Pergamon

European Journal of Cancer, Vol. 35, No. 2, pp. 256–261, 1999

Printed in Great Britain
0959-8049/99/$ - see front matter

PII: S0959-8049(98)00367-0

Original Paper

Applicability of the Health Utilities Index to a Population of Childhood Survivors of Central Nervous System Tumours in the U.K.

A.W. Glaser,[1] W. Furlong,[4] D.A. Walker,[1] K. Fielding,[2] K. Davies,[3] D.H. Feeny[2] and R.D. Barr[5]

[1]Department of Child Health; [2]Trent Institute for Health Service Research; [3]Department of Paediatric Physiotherapy, Queen's Medical Centre, University of Nottingham, Nottingham, U.K; [4]Centre for Health Economics and Policy Analysis, McMaster University; and [5]Childrens Hospital at Chedoke McMaster, Hamilton, Ontario, Canada

This paper describes the application of a multi-attribute, preference-linked health status and health-related quality of life measurement system—the Health Utilities Index (developed in Canada)—to a group of subjects in the U.K. Children who had survived tumours of the central nervous system ($n = 30$, age 6–16 years) formed the study group. Respondents (children, parents, physicians and physiotherapists) found the activity (completion of a 15-item questionnaire) to be acceptable and not burdensome (it was accomplished easily by all children ≥ 10 years of age). Instrumental reliability was established by acceptable intra- and interobserver agreement and construct validity was supported by strong similarities between the results obtained in this study and those reported from a similar group of children in Canada. The greatest burden of morbidity was reported for the attributes of emotion and cognition (each affected in > 50% of the children). Pain was surprisingly prevalent (affected in approximately one-third of children). The finding of a large number of unique health states emphasises the complex morbidity burden experienced by these children who self-reported poorer overall health (as reflected in utility scores) than did the proxy respondents. The information obtained from this study is readily interpretable and clinically useful. The results of this study also illustrate that extreme caution must be exercised in undertaking linguistic modifications to established instruments for, in this instance, these resulted in a loss of the ability to detect the most severe emotional morbidity and reduced the comparability of results between studies. With this provision, the Health Utilities Index is evidently applicable in the U.K. and the original version has been recommended for use in brain tumour studies by the U.K. CCSG (the U.K. Children's Cancer Study Group).

Key words: quality of life, children, brain tumours
Eur J Cancer, Vol. 35, No. 2, pp. 256–261, 1999

INTRODUCTION

ASSESSMENT OF health-related quality of life (HRQL) is an important outcome measure in clinical trials [1, 2]. Determination of the overall morbidity burden of survival provides information for selecting among treatment options, for economic analyses and for describing prognosis [3, 4]. The importance of HRQL assessment is recognised increasingly and many of the major sponsors of clinical trials insist that it be considered for inclusion in new protocols [5, 6].

Rigorous scientific determination of the validity, reliability, responsiveness and appropriateness of application of HRQL measures is mandatory if the benefits of measurement are to be obtained. Incorrect application may lead to inappropriate conclusions being made regarding therapeutic options and trial outcomes [7]. Despite the burgeoning number of

Correspondence to A.W. Glaser, Department of Oncology, Great Ormond Street Hospital for Children, Great Ormond Street, London WC1N 3JH, U.K.
Received 19 Jan. 1998; revised 10 Sep. 1998; accepted 19 Oct. 1998.

publications relating to HRQL, few clinical studies specify what is being measured [8]. This reflects in part the complexity of, and lack of consensus regarding, the definition of HRQL. The World Health Organization Quality of Life Group's definition of quality of life emphasises a broad conceptual approach [9]. A more recent definition, from an American Cancer Society Workshop on quality of life in childhood cancer, is as follows: 'Quality of life is multi–dimensional. It includes, but is not limited to, the social, physical and emotional functioning of the child and adolescent and, when indicated, his/her family. Measurement of quality of life must be from the perspective of the child, adolescent and family, and it must be sensitive to the changes that occur through development' [10].

For this study, our definition of HRQL was constrained to include only those factors intrinsic to patients—a 'within the skin' approach (the extent to which one's health status permits one to engage in the activities one chooses) [11]. Many established measures of HRQL in adults cannot be used with children due to the issues of growth (an increase in size), development (the acquisition of new skills) and the dependence on others [12]. Childhood illness often manifests itself as abnormal physical, emotional or intellectual development rather than as specific clinical symptoms or signs [13]. Therefore, a useful definition of a child's health is 'the ability to participate fully in developmentally appropriate activities, and requires physical, psychological and social energy' [11, 14].

