioning). The UK version was used in the present studies. It was found to have high reliability with a UK general population (Brazier *et al*., 1992), and was able to detect low levels of ill health in people who had scored 0 (indicating good health) on the NHP.

Both questionnaires have been used in the endocrine field; for example, in hormone replacement therapy (Ryan & Rosner, 2001; Kenny *et al*., 2002), diabetes (Anderson *et al*., 1997; Benbow *et al*., 1998), thyroid disease (Burney *et al*., 1999), and Addison's disease (Lovas *et al*., 2002).

Other questionnaires

Other questionnaires were also completed in the present studies, including the General Well-being Index (Hunt & McKenna, 1992), which is the British version of the Psychological General Well-being Index (Dupuy, 1984), the Well-being Questionnaire (Bradley, 1994), and a new hormone deficiency-specific individualized QoL questionnaire (HDQoL) (Bradley, 1999), the results for which have been reported previously (McMillan & Bradley, 2000; McMillan, 2001; McMillan *et al*., 2001).

The cross-sectional survey of GH-treated and non-GH-treated adults with GHD

Recruitment procedures

All participating patients had been diagnosed with GHD and received GH replacement therapy for at least 6 months immediately prior

to the study or had not received GH treatment in the previous 6 months; they were aged between 18 and 70 years; and had received appropriate adrenal, thyroid and gonadal hormone replacement therapy, as required by their hormonal condition, for at least 12 months prior to the study. Patients might have adult-onset GHD (AOGHD) or childhood-onset GHD (COGHD). Exclusion criteria were diabetes mellitus, active malignancy or pregnancy. The patients were to be severely GH deficient, as determined by an insulin tolerance test (ITT) or pituitary function test in which insulin reduced blood glucose to ≤ 2·5 mmol/l with peak GH concentration ≤ 10 mU/l. However, there were no ITT results available for 21 patients because either an ITT was too dangerous, or an ITT had never been performed or the laboratory test results were missing from the patient notes. In 18 of these 21 cases the assumption was made that because the patients were already receiving GH treatment, were about to be given GH treatment, or had received it in the past as adults, then the severity of the deficiency had been established, and they could be included in this study. In three cases in the non-GH-treated group, clinicians were consulted about the patient's suitability for inclusion. Some patients were approached in a mailshot enclosing the questionnaires, others at weekly clinics.

Statistical analyses

Normality issues. Normality of distributions was investigated through standardized *z* (skew) values (Tabachnik & Fidell, 1983). Item data were not transformed to normality, thereby sacrificing some of the accuracy of factor analyses for the convenience of having interpretability of original units. The assumption was made that if reliability were high, the factor analysis robust, and the number of respondents sufficiently high, then a degree of non-normality was acceptable. In subgroup analyses, Mann–Whitney tests were performed on skewed variables, and *t*-tests on normal data.

Internal consistency reliability. Cronbach's alpha coefficients (Cronbach, 1951) were determined, with acceptable alphas ranging from 0·7 (Todd & Bradley, 1994) to 0·95. Acceptable item-total correlations were > 0·2 (Kline, 1993).

Factor structure. This was explored with principal components analysis. Forced one-factor solutions were obtained for individual subscales to confirm the validity of calculating subscale totals. Salient loadings were taken as ≥ 0·4, higher than the recommended minimum 0·3 (Kline, 1994), erring on the side of caution in an effort to reduce the risk of spurious loadings that owed their origin to non-normality of item distributions.

Construct validity. This was sought in some expected subgroup differences (GH-treated patients having generally better health status than non-GH-treated patients, and men better than women). Health status was expected to improve with increasing duration of treatment and correlations were undertaken (for GH-treated patients).

'Familywise' error in multiple tests. The Holm's sequential Bonferroni procedure for multiple tests, cited in Green *et al.* (1997), was adopted. A minimum significance value of 0·006 was required if eight similar statistical tests were performed on the eight subscales of the SF-36 and 0·008 for the six-subscale NHP.

Analysis was conducted using SPSS for Windows (Release 7·5).

Results

The patient sample

Of 219 questionnaires distributed, 163 were returned (74·4% response rate), but six patients did not meet all inclusion criteria (e.g. were found to have diabetes mellitus), leaving 157 data sets for analysis. Fifty-six people either declined to participate or did not respond (23·5% of GH-treated and 29·8% of non-GH-treated patients).

Ninety-one GH-treated and 66 non-GH-treated patients participated. *T*-tests showed no significant between-treatment-group differences in body mass index (BMI) or height (Table 1); χ^2-tests showed no differences in numbers of men but the difference in numbers of GH-treated women (51) and non-GH-treated women (33) was significant. The mean age of GH-treated patients (47·1 years) was significantly lower than non-GH-treated patients (51·3 years). There were significantly more people with COGHD in the GH-treated group (21) than the non-GH-treated group (9).

