

# Deriving reference values and utilities for the QoL-AGHDA in adult GHD

J. J. V. Busschbach · B. H. R. Wolffenbuttel ·  
L. Annemans · W. J. Meering ·  
M. Kołtowska-Hägström

Received: 7 May 2009 / Accepted: 25 March 2010 / Published online: 17 April 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

## Abstract

**Background** Quality of life (QoL) measures are important in growth hormone (GH) deficiency (GHD) in adults. Ideally, for use in health economics, QoL should be expressed in utilities. The aim of this study was to obtain reference values and utilities for QoL of GHD in adults in Belgium and the Netherlands.

**Methods** The study was conducted in three stages: (1) The Quality of Life-Assessment for Growth Hormone Deficiency in Adults (QoL-AGHDA) and the EQ-5D were

administered in a representative sample of 6,875 individuals from the Belgian and 1,400 individuals from the general Dutch population. The EQ-5D<sub>index</sub> can be used to estimate utilities. Using a regression, utilities were predicted from the QoL-AGHDA. (2) QoL-AGHDA scores were obtained from 299 Belgian and 234 Dutch adult patients with GHD and no GH replacement. These scores were converted to utilities and compared the burden of disease with other patient groups. (3) To test the criterion validity, the 'standard' EQ-5D<sub>index</sub> was used in a subsample of 64 Dutch GHD patients and compared with the predicted utilities.

**Results** We obtained data from 1,026 Belgian (response rate = 15%) and 1,038 Dutch respondents (response rate = 74%). The Belgian mean QoL-AGHDA value was 6.95 (90% range = 14.00), and the Dutch mean was 5.48 (range = 13.00). The  $R^2$  of the regression model to predict the EQ-5D<sub>index</sub> was 0.360 (Belgium) and 0.482 (the Netherlands). We demonstrated a considerable burden of disease in GHD patients, comparable to patients with hypertension or with type II diabetes. The criterion validity was 0.407 (intraclass correlation, ICC).

**Conclusions** Interventions in GHD can now be evaluated more validly in Belgium and the Netherlands.

**Keywords** Growth hormone/\*deficiency · Quality of life · Quality-adjusted life years · Questionnaires · Reference values

**JEL Classification** i · i1 · i19

## Background

The availability of general population values for Quality of Life (QoL) measures is important for clinical assessment of

---

J. J. V. Busschbach (✉)  
Medical Psychology and Psychotherapy, Erasmus MC,  
P.O. Box 2040, 3000 CA Rotterdam, The Netherlands  
e-mail: J.vanbusschbach@erasmusmc.nl

B. H. R. Wolffenbuttel  
Department of Endocrinology, University Medical Centre  
Groningen, University of Groningen, HPC AA31,  
P.O. Box 30001, 9700 RB Groningen, The Netherlands  
e-mail: bwo@int.umcg.nl

L. Annemans  
Department of Public Health, Faculty of Medicine and Health  
Sciences, Ghent University, De Pintelaan 185 Bl.A-1,  
9000 Ghent, Belgium  
e-mail: Lieven.Annemans@UGent.be

L. Annemans  
Brussels University (VUB), Brussels, Belgium

W. J. Meering  
Pfizer BV, Rivium Westlaan 142, 2909 LD Capelle a/d IJssel,  
The Netherlands  
e-mail: WillemJan.Meering@pfizer.com

M. Kołtowska-Hägström  
KIMS Medical Outcomes, Pfizer Endocrine Care,  
Vetenskapsvägen 10, SE-191 90 Sollentuna, Sweden  
e-mail: maria.koltowska-haggstrom@pfizer.com

patients and may facilitate health economic evaluations. Ideally, for use in health economics QoL should be expressed in ‘utilities’, but utilities can often not be produced by applying conventional clinical QoL measures. The Quality of Life-Assessment for Growth Hormone Deficiency in Adults (QoL-AGHDA) is one of the most used conventional clinical QoL measures in adult growth hormone deficiency (GHD). The aim of this study was to obtain general population values for the QoL-AGHDA, to develop a model to obtain utility values from the QoL-AGHDA and to further explore the burden of disease of GHD in adults in Belgium and the Netherlands.