The Health Utilities Index (HUI) is a family of generic multi-attribute health status and HRQL measurement systems developed at McMaster University in Canada [15]. They have been applied to a wide variety of childhood populations including extremely low birthweight infants, survivors of paediatric intensive care and children with cancer [16–18]. HUI health status assessment data are linked to multi–attribute utility functions to calculate preference scores of HRQL for individual health states [19].

The primary objective of this study was to assess the applicability of the HUI to the measurement of health status in survivors of childhood cancer in the U.K. For this purpose the HUI was applied to a cohort of childhood survivors of central nervous system tumours in a neuro-oncology unit in the U.K. similar to that studied by Barr and colleagues in Canada and described in a companion paper [20]. Applicability was assessed in terms of acceptability, respondent burden, reliability, responsiveness, validity, usefulness and interpretability.

The secondary objective was to determine the impact of wording modifications made to the established questionnaire. For this purpose, a mapping exercise was performed to convert Nottingham response options to those of the original HUI2 and HUI3 health status classification schemes. Two questions (numbers 7 and 14) of the original HUI 15-item questionnaire were modified by the Nottingham group. The Nottingham version, in which the two questions addressing emotional function are identical, was used in this and other studies [21–23].

PATIENTS AND METHODS

Subjects

All children who met the following inclusion criteria were considered eligible for this study: (1) had been treated at the Queen's Medical Centre, Nottingham for a primary tumour of the central nervous system; (2) were off-treatment for at least 1 year, (3) were over 5 years of age; and (4) attended the local neuro-oncology follow-up clinic.

Questionnaire

A modified version of the original 15-item HUI questionnaire was used. Each question has four to six response options. To determine the appropriate mapping between the revised options and the HUI attribute levels, 40 paediatric and adult physiotherapists, working at the Queen's Medical Centre in Nottingham, were sent a questionnaire asking them to draw lines joining equivalent descriptors of emotional level between the Nottingham response options and the original version of the two questions. The original version response options correspond exactly to the definitions of HUI2 and HUI3 attribute levels of emotional function. The Nottingham and original responses were presented adjacent to each other below the stem question, which is identical in both versions. The mapping results were returned anonymously in provided FREEPOST envelopes. The results of this exercise provided HUI2 and HUI3 levels of emotional function for responses to the Nottingham questionnaire. Responses to the other 13 questions were converted to health state vectors in the HUI2 and HUI3 formats by an established algorithm [24]. Single–attribute and global utility scores were calculated using the HUI2 utility function [19, 25].

Health status assessment survey

The questionnaire was completed by all parents, and (independently) by patients aged 10 years or more, in the clinic waiting room prior to being seen by their clinician. Patients, aged 5–9 years inclusive, completed the same questionnaire with the help of a play specialist. These subjects were felt to have inadequate comprehension skills to complete the questionnaire independently. The play specialist was known to all of the children and she acted as an emotionally uninvolved, yet informed facilitator. Following the routine clinic review, one of the physicians ($n = 3$) and the paediatric physiotherapist each completed identical questionnaires independently.

Analyses

The categorical paired responses to the two identical questions about emotion (questions 7 and 14, as modified by the Nottingham group, were worded identically) were analysed by per cent agreement and the kappa statistic, a chance corrected measure of association [26], to assess intrarater agreement. The corresponding single-attribute utility scores were analysed by intraclass correlation coefficients (ICC) [27] also to assess intrarater agreement. Interobserver agreement about functional levels for each HUI2 and HUI3 attribute was assessed by per cent agreement and the kappa statistic. Single-attribute and global health status utility scores are continuous measures and the interobserver agreement of these scores was assessed using ICC.

Assessments of face validity and of construct validity were undertaken. The results reported by Barr and colleagues [20] were compared with those from Nottingham using differences in frequencies of attribute levels, the chi–square test with Yates correction where appropriate, and Fisher's exact test.

The reader is referred to the companion paper by Barr and colleagues [20] for details of methods used in both studies,

such as the HUI scoring systems and the guidelines used for the interpretation of statistics on observer agreement.