Questionnaire completion rates were high: 98·5% (NHP) and 99·1% (SF-36) indicating high acceptability to respondents. Non-normality (distributions skewed towards good health status), was found in a much greater proportion of NHP items (36/38) than SF-36 items (20/36). In fact, 31/38 dichotomous NHP items had ≥ 75% 'No' responses, indicating good health, and 23·7% of respondents scored 'No' on all 38 NHP items. The highest levels of dysfunction in the whole patient sample, found by the NHP, were for Energy (mean 34·28 ± 39·48, $N = 153$) and Sleep (21·34 ± 30·65, $N = 153$). The SF-36 found the highest dysfunction in Vitality (mean 49·59 ± 23·8) and General Health (56·0 ± 25·04). However, when compared with the results for the UK general population (Garratt *et al.*, 1993), the patient sample as a whole was found to have significantly worse perceived functioning in SF-36 subscales of General Health, Bodily Pain, Social Functioning, Physical Functioning, Roles Emotional and Physical, and Vitality (Table 2). These results were similar to those found by other researchers when comparing healthy controls and patients with GHD, e.g. in the Belgian population (Hakkaart-van Roijen *et al.*, 1998).

Table 1 Characteristics of patients in the cross-sectional survey

	GH treatment ($N = 91$)	No GH treatment ($N = 66$)	P
Women	51	33	0·05
Men	40	33	
COGHD	21	9	< 0·05
AOGHD	70	57	
Isolated GHD	5	1	
Multiple hormone deficiency	86	65	
Mean age (years) [range]	47·1 ± 12·6 [23·7–70·9]	51·3 ± 12·4 [23·8–70·9]	< 0·05
Mean duration GHD (years) (AOGHD)	13·0 ± 6·8	13·2 ± 7·8	
BMI (kg/m^2)	27·2 ± 5·5	28·0 ± 5·3	
Height (cm)	167·5 ± 10·5	168·7 ± 10·6	

Table 2 SF-36 subscale means and comparison with UK general population

Subscale	N	Mean	UK mean*	P
Bodily Pain	156	68·99 ± 7·48	76·9	< 0·001
General Health	152	56·00 ± 25·04	68·7	< 0·001
Mental Health	156	70·79 ± 19·56	73·7	
Physical Functioning	157	72·06 ± 24·22	79·2	< 0·001
Role-Emotional	156	64·96 ± 39·49	75·0	< 0·01
Role-Physical	155	58·92 ± 40·23	76·5	< 0·001
Social Functioning	156	73·56 ± 26·86	78·6	< 0·05
Vitality	157	49·59 ± 23·8	61·2	< 0·001

Subscale range 0–100, with higher score indicating better health status/functioning.
*Means for a healthy UK general population, from Garratt *et al*. (1993).

Reliability analyses

NHP. All subscale alpha coefficients were acceptable (range 0·74–0·89) and corrected item-total correlations within subscales were satisfactory (range 0·22–0·62). Reliability analysis of the whole NHP scale found a high Cronbach's alpha (0·92) but also indicated the unacceptability of the NHP Total Score occasionally reported in the literature, as corrected item-total correlations for two items failed to reach the 0·2 criterion.

SF-36. All subscale alphas were high for short scales of from two to 10 items (range 0·78–0·94) and corrected item-total correlations within subscales were satisfactory (range 0·43–0·89).

Factor analysis

NHP. Forced one-factor analyses of the individual subscales found satisfactory loadings ≥ 0·4 for all except one Physical Mobility item that loaded at 0·321. A forced one-factor analysis of the whole scale found 10/38 items loading at < 0·4, again indicating the invalidity of calculating an overall NHP Total Score.

SF-36. Forced one-factor analyses of subscales found all loadings satisfactory (≥ 0·57).

With the exception of NHP Physical Mobility, analyses supported the calculation of NHP and SF-36 subscale scores for adults with GHD.

Subgroup differences: GH-treatment group differences

NHP. There were no significant differences on Mann–Whitney tests between the treatment groups (P-values in the range 0·16 to 0·99), although trends were in the expected direction of those receiving GH treatment having better health status than non-GH-treated patients on all subscales except Social Isolation.

SF-36. T-tests were performed on all SF-36 subscales but the skewed Physical and Social Functioning subscales, for which Mann–Whitney tests were conducted. There were no significant between-treatment-group differences: P-values were in the range 0·09 (Mental Health) to 0·83 (Role-Emotional). GH-treated patients had slightly better functioning than non-GH-treated patients in all subscales but Bodily Pain, Mental Health and Physical Functioning. A further analysis, excluding the 21 patients who did not meet the initial inclusion criteria (no ITT test result available), did not find any significant between-treatment-group differences.

Correlations with duration of GH treatment in GH-treated patients

NHP. No correlations were significant once Bonferroni corrections had been applied.