Adults with longstanding, untreated GHD have multiple somatic impairments, including altered blood biochemistry, metabolism, body composition, and muscular and aerobic performance. They experience functional limitations, diminished productivity, social isolation, excessive fatigue and poorer QoL [1]. Given the impact of GHD on QoL, the assessment of QoL is crucial in the appraisal of the medical need and the effects of interventions. The aim of this study was to enhance the assessment of QoL in GHD in three ways: (1) First, we collected Belgian and Dutch reference values for the QoL-AGHDA based on samples from both the general population and patient samples. The QoL-AGHDA is not legitimate for health economic evaluation; therefore, it is necessary to make relevant adaptations. To support health policy evaluations of intervention in GHD, we made the QoL-AGHDA suitable for health economic evaluations in the Netherlands and Belgium. (2) Second, we compared the burden of disease of patients with GHD with other patient groups. (3) We tested the validity of the QoL-AGHDA outcomes for use in economic evaluations. Because the health economic aspects of this study are less known than the collection of reference values in patients and the general population, we have describe these aspects in more detail below.

### Adapting the QoL-AGHDA for QALY analysis

Health economic evaluations are becoming increasingly important in health policy-making in the Netherlands and Belgium [2–4]. For a valid health economic evaluation of GH substitution, convincing estimates of QoL are crucial but undeveloped [5, 6]. In health economics QoL outcomes are preferably measured in quality-adjusted life years (QALYs). QALYs are estimated by multiplying the number of life years with a quality of life value of ‘utility’ that has the value 1.00 for ‘perfect health’ and 0.00 for ‘a health state comparable to the value of death’. For instance, if the utility for sitting in a wheelchair has a value of 0.6, than 2 life years in a wheelchair stands for 1.2 QALY. QALYs are particular useful to provide a common metric of burden of

disease across the entire spectrum of diseases. For instance, with QALYs it becomes possible to compare the additional effects of a wheelchair on QoL with a life-saving operation. The most complex aspect of QALYs is the estimation of the utility. Fortunately, there are special QoL questionnaires like the EQ-5D that can provide this utility on a routine basis [7]. Most questionnaires like the EQ-5D are generic questionnaires; they can be used in a wide variety of diseases. A drawback of this generic feature is that these one-size-fits-all questionnaires might miss disease-specific symptoms that are relevant for QoL of these patients. A recent development is to validate disease-specific questionnaires in such a way that these questionnaires can also provide a utility that can be used in QALY analyses [8]. Such validations can be useful when the generic QoL questionnaires are found to be insufficiently sensitive to pick up relevant changes in QoL or when generic instruments have not been included in an investigation. Both motivations apply in GHD research [6]. For this reason, researchers in the UK and Sweden have linked the QoL-AGHDA and the EQ-5D in a way that QoL-AGHDA data collected in GHD patients can be used in QALY-type calculations [9, 10]. To achieve this, in both the UK and Sweden, investigators used a regression of the EQ-5D<sub>index</sub> on the QoL-AGHDA scores in samples of the general population. After such regression in the general population, the investigators could estimate the EQ-5D indexes from patient QoL-AGHDA scores. This method assumes that the relationship between the QoL-AGHDA and the EQ-5D in the general population is a valid predictor of the same relationship in patients. This crucial assumption can be tested if the EQ-5D and the QoL-AGHDA are both administered in a subsample of patients. Testing this assumption forms part of this paper.

### Comparing the burden of GHD with other diseases

Because the scores on the QoL-AGHDA only relate to disease-specific factors, it is difficult to relate the QoL-AGHDA scores to burden of disease in other patient groups. When the burden of GHD is measured using the EQ-5D or when the score of the QoL-AGHDA is transformed into an EQ-5D<sub>index</sub>, such comparison becomes possible. The possibility of describing the burden of GHD relative to other diseases is important, as ‘burden’ is recognised as an important factor in reimbursement decisions [11]. In the Netherlands, for example, this recognition has recently resulted in a formal recommendation to incorporate burden of disease in reimbursement decisions [12]. In this article we estimated the burden of GHD in generic terms and compared it with burden in other diseases.