RESULTS

30 subjects were identified as being eligible: 10 male and 20 female patients. The subjects' ages at diagnosis ranged from 1 to 13 years, mean 6.4 years. The subjects' ages at assessment ranged from 6 to 16 years, mean 10.5 years. The time from diagnosis to assessment ranged from 1 to 10 years, mean 4.1 years. The tumour types were astrocytoma/glioma ($n = 15$), primitive neuro–ectodermal tumour/medulloblastoma ($n = 5$), ependymoma ($n = 4$) and others ($n = 6$). All subjects had undergone at least one neurosurgical procedure, 11 (37%) received radiotherapy and 10 (33%) received chemotherapy. The characteristics of the patient sample are similar to those reported in the accompanying paper by Barr and colleagues [20].

All parents and 28 children (93%) completed questionnaires. 9 children were aged between 5 and 9 years inclusive. 7 of them completed questionnaires with the play specialist. Questionnaires were not administered to 2 children because the play specialist was unavailable to see them. Questionnaires were answered by the physiotherapist on all cases (100%), whilst on 27 (90%) patients questionnaires were filled in by the physicians. Four questionnaires were not completed correctly by parents, allowing 26 to be used for the generation of utility scores. The questionnaire from 1 child was not completed correctly, allowing 27 to be analysed for utility scores. Two physiotherapist and four physician questionnaires were not complete. Two types of problems with completion were identified. Those from the physicians and the physiotherapist were due to more than one page being turned over at once, resulting in unanswered questions. The others ($n = 5$) occurred when parent or child respondents assigned more than one response per question.

Morbidity

The results of the substudy to determine the appropriate HUI2 and HUI3 emotional attribute levels from modified questionnaire responses demonstrated that there was very good consensus among physiotherapists about the mapping required (data not shown). The physiotherapist was designated as the primary assessor as all study subjects were evaluated by this individual.

The proportions of children without morbidity or with morbidity in one or more attributes are listed in Table 1. Twenty–one per cent of survivors had no morbidity described by the HUI2 with the same proportion falling into this group according to the HUI3. These percentages are very similar to the results from the Hamilton cohort [20]. The severity of functional deficits is illustrated in Table 2 . As predicted by clinical experience, the attributes affected most frequently were emotion and cognition. A greater proportion ($P = 0.019$) of the Nottingham cohort was reported to have morbidity in the attribute of emotion than in a similar cohort reported by Barr and colleagues [20]. Differences in the frequency distributions of levels 1 and 2 for emotion may be due entirely to rewording of the questionnaires related to this attribute. The severity of cognitive morbidity is especially notable, although the proportion of the Nottingham cohort affected is smaller than in Hamilton, perhaps reflecting a smaller proportion exposed to radiotherapy in the Nottingham patients. The other statistically significant differences between the cohorts were in HUI2 self–care ($P = 0.014$) and HUI3 dexterity ($P = 0.027$). In both of these complementary studies pain was surprisingly prevalent.

A comparison of children who received radiotherapy with those who did not has not been performed as only 11 children were exposed to this therapeutic modality. A breakdown of cases by site of radiotherapy and status of disease provided cells which were either empty or contained very small numbers.

Summary statistics of HUI2 and HUI3 HRQL utility scores, derived from information on health status provided by the various types of assessors, are shown in Table 3. Mean HUI2 scores ranged from a minimum of 0.78 (based on child reports) to a maximum of 0.89 (based on physician reports). The corresponding range of HUI3 scores was 0.66 to 0.83.

Reliability

The results for analyses of intrarater agreement (test–retest reliability) for levels of emotional function are provided in Table 4. There were moderate to high levels of agreement, of both categorical (i.e. attribute levels) and interval scale variables (i.e. single-attribute utility scores), for all types of assessors. Inter-rater agreement for single-attribute utility scores is displayed in Table 5 and for global utility scores in Table 6 . In general, the patterns of agreement between types of raters for attribute levels were similar to those reported by Barr and colleagues [20], except for assessments of pain in which there was lower agreement between child and physician assessments than was obtained in Canada. ICCs indicate 'fair' or better agreement between types of raters, except for that between the physician and the children.