SF-36. Mental Health correlated significantly with duration of GH treatment ($r = 0{\cdot}29$, $P = 0{\cdot}007$, $N = 87$) as did Vitality ($r = 0{\cdot}33$, $P = 0{\cdot}002$, $N = 88$), indicating, as expected, improved mental health and vitality with increased duration of treatment.

Age of onset of GHD

Neither questionnaire found significant differences in health status between those with AOGHD or COGHD once Bonferroni corrections had been applied (full results not supplied).

Sex differences

Men had significantly better NHP Physical Mobility and significantly better health status in all SF-36 subscales except for Mental Health compared with women (Table 3). These are similar to results for the UK general population (Brazier *et al*., 1992).

The randomized placebo-controlled trial of withdrawal of GH treatment from GH-treated adults

Study design

Patients were allocated to placebo or continued treatment with GH in a randomized, double-blind, placebo-controlled, parallel-groups study, in which patients self-administered either GH (0·125–0·25 IU/kg body weight/week) or placebo for a period of 3 months. Lilly Industries Ltd supplied the vials of GH (Humatrope), which were indistinguishable from placebo. IGF-I was measured by double-antibody radioimmunoassay after acid/ethanol extraction, using a commercially available reagent pack [Amersham, Arlington Hts, III, within-assay coefficient of variance (CV) < 5%]. Questionnaires were given at baseline and end-point, 3 months later. All patients were attending the Endocrine Clinic of St Thomas' Hospital, London.

Table 3 Means for men and women in the cross-sectional survey

	Men	Women	*P*
NHP			
Emotional Reactions	11·92 ± 18·49	19·79 ± 27·37	
Energy	26·39 ± 35·9	41·28 ± 41·37	
Pain	7·70 ± 19·6	13·91 ± 24·97	
Physical Mobility	7·44 ± 15·96	12·59 ± 18·01	< 0·01
Sleep	17·70 ± 29·06	24·65 ± 31·85	
Social Isolation	13·58 ± 23·78	16·98 ± 26·2	
SF-36			
Bodily Pain	78·06 ± 23·99	61·23 ± 28·03	< 0·001
General Health	62·28 ± 22·87	50·35 ± 25·70	< 0·01
Mental Health	73·81 ± 18·32	68·14 ± 20·33	
Physical Functioning	80·43 ± 19·75	64·79 ± 25·47	< 0·001
Role-Emotional	73·97 ± 35·24	57·03 ± 41·49	< 0·05
Role-Physical	71·00 ± 37·84	48·17 ± 39·44	< 0·001
Social Functioning	78·60 ± 25·89	69·13 ± 27·07	< 0·05
Vitality	54·38 ± 23·76	45·42 ± 23·18	< 0·05
Minimum *N*	68 (NHP) 72 (SF-36)	77 (NHP) 80 (SF-36)	

Subscale scores range 0–100, higher score indicating better health status on SF-36, but poorer health status on NHP.
Values are means ± SD.

Recruitment procedures

Inclusion and exclusion criteria were similar to those in the cross-sectional study except that all participating patients had received GH replacement therapy for at least 6 months immediately prior to the study; were aged between 22 and 70 years; and were taking adequate contraception, if women of childbearing age and potential. All patients had endogenous GH levels of ≤ 7·7 mU/l at blood glucose ≤ 2·0 mmol/l on provocative testing. Additional exclusion criteria were: clinically significant pulmonary, cardiothoracic, renal or neuromuscular disease; clinically apparent chromosomal or genetic malformation syndromes; a history of alcohol or drug abuse; were unlikely to comply with the protocol; were taking oral and parenteral steroids, other than in replacement doses, that were likely to suppress the action of GH. Prospective participants were approached personally by clinicians.

Additional measures

Interviews were also conducted, results for which have been reported (McMillan, 2001). The number of serious negative life events and difficulties occurring in the 12 months prior to this study was assessed in the interviews using a short checklist (unpublished, available from the author, Dr Bernice Andrews, at the Department of Psychology, Royal Holloway, UK), modified from a screening checklist for stressful life events and chronic difficulties (Costello & Devins, 1998).

Hypotheses

When adults with GHD are given GH treatment, improvements in health status have been reported. There was general clinical expectation of deterioration in physiological/metabolic factors within the 3-month study period and that this might be accompanied by reduced health status (Sönksen *et al*., 1991). It was therefore hypothesized that after 3 months' withdrawal of GH the placebo group would exhibit, relative to baseline, decreased Energy, Emotional Reactions, Physical Mobility, Sleep and increased Social Isolation (higher scores on NHP); and decreased General Health, Mental Health, Physical Functioning, Social Functioning, and Vitality (reduced scores on SF-36), but GH-treated patients would show little change.

Statistical analyses

T-tests or Mann–Whitney tests were conducted on the differences between the treatment-group change scores over the withdrawal period. Bonferroni corrections were not applied to reduce the chance of Type II errors in a study with small sample size, and as it was unlikely that a significant result that had been predicted would be obtained by chance. The required significance level was set at $< 0{\cdot}05$.