## Materials and methods

### Questionnaires

#### *QoL-AGHDA*

One of the most frequently used QoL assessment instruments in adult GHD is the QoL-AGHDA. The QoL-AGHDA is a disease-specific instrument based on the concept that “QoL is the degree to which human needs are satisfied” [13]. The QoL-AGHDA was developed following in-depth interviews in adult patients with GHD and consists of 25 items with yes/no answers, acknowledging or denying GHD-related problems [14]. Examples of questionings are: “I feel a strong need to sleep during the day” and “I often feel lonely even when I am with other people”. All items are listed in Table 2. The most used scoring mode is an overall sum score, without any references to a possible dimensional structure. In line with that approach, Rasch models are often employed, and much attention is given to the unidimensionality of the sum score, which represents the ‘need’-driven perspective on QoL [15]. A high QoL-AGHDA score denotes that ‘less needs are satisfied’, and thus a lower QoL, with a score of 25 as maximum. The high number of items, all loading on an unidimensional score, contributes to a high reliability (test-retest, Spearman rank  $\geq 0.86$ ) and a high level of internal consistency (Cronbach’s  $\alpha \geq 0.88$ ) [14]. Mean values in the normal population range between 4 and 7 in various international studies [16]. A score of 11 or more on the QoL-AGHDA is one of the UK National Institute for Health and Clinical Excellence’s (NICE) requirements for GH replacement therapy [17]. The total sum score decreases when patient receive GH replacement therapy and approaches the mean value of the normal population [16].

The QoL-AGHDA is a typical disease-specific instrument, as it only targets one group of patients. This is one of the features that distinguish it from generic instruments like the EQ-5D. Because the QoL-AGHDA is tailored for GHD-related QoL, one can safely assume that its sensitivity is higher than a generic instrument. A downside of the specificity is that scores of the QoL-AGHDA are difficult to compare with scores from other QoL instruments. Moreover, unlike the EuroQol, the QoL-AGHDA does not provide a utility score necessary for health economic analysis. If one can transform QoL-AGHDA scores into EuroQol scores, then outcomes of the QoL-AGHDA would be comparable with other (generic) QoL instruments, and utility for the use in health economic analyses would be available. In the present research effort, we want to establish just that.

#### *EQ-5D*

The EuroQol EQ-5D is a five-item generic QoL questionnaire specially designed for health economic evaluations that involves QoL estimates. The five items cover the five dimensions of QoL: mobility, self care, usual activities, pain/discomfort and anxiety/depression. The answers on the five items are summarized into an index on which 1.00 represents ‘full health’ and 0.00 the value of ‘being dead’ [7]. The transformation of the five items into an index was done based on formal national standards. In Belgium this standard is based on a ‘visual analogue scale’ valuation study, whereas the Dutch standard is based on a ‘time trade-off’ [18, 19].

### Patient samples

The study involves three separate samples:

- (1) The first cohort was elicited from the general population of Belgium and the Netherlands. These subjects filled in both the QoL-AGHDA and EQ-5D. The QoL-AGHDA data from these two samples were used for deriving national population reference values. The regression of the EQ-5D on the QoL-AGHDA was used to transform QoL-AGHDA scores into utilities that could be used in health economic analyses.
- (2) The second cohort comprised Belgian and Dutch patients with GHD who contributed QoL data to KIMS (Pfizer International Metabolic Database) [20]. These patients filled in only the QoL-AGHDA prior to growth hormone treatment, providing scores for the Belgian and Dutch patient populations.
- (3) The last cohort was a subsample of 64 Dutch patients with GHD who filled in both the QoL-AGHDA and EQ-5D. This sample was used to test the criterion validity of the regression analysis of the EQ-5D<sub>index</sub> score on the QoL-AGHDA made in the Dutch general population.

#### *Population reference QoL-AGHDA and QALY values*

Members of the Belgian population were sampled in January 2007 by InSites Consulting, a professional institute for public opinion and marketing research. Out of the total panel of ~500,000 individuals, 6,875 individuals aged over 18 years were selected and invited to take part in the online survey. The invited sample was representative of the national distribution of age, sex and language (region). These socioeconomic variables are the usual variables to test whether a sample is representative of the population.

Apart from these variables, it is more important that the sample should be representative of the distribution of the variable of interest, in this case 'health'. We therefore asked the respondents to fill in the 'general health question', a question with a five-level Likert scale, which is often asked as part of national statistical monitors. In Belgium this monitor was undertaken by the Belgian National Institute of Statistics (<http://www.statbel.fgov.be>). In this way we could check whether our sample was representative of 'general health'.

Members of the Dutch population were sampled by TNS NIPO, a professional Dutch institute for public opinion and marketing research. This research institute has a panel of over 200,000 individuals who are interviewed regularly via the Internet. In July 2005, a sample of 1,400 individuals aged over 18 years was drawn from this panel. The sample was made representative of the national distribution of age, sex, region, urbanisation and social class. Like in the Belgium sample, we asked the general health question to check whether the sample was representative for general health. The distribution of the outcome of the general health question was compared to the figures of the Central Bureau of Statistics ([www.cbs.nl](http://www.cbs.nl)).