Table 1. *Proportion* (%) *of children by number of attributes affected**

HUI2									
Number of attributes†	0	1	2	3	4	5	6		
Proportion	21%	11%	21%	29%	18%	0%	0%		
Δ (H – N)	– 3%	+ 18%	+ 6%	– 19%	– 11%	+ 7%	+ 2%		
HUI3									
Number of attributes	0	1	2	3	4	5	6	7	8
Proportion‡	21%	18%	18%	29%	11%	0%	4%	0%	0%
Δ (H – N)	+ 1%	– 3%	+ 2%	– 2%	– 9%	+ 7%	+ 1%	+ 2%	0%

*Defined as the Health Utilities Index (HUI) level > 1 (other than normal). †Missing for one individual. ‡Sum equals 101% due to rounding. Δ (H – N) = Difference in results between Hamilton and Nottingham.

Table 2. Frequency distributions of attribute levels. Proportions (%) of children

HUI2						
Level	Sensation (%)	Mobility (%)	Emotion* (%)	Cognition (%)	Self-care (%)	Pain (%)
1	76	79	43	48	69	66
Δ (H – N)	– 15	1	30	– 14	16	7
2	10	10	54	48	17	31
Δ (H – N)	7	0	– 30	8	– 10	– 7
3	14	10	3	3	14	3
Δ (H – N)	3	– 8	– 1	4	– 14	– 1
4	0	0	0	0	0	0
Δ (H – N)	0	0	5	2	7	0
5	NA	0	NA	NA	NA	0
Δ (H – N)		0				0

HUI3								
Level	Vision (%)	Hearing (%)	Speech (%)	Ambulation (%)	Dexterity (%)	Emotion (%)	Cognition (%)	Pain (%)
1	76	97	90	79	86	43	48	72
Δ (H – N)	– 8	1	– 5	1	– 20	30	– 14	– 4
2	17	0	3	10	3	54	10	17
Δ (H – N)	0	0	9	0	19	– 30	0	10
3	0	0	7	7	0	3	17	10
Δ (H – N)	2	0	– 7	– 7	7	– 1	– 7	– 5
4	0	0	0	3	7	0	21	0
Δ (H – N)	2	2	0	– 1	– 5	0	16	0
5	7	3	0	0	3	NA	3	0
Δ (H – N)	0	– 3	2	2	– 3		4	0
6	0	0	NA	0	0	NA	0	NA
Δ (H – N)	2	0		5	2		2	

*Missing for one individual. NA not applicable. Δ (H – N) = Difference in results between Hamilton and Nottingham.

DISCUSSION

Quantification of the overall morbidity burden of survival from central nervous system tumours in childhood has been performed using a multi-attribute utility approach to health status measurement. The applicability of a North American instrument to a U.K. population has been examined.

High completion rates demonstrate the acceptability and low respondent burden of the questionnaire to respondents in the U.K. Errors in completion suggest that clarification of instructions may be beneficial, questionnaires should be checked immediately upon return and errors corrected before respondents are given leave. No respondents objected to the questionnaire and prompt completion did not delay the routine out-patient clinic schedule. Good test–retest reliability has been reported previously for the 15-item questionnaire in the U.K. [22] and was confirmed in the present study. Inter-rater agreement is shown now to be acceptable for both single-attribute measures and global utility scores. These findings are in accord with other observations on the sound psychometric properties of the HUI [15]. Further evidence in support of the face validity and construct validity of this system comes from the results of the present study.

The use of a multi-attribute health status assessment system has confirmed that survivors of central nervous system tumours experience difficulties in multiple domains. The findings that 50% of patients had cognitive difficulties, at

Table 3. Utility scores for global health status assessments by type of rater

	n	Mean	S.D.	Median	Minimum	Maximum
HU12						
Physiotherapist	28	0.85	0.13	0.88	0.56	1.00
Parents	26	0.82	0.18	0.86	0.25	1.00
Physicians	23	0.89	0.11	0.90	0.65	1.00
Children	25	0.78	0.18	0.80	0.36	1.00
HU13						
Physiotherapist	28	0.76	0.24	0.80	0.01	1.00
Parents	27	0.72	0.29	0.80	0.24	1.00
Physicians	24	0.83	0.17	0.86	0.57	1.00
Children	26	0.66	0.28	0.68	0.19	1.00

S.D., standard deviation.