Results

The patient sample

Only 66 of the 144 adults with GHD being treated with GH at St Thomas' Hospital fulfilled the stringent inclusion criteria and were approached. However, 62% declined to participate, the majority because they were satisfied with their current treatment – not wishing to risk withdrawal, with expected return of undesirable symptoms of GHD. Twenty-two patients were finally recruited into the study, but one patient withdrew from psychological aspects of the study just prior to baseline, and another (placebo group) withdrew suddenly and prematurely from the study, about 2 weeks prior to end-point, owing to adverse symptoms. There was no significant difference in the number of men (23) and women (18) who declined to participate in the study, or in the age of those who refused and those who participated. The study sample was therefore representative of the patient pool in terms of age and sex.

Table 4 shows patient characteristics. At baseline, the placebo group had significantly higher BMI ($31{\cdot}3$ kg/m^2) compared with the GH-treatment group ($24{\cdot}7$ kg/m^2).

Biochemical changes

Three months after baseline the serum total IGF-I of placebo-treated patients fell from normal, age-related levels (mean $26{\cdot}6 \pm 3{\cdot}8$ nmol/l) to levels indicative of severe GHD ($11{\cdot}6 \pm 1{\cdot}9$ nmol/l) ($P < 0{\cdot}001$). Only a small, nonsignificant decrease was noted in GH-treated patients.

One patient in the GH-treatment group reported several adverse symptoms over the course of the withdrawal period. A large drop ($-20{\cdot}2$ nmol/l or 39·3%) in his IGF-I levels were noted over the study, well outside the mean change of $-1{\cdot}74 \pm 10{\cdot}68$ for the GH-treatment group and more than the mean change of $-15{\cdot}02 \pm 12{\cdot}38$ in IGF-I levels for the placebo group. However, his IGF-I level at end-point was in the normal range. After the study codes had been broken, the patient was asked if there were any reasons for this and he claimed that the injection pen may have malfunctioned. It is also possible that he did not fully adhere to the injection regimen, as he had a history of nonadherence. Analyses were therefore undertaken excluding the data of this anomalous case.

Questionnaire data

NHP. No significant between-group differences in change scores were found in Mann–Whitney tests.

Table 4 Characteristics of participants in the GH-withdrawal study

	Placebo-treated ($N = 12$)	GH-treated ($N = 9$)	P
Men	6	4	
Women	6	5	
Mean age at baseline (years) [range]	45·8 [25–68]	43·8 [25–66]	
Mean duration of GH treatment (months) [range]	61 [12–132]	60 [18–132]	
COGHD : AOGHD ratio	2 : 10	4 : 5	
BMI at baseline (kg/m^2)	$31{\cdot}3 \pm 8{\cdot}3$	$24{\cdot}7 \pm 2{\cdot}9$	$< 0{\cdot}05$
Isolated GHD	0	2	
Gonadal hormone deficiency	8	6	
Thyroid hormone deficiency	9	5	
Corticosteroid deficiency	10	5	
Antidiuretic hormone deficiency	1	1	
Acromegaly	0	1	
Cushing's disease	4	1	
Craniopharyngioma	1	1	
Chromophobe adenoma	4	1	
Macroprolactinoma	1	1	
Prolactinoma	1	1	
Traumatic hypopituitarism	1	0	

Table 5 GH-treatment group means over the GH-withdrawal period

	Placebo-treated		GH-treated		
	Baseline	End-point	Baseline	End-point	*P*
NHP					
Emotional Reactions	13·86 ± 24·46	14·42 ± 24·91	3·78 ± 7·43	0·00 ± 0·0	
Energy	24·95 ± 35·06	26·33 ± 37·89	7·90 ± 15·18	7·90 ± 15·18	
Pain	4·12 ± 10·85	5·17 ± 17·14	4·77 ± 9·44	5·29 ± 12·77	
Physical Mobility	5·42 ± 15·74	7·15 ± 23·73	2·75 ± 7·77	5·42 ± 11·60	
Sleep	17·43 ± 25·98	21·67 ± 29·84	5·16 ± 7·20	0·00 ± 0·0	
Social Isolation	13·13 ± 24·11	13·09 ± 26·10	9·68 ± 14·08	0·00 ± 0·0	
SF-36					
Bodily Pain	72·08 ± 20·57	71·55 ± 28·00	70·00 ± 14·05	81·50 ± 27·52	
General Health	63·50 ± 19·58	57·00 ± 23·19	62·75 ± 23·89	66·12 ± 25·16	0·013*
Mental Health	75·00 ± 15·92	69·45 ± 21·04	77·50 ± 20·05	84·50 ± 7·23	0·027*
Physical Functioning	78·52 ± 26·43	72·27 ± 30·20	83·13 ± 14·13	85·63 ± 15·45	
Role-Emotional	77·78 ± 35·77	63·33 ± 48·30	87·50 ± 35·36	91·67 ± 23·57	
Role-Physical	70·83 ± 39·65	72·73 ± 41·01	84·38 ± 18·60	90·63 ± 12·94	
Social Functioning	82·95 ± 25·78	77·27 ± 28·95	93·75 ± 9·45	90·63 ± 12·94	
Vitality	60·00 ± 22·76	51·82 ± 28·40	59·38 ± 15·22	60·00 ± 13·63	
Minimum *N*	11 (NHP)	11 (NHP)	8 (NHP)	6 (NHP)	
	11 (SF-36)	10 (SF-36)	8 (SF-36)	8 (SF-36)	