#### *The QoL-AGHDA in GHD patients*

Patient data from Belgium and the Netherlands were retrieved from the KIMS database. The KIMS database is part of an international pharmaco-epidemiological survey, launched in 1994 at the request of endocrinologists and health-care decision-makers to monitor the outcomes and safety of long-term GH replacement therapy (Genotropin®) in adults with GHD being treated in a conventional clinical setting. The study to date contains data on more than 14,000 patients from 31 countries. The other aims of KIMS are to improve understanding of the consequence of GHD in adult hypopituitarism and to contribute to optimisation of GH replacement [20].

The Belgian patient sample consisted of 370 participants [mean (SD) age at entry into KIMS: 43 (15.1) years] and the Dutch sample comprised 286 participants [44 (15.5) years]. The majority of patients in both countries acquired pituitary insufficiency and consequent GHD during adulthood (77 and 76% in Belgium and the Netherlands, respectively). The severity of hypopituitarism, expressed as a number of pituitary hormone deficits, varied from isolated GHD to panhypopituitarism (18% of patients). The proportion of the number of deficits was similar in both countries except for isolated GHD. Isolated GHD was less common in Belgian patients (8%) than in Dutch patients (13%). The profile of co-morbidities did not differ between countries with 13–16% of patients reporting hypertension. In both groups half of the patients developed

hypopituitarism due to the long-term consequences of surgery for pituitary adenoma with non-functioning pituitary adenoma being most frequent (~30%). QoL was assessed by the QoL-AGHDA at the entry into the KIMS in all patients and before GH replacement therapy was started.

#### *Patient subsample to test validity of regression model*

A subsample of 64 Dutch GHD patients filled in both the EQ-5D and the QoL-AGHDA. Because these patients filled in both questionnaires, we could test whether the predicted EQ-5D<sub>index</sub> (on the basis of the regression of the EQ-5D on the QoL-AGHDA in the general population) was indeed comparable with the 'standard' EQ-5D<sub>index</sub> (completed directly by the patients). The mean age of the Dutch patients who filled in both the EQ-5D and the QoL-AGHDA was 42.6 (SD = 14.6), and 39% of the patients were female.

#### *Analyses*

Scores elicited from disease-specific questionnaires such as the QoL-AGHDA in the general population are most likely to be skewed, because the questionnaire focuses on a specific pathology. The use of mean and standard deviation to determine 'deviations from normal' is therefore limited. For this reason we also present the 90 and 95% cutoff points of the normal population, which could serve as clinical benchmarks. The reference values of the general population samples were weighted for 'general health' to adjust for differences between the samples and general population health.

We used a multiple regression model to estimate the EQ-5D<sub>index</sub> from the individual 25 QoL-AGHDA items in the Belgian and Dutch general population data. All yes/no answers on the QoL-AGHDA were added as dummies. In addition to these 25 items, sex and age were added to the model. In this investigation we were not aiming to explain the relationship between the QoL-AGHDA and EQ-5D, but only trying to achieve the best prediction possible. For this reason the most parsimonious model was not an issue here [21]. Thus, the model included all 25 items of the QoL-AGHDA, irrespective of whether the univariate contribution was statistically significant. We did two sensitivity analyses. To illustrate what would happen if we had tried to achieve the most parsimonious model, we removed all non-significant contributions from the Dutch model in a 'backwards elimination analysis'. In addition, we made a simple regression of the EQ-5D<sub>index</sub> on the QoL-AGHDA total sum score.

Criterion validity between the estimated EQ-5D<sub>index</sub> from the QoL-AGHDA and the 'standard' EQ-5D<sub>index</sub> was tested by administering both questionnaires in a subsample of patients in the Netherlands. The relationship between the

estimated EQ-5D<sub>index</sub> and the ‘standard’ EQ-5D<sub>index</sub> was described by a single-measure intraclass correlation coefficient (ICC) using a two-way fixed effect model (the two questionnaires are the fixed sample) with ‘absolute agreement’ (SPSS, version 15). We also tested the difference in mean values with a paired *t*-test.