Table 4. Intrarater agreement (test–retest reliability) of the Health Utilities Index (HUI) levels of emotion

	Attribute level data			Utility score data	
	No. of paired respondents	Per cent agreement	Kappa statistic*	Mean difference	ICC
Physiotherapist	29	76	0.58	0.024	0.73
Parents	30	90	0.84	0.001	0.87
Physicians	26	89	0.79	0.019	0.88
Children	28	68	0.54	0.023	0.62

*Cichetti and quadratic weights were applied and provided similar results. As in the accompanying paper of Barr and colleagues [20], only the quadratic weighted results are displayed. ICC intraclass correlation coefficient (value > 0.5 is indicative of a high level of agreement).

Table 5. Measures of agreement in assessments of morbidity on HUI2 single attributes by pairs of raters

Rater pair	*n*		Attribute					
			Sensation	Mobility	Emotion	Cognition	Self-care	Pain
Physio versus parents	25	PA	85	76	64	83	75	76
		K	0.62	0.44	NS	0.66	0.47	0.52
Physio versus children	25	PA	67	73	69	85	78	58
		K	0.32	NS	0.37	0.70	0.43	NS
Physio versus physicians	21	PA	88	84	44	62	88	62
		K	0.68	0.61	NS	NS	0.69	NS
Parents versus physicians	20	PA	87	92	69	65	79	65
		K	0.68	0.84	NS	NS	0.56	0.34
Children versus physicians	19	PA	70	90	50	46	91	52
		K	0.38	0.77	NS	NS	0.78	NS
Children versus parents	24	PA	77	88	67	67	77	81
		K	0.54	0.72	0.37	NS	0.47	0.62

n number of paired assessments; Physio, physiotherapist; K, Kappa statistic (for interpretation, see [20]); PA per cent agreement; NS not significantly different from 0.00.

least an equivalent proportion experienced emotional problems and over 20% had impaired mobility are not unexpected. These figures are in keeping with detailed single dimension morbidity studies of survivors of childhood central nervous system tumours [28, 29]. Additionally, use of the HUI identified that approximately 30% of children were perceived to experience pain which is not a commonly reported form of morbidity in survivors of central nervous system tumours [30]. These observations address the usefulness of this generic and comprehensive approach to health status measurement.

As to interpretability, comparison of the present findings can be made with the results of measurement of health status and HRQL by the HUI in the general population [31], in a retrospective study of survivors of extremely low birthweight (mean HUI2 score = 0.82) and a control group (mean HUI2 score = 0.95) at age 8 years [16], and in children who experienced head injuries [32]. It is interesting to note that the HUI2 mean score reported for extremely low birthweight survivors is the same as the mean score of patients in this study based on parental assessments. The results obtained in this study reveal marked similarities with the morbidity identified in a similar population of children in Canada (Table 7). The burden of cognitive impairment is particularly striking and supports a wealth of information from neuropsychological studies in other children [33].

Greater emotional morbidity has been reported in this study, compared with that reported by Barr and colleagues [20] for a similar group of subjects in Canada. However, the differences are confined entirely to frequency distributions of levels 1 and 2, and this effect is considered most likely due to the linguistic modification, rather than an actual difference in emotional status.

The results of the mapping exercise are not recommended for use in other studies in which the Nottingham-modified questionnaire has been used, because the exercise was completed by a sample of physiotherapists who were not well acquainted with the conceptual foundation of the HUI and may not generalise to other groups of assessors. Furthermore, the results of the mapping exercise demonstrate that level 5 in the emotional attributes in both HUI2 and HUI3 is not available if these linguistic modifications are used. The response option describing the lowest level of emotion was not selected in this study nor in our previous work on survivors of childhood central nervous system tumours [23]. Nevertheless, this level of emotional morbidity has been identified in other studies using the HUI, such as those following extremely low birthweight neonates through childhood into adolescence [16, 34]. This study demonstrates that extreme caution is required even if minor changes in the wording of questionnaires are made as the weighting and interpretation of items may be altered significantly. Such modifications are thus inadvisable.

Table 6. Inter-rater agreement of HUI2 (utility) scores for global health-related quality of life

Rater pair	*n*	ICC
Physio versus parents	25	0.75
Physio versus children	25	0.40
Physio versus physicians	23	0.66
Parents versus physicians	20	0.53
Children versus physicians	19	0.15
Children versus parents	24	0.57

n, number of paired assessments; ICC, Intraclass correlation coefficient (ICC > 0.5 is indicative of a strong correlation (high level of agreement)); Physio, physiotherapist.