*Two-tailed. Values are means ± SD.
Subscale scores range 0–100, higher score indicating better health status on SF-36, but poorer health status on NHP.

SF-36. As expected, a significant between-group difference in change was found in SF-36 General Health [$t(17) = 2{\cdot}76$, $P = 0{\cdot}013$, two-tailed], with scores of placebo-treated patients dropping from 63·5 ± 19·6 at baseline to 57·0 ± 23·2 but scores of GH-treated patients increasing from 62·8 ± 23·9 to 66·1 ± 25·2, and in SF-36 Mental Health [$t(17) = 2{\cdot}41$, $P = 0{\cdot}027$, two-tailed], with scores of placebo-treated patients dropping from 75·0 ± 15·9 at baseline to 69·5 ± 21·0 but scores of GH-treated patients increasing from 77·5 ± 20·1 to 84·5 ± 7·2. Patients withdrawn from GH reported reduced general health and mental health at end-point. (Table 5).

Analysis of the single health transition item showed that, by end-point, four of the 12 placebo-treated patients (33·3%), but only one patient in the GH-treated group (the outlier whose IGF-I levels dropped considerably), considered their health worse than 1 year before.

Discussion

The NHP has been one of the most frequently used questionnaires in research into adult GHD, yet it is known to have certain disadvantages. It is designed for more severe illness and therefore has highly skewed distributions in relatively healthy populations, with reduced sensitivity to between-group differences, reduced sensitivity to change (Bowling, 1991) and inability to detect low levels of morbidity. Such criticisms were confirmed in the present studies. For all individual NHP items the majority of respondents in the cross-sectional survey answered 'No', that they perceived no health problem in that area; indeed almost a quarter of respondents reported no health problem in any of the 38 items. The wider range of SF-36 response options means that more information is available about different degrees of impairment, thus the SF-36 had greater discriminatory power. Responses to individual items of the SF-36 showed that most respondents in the cross-sectional survey were not limited at all in moderate activities and only felt limited in vigorous activities such as climbing several flights of stairs. However, patients had significantly worse health status for most SF-36 subscales than the UK general population.

Both NHP and SF-36 have good acceptability to adults with GHD (as indicated by high completion rates), and good internal consistency reliability. However, reliability and factor analyses of the whole 38-item NHP scale confirmed the inadvisability of calculating an NHP Total Score, as occasionally reported in the literature.

In the cross-sectional study the SF-36 showed more sensitivity to expected sex differences than the NHP. Women had significantly worse health status than men on all SF-36 subscales but Mental Health, but on only one NHP subscale (Physical Mobility). In the general population, women's health status is worse

than men's (Hunt & McKenna, 1989; Brazier *et al*., 1992), and in adult GHD women have also tended to exhibit worse health status (NHP) than men (Burman *et al*., 1995). No significant differences in health status were found between patients with COGHD and those with AOGHD although these are different clinical entities (Attanasio *et al*., 1997): CO patients are more underdeveloped physically and AO patients have greater lipid abnormalities. There have been few studies comparing health status in the two groups. Although one study found no difference in health-related QoL (Abs *et al*., 1999), another study found significantly reduced physical mobility and energy in AOGHD than COGHD (Attanasio *et al*., 1997).

Although the SF-36 Mental Health and Vitality subscales showed correlations indicative of improving health status for GH-treated patients with longer periods of GH treatment, neither questionnaire found significant GH treatment-group differences. It could be that neither questionnaire was sensitive to treatment-group differences, or that there were no real differences in perceived health status between the two groups. It may be that people prescribed GH treatment present with more serious symptoms of mental and physical ill-health but GH treatment improves their health status to a level little different from people with GHD who are untreated, perhaps because they are managing well without it. The results seem to indicate that an appropriate subgroup of patients was prescribed GH treatment – perhaps the result of a combination of clinical acumen and patient pressure from those adversely affected by GHD. In the study as a whole, however, significant findings were limited to a few selected comparisons (e.g. sex differences and differences from the healthy general population). Insufficient data were available to conduct further analyses of between-group differences.

Recruitment difficulties for the GH-withdrawal study indicated that the majority of patients did not want to risk withdrawal from GH treatment, and the anticipated negative effect that it would have on their lives. The small pool of patients suitable for inclusion in the GH-withdrawal study, combined with a high refusal rate, resulted in a small sample of 21 patients, and low power of analysis. One placebo-treated patient withdrew from the study just prior to end-point owing to adverse symptoms. The possible nonadherence to the study protocol of one GH-treated patient further reduced the data available for analysis. While it might have been more logical and less problematic to investigate the sensitivity of the questionnaires to change with GH replacement, for which there would be no recruitment problems, there was no opportunity for such a study at the time.

In the GH-withdrawal study, where health status of placebo-treated patients was expected to worsen, the NHP found no significant between-treatment-group differences. The SF-36, on the other hand, had two significant findings in the expected direction (of worsening health in the placebo group) in the General Health and Mental Health subscales. In a study such as the GH-withdrawal study, where power was low, the NHP was not sufficiently sensitive to change. It is possible that some of the negative findings for the benefits of GH treatment in the literature might have been the result of the lack of sensitivity of the questionnaire chosen to measure change, the NHP.

The relationship between hormonal abnormalities and psychological consequences is complex (Fava *et al*., 1993). These studies did not cover the effects of patients' other hormone deficiencies or endocrine disease (if any). Patients with diabetes were excluded from the studies, and only those patients who were being treated with appropriate adrenal, thyroid and gonadal hormone replacement therapy were included. However, there is evidence that some patients still exhibit psychological symptoms (particularly depression or anxiety) even if their hormone levels have been normalized with therapy (O'Malley *et al*., 2000; Sonino & Fava, 2001; Lovas *et al*., 2002). Full details of other hormone deficiencies were not recorded in the cross-sectional study, although they are available for the longitudinal study. It is not possible to say whether any psychological distress was due to another endocrine condition, but this should be borne in mind when interpreting results.

Conclusion

In conclusion, the present studies have shown that the SF-36 is a better measure than the NHP of health status of people with GH deficiency because of its greater discriminatory power, with the ability to detect lesser degrees of disability. It also has superior sensitivity to some subgroup differences and superior sensitivity to change compared with the NHP. The SF-36 is highly acceptable to respondents, and has very good internal consistency reliability. The SF-36 is recommended to measure the health status of adults with GH deficiency.

Questionnaire availability

The NHP is available from Dr Stephen McKenna, Galen Research, Enterprise House, Manchester Science Park, Lloyd Street North, Manchester M15 6SE, UK.

The SF-36 is administered by the Medical Outcomes Trust, 8 Park Plaza #503, Boston, MA 02116, USA. Licence agreement details are obtainable from www.qualitymetric.com.

Acknowledgements

C.V.M. was funded by a research studentship from Lilly Industries Ltd, which also provided partial funding for J.G. We acknowledge valuable assistance from Louise Breen, Research Nurse at the Endocrine Clinic of St Thomas' Hospital, and from John Valentine, statistician at Royal Holloway, University of London, and the essential contributions of participants in the studies.

References

Abs, R., Bengtsson, B.A., Hernberg-Stahl, E., Monson, J.P., Tauber, J.P., Wilton, P. & Wuster, C. (1999) GH replacement in 1034 growth hormone deficient hypopituitary adults: demographic and clinical characteristics, dosing and safety. *Clinical Endocrinology*, **50**, 703–713.

Anderson, R.M., Fitzgerald, J.T., Wisdom, K., Davis, W.K. & Hiss, R.G. (1997) A comparison of global versus disease-specific quality-of-life measures in patients with NIDDM. *Diabetes Care*, **20**, 299–305.

Attanasio, A.F., Lamberts, S.W.J., Matranga, A.M.C., Birkett, M.A., Bates, P.C., Valk, N.K., Hilsted, J., Bengtsson, B.A. & Strasburger, C.J. (1997) Adult growth hormone (GH)-deficient patients demonstrate heterogeneity between childhood onset and adult onset before and during human GH treatment. *Journal of Clinical Endocrinology and Metabolism*, **82**, 82–88.

Baum, H.B.A., Katznelson, L., Sherman, J.C., Biller, B.M.K., Hayden, D.L., Schoenfeld, D.A., Cannistraro, K.E. & Klibanski, A. (1998) Effects of physiological growth hormone (GH) therapy on cognition and quality of life in patients with adult-onset GH deficiency. *Journal of Clinical Endocrinology and Metabolism*, **83**, 3184–3189.

Benbow, S.J., Wallymahmed, M.E. & MacFarlane, I.A. (1998) Diabetic peripheral neuropathy and quality of life. *QJM: Monthly Journal of the Association of Physicians*, **91**, 733–737.

Bengtsson, B.A., Johannsson, G., Shalet, S.M., Simpson, H. & Sönksen, P.H. (2000) Treatment of growth hormone deficiency in adults. *Journal of Clinical Endocrinology and Metabolism*, **85**, 933–937.