Table 7. Similarities in health status between patients in Canada and the U.K.

	Hamilton [20]	Nottingham (current study)
Prevalence of no morbidity		
HUI2	18%	21%
HUI3	22%	21%
No. of unique health states		
HUI2	22/41 (0.54/patient)	19/30 (0.63/patient)
HUI3	29/41 (0.71/patient)	21/30 (0.70/patient)
Prevalence of cognitive morbidity		
HUI2	66%	52%
HUI3	66%	52%
Prevalence of pain		
HUI2	27%	34%
HUI3	32%	28%

A paediatric physiotherapist was designated as the primary assessor for this report. The point estimate of the ICC for agreement between the responses of the physiotherapist and the parents is identical to that of the nurse and the parents reported by Barr and colleagues [20]. The physiotherapist's scores had a similar level of agreement with those of the physicians compared with those reported for the nurse and physicians in the study by Barr and colleagues [20]. There is a hint that the nurse agreed more with the information provided by the children than did the physiotherapist with the children in this study, but this difference was not statistically significant. The low scores obtained from reports provided by the children themselves, in the U.K. study, may explain this finding and the lower correlation between children's scores and those derived from the reports of the physicians.

Agreement between types of respondents in this paper and the accompanying paper by Barr and colleagues [20] is, typically, only fair to moderate. We are not surprised by this result and we note that there is no generally accepted gold-standard perspective for assessing health status and HRQL. For the purpose of these studies we identified 'primary' assessors, but this should not be interpreted as meaning that we endorse the particular viewpoints of physiotherapists and nurses more than the viewpoints of others. In general, we recommend that assessments provided by all types of assessors be respected and used appropriately.

This pilot study has demonstrated the applicability of the HUI to a population of childhood survivors of central nervous system tumours in the U.K., and that the wording of the original HUI questionnaire should be retained. These results, in conjunction with the growing literature regarding the measurement properties of the HUI system, have led to the recommendation that the HUI be included as part of the clinical follow-up in all United Kingdom Children's Cancer Study Group central nervous system trials [22, 23].

1. Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *N Engl J Med* 1996, **334**, 835–840.
2. American Society of Clinical Oncology. Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. *J Clin Oncol* 1996, **14**, 671–679.
3. Fallowfield L. Quality of quality of life data. *Lancet* 1996, **348**, 421–422.
4. Barr RE, Furlong W, Henwood J, *et al*. Economic evaluation of allogeneic bone marrow transplantation: a rudimentary model to generate estimates for the timely formulation of clinical policy. *J Clin Oncol* 1996, **14**, 1413–1420.
5. Sadura A, Pater J, Osoba D, *et al*. Quality of life assessment: patient compliance with questionnaire completion. *JNCI* 1992, **84**, 1023–1026.
6. Medical Research Council. The Assessment of MRC Trials 1996/1997. London, Medical Research Council, 1996.
7. Glaser A, Walker D. Outcome assessments and air ambulance services. *Lancet* 1996, **347**, 1843.
8. Gill TM, Feinstein AR. A critical appraisal of the quality of life measurements. *JAMA* 1994, **272**, 619–626.
9. WHOQOL Group. *Measuring Quality of Life: The Development of the World Health Organization Quality of Life Instrument (WHOQOL)*. Geneva, WHO, 1993.
10. Bradlyn AS, Ritchey AK, Harris CV, *et al*. Quality of life research in pediatric oncology. Research methods and barriers. *Cancer* 1996, **78**, 1333–1339.
11. Feeny D, Furlong W, Boyle M, *et al*. Multi-attribute health status classification systems: Health Utilities Index. *Pharmaco Economics* 1995, **7**, 490–502.
12. Jenney M, Kane RL, Lurie N. Developing a measure of health outcomes in survivors of childhood cancer: a review of the issues. *Med Pediatr Oncol* 1995, **24**, 145–153.
13. Starfield B, Bergner M, Ensminger M, *et al*. Adolescent health status measurement: development of the child health and illness profile. *Pediatrics* 1993, **91**, 430–435.
14. Pantel RH, Lewis CC. Measuring the impact of medical care on children. *J Chronic Dis* 1987, **40**(Suppl. 1), 99s–108s.
15. Feeny DH, Torrance GW, Furlong WJ. Health Utilities Index. In Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd edn. Philadelphia, Lippincott–Raven Press, 1996, 239–252.
16. Saigal S, Feeny D, Furlong W, *et al*. Comparison of the health–related quality of life of extremely low birthweight children and a reference group of children at age eight years. *J Pediatr* 1994, **125**, 418–425.
17. Gemke RJB, Bonsel JB, van Vught AJ. Long-term survival and state of health after paediatric intensive care. *Arch Dis Child* 1995, **73**, 196–201.
18. Barr RD, Feeny D, Furlong W, *et al*. A preference–based approach to health–related quality of life for children with cancer. *Int J Pediatr Hematol Oncol* 1995, **2**, 305–315.
19. Torrance GW, Furlong W, Feeny D. Multi-attribute prefence functions. Health Utilities Index. *Pharmaco Economics* 1995, **7**, 503–520.
20. Barr RD, Simpson T, Whitton AC, *et al*. Health–related quality of life in survivors of tumours of the central nervous system in childhood—a preference–based approach to measurement in a cross–sectional study. *Eur J Cancer* 1999, **35**, 248–255.
21. Billson A, Walker DA. Assessment of health status in survivors of cancer. *Arch Dis Child* 1994, **70**, 200–204.
22. Glaser AW, Davies K, Walker D, *et al*. Influence of proxy respondents and mode of administration on health status assessment following central nervous system tumours in childhood. *Qual Life Res* 1997, **6**, 43–53.
23. Glaser AW, Rashid NFNA, Chin Lynn U, *et al*. School behaviour and health status after central nervous system tumours. *Br J Cancer* 1997, **76**, 643–650.
24. Furlong W, Torrance GW, Feeny D. Health Utilities Index: algorithm for determining Mark II/Mark III health status classification levels, health states and health state utility scores from 1992-10-20 self–administered health status questionnaire. McMaster University, 14 March, 1996.
25. Torrance GW, Feeny DH, Furlong WJ, *et al*. Multi–attribute preference functions for a comprehensive health status classification system: Health Utilities Index Mark 2. *Med Care* 1996, **34**, 702–722.
26. Cohen J. A co-efficient of agreement for nominal scales. *Edu Psychol Meas* 1960, **20**, 37–46.
27. Shrout P, Fleiss JL. Intra-class correlations: uses in assessing rater reliability. *Psychol Bull* 1979, **86**, 420–428.
28. Mulhern RK, Heideman RL, Khatib ZA, *et al*. Quality of survival among children treated for brain stem glioma. *Pediatr Neurosurg* 1994, **20**, 226–232.
29. Lannering B, Marky I, Lundberg A, *et al*. Long–term sequelae after pediatric brain tumours: their effect on disability and quality of life. *Med Pediatr Oncol* 1990, **18**, 304–310.
30. Whitton AC, Rhydderch H, Furlong W, *et al*. Self–reported comprehensive health status of adult brain tumour patients using the Health Utilities Index. *Cancer* 1997, **80**, 258–265.
31. Barr RD, Pai MKR, Weitzman S, *et al*. A multi–attribute approach to health status measurement and clinical management—illustrated by an application to brain tumours in childhood. *Int J Oncol* 1994, **4**, 639–648.
32. Glaser A, Buxton N, Fielding K, *et al*. Prediction of outcome and health status assessment following major head injuries in childhood. 7th World Congress of Intensive and Critical Care Medicine. Ottawa, July 1997 (abstract).
33. Mulhern RK. Neuropsychological late effects. In Bearison DJ, Mulhern RK, eds. *Pediatric Psycho–oncology. Psychological Perspectives on Children with Cancer*, New York, Oxford University Press, 1994, 99–121.
34. Saigal S, Feeny D, Rosenbaum P, *et al*. Self–perceived health status and health–related quality of life of extremely low birthweight infants at adolescence. *JAMA* 1996, **276**, 453–459.

Acknowledgements—Grants were received from the University of Nottingham Medical School Trust Funds, Nottingham Brain Tumour Research Fund and the Rank Foundation (A.W.G.).