Bowling, A. (1991) *Measuring Health*. Open University Press, Buckingham.

Bradley, C. (1994) The Well-being Questionnaire. In *Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice* (ed. C. Bradley), pp. 89–109. Harwood Academic Publishers, Chur, Switzerland.

Bradley, C. (1999) Achieving accessibility with quality: questionnaire measurement of condition-specific individualised quality of life. *Proceedings of the British Psychological Society*, **7**, 143.

Bradley, C. (2001) Importance of differentiating health status from quality of life. *Lancet*, **357**, 7–8.

Brazier, J.E., Harper, R., Jones, N.M.B., Ocathain, A., Thomas, K.J., Usherwood, T. & Westlake, L. (1992) Validating the SF-36 Health Survey Questionnaire – new outcome measure for primary care. *British Medical Journal*, **305**, 160–164.

Burman, P., Broman, J.E., Hetta, J., Wiklund, I., Erfurth, E.M., Hagg, E. & Karlsson, F.A. (1995) Quality-of-life in adults with growth-hormone (GH) deficiency – response to treatment with recombinant human GH in a placebo-controlled 21-month trial. *Journal of Clinical Endocrinology and Metabolism*, **80**, 3585–3590.

Burney, R.E., Jones, K.R., Christy, B. & Thompson, N.W. (1999) Health status improvement after surgical correction of primary hyperparathyroidism in patients with high and low preoperative calcium levels. *Surgery*, **125**, 608–614.

Carroll, P.V., Christ, E.R. & Growth Hormone Research Society Scientific Committee (1998) Growth hormone deficiency in adulthood and the effects of growth hormone replacement: a review. *Journal of Clinical Endocrinology and Metabolism*, **83**, 382–395.

Carroll, P.V., Littlewood, R., Weissberger, A.J., Bogalho, P., McGauley, G., Sönksen, P.H. & Russell-Jones, D.L. (1997) The effects of two doses of replacement growth hormone on the biochemical, body composition and psychological profiles of growth hormone-deficient adults. *European Journal of Endocrinology*, **137**, 146–153.

Costello, C.G. & Devins, G.M. (1998) Two stage screening for stressful life events and chronic difficulties. *Canadian Journal of Behavioral Science*, **20**, 85–92.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297–334.

Cuneo, R.C., Judd, S., Wallace, J.D., Perry-Keene, D., Burger, H., Lim-Tio, S., Strauss, B., Stockigt, J., Topliss, D., Alford, F., Hew, L., Bode, H., Conway, A., Handelsman, D., Dunn, S., Boyages, S., Cheung, N.W. & Hurley, D. (1998) The Australian multicenter trial of growth hormone (GH) treatment in GH-deficient adults. *Journal of Clinical Endocrinology and Metabolism*, **83**, 107–116.

Dupuy, H.J., The Psychological General Well-being Index (PGWB). (1984) In *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies* (eds. N.K. Wenger, M.E. Mattson, C.D. Furburg, J. Elinson) Le Jacq. Publishing Inc., New York.

Fava, G.A., Sonino, N. & Morphy, M.A. (1993) Psychosomatic view of endocrine disorders. *Psychotherapy and Psychosomatics.*, **59**, 20–33.

Garratt, A.M., Ruta, D.A., Abdalla, M.I., Buckingham, J.K. & Russell, I.T. (1993) The SF-36 Health Survey Questionnaire – an outcome measure suitable for routine use within the NHS. *British Medical Journal*, **306**, 1440–1444.

Green, S.B., Salkind, N.J. & Akey, T.M. (1997) *Using SPSS for Windows: Analyzing and Understanding Data*. Prentice Hall, New Jersey.

Hakkaart-van Roijen, L., Beckers, A., Stevenaert, A. & Rutten, F.F.H. (1998) The burden of illness of hypopituitary adults with growth hormone deficiency. *Pharmacoeconomics*, **14**, 395–403.

Hunt, S.M., McEwen, J. & McKenna, S.P. (1985) Measuring health status – a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners*, **35**, 185–188.

Hunt, S.M. & McKenna, S.P. (1989) The Nottingham Health Profile. In *European Guide to the Nottingham Health Profile*. The European Group for Quality of Life and Health Measurement. Montpellier, France.

Hunt, S.M. & McKenna, S.P. (1992) A British adaptation of the General Well-being Index: a new tool for clinical research. *British Journal of Medical Economics*, **2**, 49–60.

Hunt, S.M., McKenna, S.P. & Doward, L.C. (1993) Preliminary report on the development of a disease-specific instrument for assessing quality-of-life of adults with growth hormone deficiency. *Acta Endocrinologica*, **128**, 37–40.

Kenny, A.M., Bellantonio, S., Gruman, C.A., Acosta, R.D. & Prestwood, K.M. (2002) Effects of transdermal testosterone on cognitive function and health perception in older men with low bioavailable testosterone levels. *Journal of Gerontology Series A – Biological Sciences and Medical Sciences*, **57**, M321–M325.

Kline, P. (1993) *A Handbook of Test Construction*. Routledge, London.

Kline, P. (1994) *An Easy Guide to Factor Analysis*. Routledge, London.

Lovas, K., Loge, J.H. & Husebye, E.S. (2002) Subjective health status in Norwegian patients with Addison's disease. *Clinical Endocrinology*, **56**, 581–588.

Mardh, G., Lundin, K., Borg, G., Jonsson, B. & Lindeberg, A. (1994) Growth hormone replacement therapy in adult hypopituitary patients with growth hormone deficiency: combined data from 12 European placebo-controlled clinical trials. *Endocrinology and Metabolism*, **1** (Suppl. A), 43–49.

McGauley, G., Cuneo, R.C., Salomon, F. & Sönksen, P. (1990) Psychosocial well being before and after growth hormone treatment in adults with growth hormone deficiency. *Hormone Research*, **33** (Suppl. 4), 52–54.

McMillan, C.V. (2001) A psychometric evaluation of measures of quality of life and related health outcomes in adults with growth hormone deficiency. PhD Thesis, Royal Holloway, University of London, UK.

McMillan, C. & Bradley, C. (2000) A comparison of the psychometric properties of quality of life and related measures (GWBI, W-BQ12, NHP, SF-36 and HDQoL) for use in research into adult growth hormone deficiency (GHD). *Quality of Life Research*, **9**, 1324 (Abstract 1352).

McMillan, C.V., Gibney, J., Russell-Jones, D.L., Sönksen, P.H. & Bradley, C. (2001) Psychological effects of withdrawal of growth hormone (GH) therapy from GH-deficient adults. *Endocrine Abstracts*, **1**, 111.

O'Malley, B., Hickey, J. & Nevens, E. (2000) Thyroid dysfunction – weight problems and the psyche: the patients' perspective. *Journal of Human Nutrition and Dietetics*, **13**, 243–248.

Powrie, J., Weissberger, A. & Sönksen, P. (1995) Growth hormone replacement therapy for growth hormone-deficient adults. *Drugs*, **49**, 656–663.

Ryan, N. & Rosner, A. (2001) Quality of life and costs associated with micronized progesterone and medroxyprogesterone acetate in hormone replacement therapy for nonhysterectomized, postmenopausal women. *Clinical Therapeutics*, **23**, 1099–1115.

Smith, K.W., Avis, N.E. & Assmann, S.F. (1999) Distinguishing between quality of life and health status in quality of life research: a meta-analysis. *Quality of Life Research*, **8**, 447–459.

Sonino, N. & Fava, G.A. (2001) Psychiatric disorders associated with Cushing's syndrome. Epidemiology, pathophysiology and treatment. *CNS Drugs*, **15**, 361–373.

Sönksen, P., Cuneo, R.C., Salomon, F., McGauley, G., Wiles, C.M., Wilmshurst, P., Byrne, C., Hesp, R., Lowy, C. & Weissberger, A. (1991) Growth-hormone therapy in adults with growth-hormone deficiency. *Acta Paediatrica Scandinavica*, **379** (Suppl.), 139–146.

Tabachnik, B.G. & Fidell, L.S. (1983) *Using Multivariate Statistics*, 1st edn. Harper & Row, New York.

Todd, C. & Bradley, C. (1994) Evaluating the design and development of psychological scales. In *Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice* (ed. C. Bradley), pp. 15–42. Harwood Academic Publishers, Chur, Switzerland.

Wallymahmed, M.E., Baker, G.A., Humphris, G., Dewey, M. & MacFarlane, I.A. (1996) The development, reliability and validity of a disease specific quality of life model for adults with growth hormone deficiency. *Clinical Endocrinology*, **44**, 403–411.

Wallymahmed, M.E., Foy, P., Shaw, D., Hutcheon, R., Edwards, R.H.T. & MacFarlane, I.A. (1997) Quality of life, body composition and muscle strength in adult growth hormone deficiency: the influence of growth hormone replacement therapy for up to 3 years. *Clinical Endocrinology*, **47**, 439–446.

Ware, J.E. & Sherbourne, C.D. (1992) The MOS 36-Item Short-Form Health Survey (SF-36). 1. Conceptual framework and item selection. *Medical Care*, **30**, 473–483.

Whitehead, H.M., Boreham, C., McIlrath, E.M., Sheridan, B., Kennedy, L., Atkinson, A.B. & Hadden, D.R. (1992) Growth-hormone treatment of adults with growth-hormone deficiency – results of a 13-month placebo controlled cross-over study. *Clinical Endocrinology*, **36**, 45–52.

Wiren, L., Bengtsson, B.A. & Johannsson, G. (1998) Beneficial effects of long-term GH replacement therapy on quality of life in adults with GH deficiency. *Clinical Endocrinology*, **48**, 613–620.