## Results

### Population reference values QoL-AGHDA and QALY values

Complete QoL-AGHDA data were obtained for 1,026 of the interviewees (response rate = 15%) of the Belgian population and 1,038 (response rate = 74%) individuals of the Dutch general population. The Belgian sample consisted of 40% French speaking responders and 60% Flemish, which represents the real distribution in Belgium. The gender and age distributions were almost equal to data from the national institutes of statistics of Belgium (<http://www.statbel.fgov.be>) and the Netherlands ([www.cbs.nl](http://www.cbs.nl)). The only exception was an underrepresentation of the middle age groups (40–60 years) in the Dutch sample, with 38% belonging to that age group although it represents 46% of the general population in the Netherlands. As the reference values are presented in age categories, this underrepresentation did not jeopardize the representation per age group. A different issue was overall health: Overall health is by definition related to health-related QoL, but there is, unlike age, no rationale to present reference values for ‘subgroups of general health’. In Table 1 it can be seen that some of the Belgian age groups were healthier, while other subgroup responders were less healthy compared to the figures of the National Institute of Statistics. Overall, the Dutch sample appears a bit healthier than the general population (Table 1). The reference values of the general population samples were therefore weighted by the reported ‘general health’.

The regression of the EQ-5D time trade-off index score on the 25 items of the QoL-AGHDA is presented in Table 2. The  $R^2$  in the Dutch general population sample was 0.482 and 0.360 in the Belgian sample.

In the first sensitivity analysis, the ‘backwards elimination analysis’, 13 of the 25 variables of the QoL-AGHDA were deleted from the model, and the  $R^2$  dropped slightly from 0.482 to 0.477. The simple regression of the EQ-5D<sub>index</sub> on the QoL-AGHDA total sum score produced a much lower  $R^2$  of 0.356.

### The QoL-AGHDA in GHD patients

Table 3 presents the patient data. Men report fewer problems than women ( $p < 0.001$ ), which is often the case in

QoL data [22]. Patients in Belgium seemed to report more problems than in the Netherlands, but this difference was not statistically significant ( $p = 0.073$ ). The differences between the countries in the EQ-5D<sub>index</sub> derived from the QoL-AGHDA must be interpreted with caution, as the difference can be attributed to both difference in ‘true patient score’ and also to other differences like the valuation function in Table 2 (see also the “Discussion” section). In Fig. 1, the estimated patient EQ-5D score is compared with the EQ-5D<sub>index</sub> of asthma [23], hypertension [23], type II diabetes [24], low back pain [23], Parkinson’s disease [25] and intermittent claudication [26]. The results suggest that the burden of GHD (mean value 0.7635) is between the burden of asthma (0.79) and diabetes (0.69), and well below the average of the general Dutch population in this study (0.88).

### Patient subsample to test the validity of the regression model

The correlation between the estimated EQ-5D<sub>index</sub> on the basis of the QoL-AGHDA and the ‘standard’ EQ-5D<sub>index</sub> was 0.407 (intraclass correlation). The difference between the estimated mean EQ-5D<sub>index</sub> on the basis of the QoL-AGHDA (0.7387) and the mean of ‘standard’ EQ-5D<sub>index</sub> (on the basis of the ‘standard’ EQ-5D in patients: 0.7027) did not reach a statistically significant difference ( $t = 1.132$ ,  $p = 0.262$ ).

## Discussion

We provided Dutch and Belgian reference values for the QoL-AGHDA and adapted the QoL-AGHDA for use in health economic evaluations. We demonstrated the criterion validity of the health economic valuation method, and we showed that the burden of GHD is considerable.

### Theoretical reflection

It is tempting to try to interpret the slightly lower QoL values of the patients in Belgium compared to the Dutch. It is, however, difficult to understand these differences, because there are several possible explanations. For instance, patients in Belgium may be more inclined to express their QoL limitations. Also a subtle difference in the French and Dutch translations of the questionnaire may cause differences. Another reason might be the structural difference between the estimated scores for the two EQ-5D indices in Belgium and the Netherlands, i.e., between values derived using the visual analogue scale technique in Belgium and those in the Netherlands derived using the time trade-off technique [27]. Moreover, subtle differences

**Table 1** Mean values of the QoL-AGHDA for the Belgian and Dutch population

	N	Mean (SD) QoL-AGHDA		Score worse 90%	Score worse 95%	Score worse 90%	Score worse 95%	Sample Report "less than good health" (%)	National Bureau of Statistics (2005) Report "less than good health" (%)	
		Non- weighted	Weighted	Non- weighted	Non- weighted	Weighted	Weighted			
<i>Belgian population</i>										
Overall	1,026	6.79 (5.47)	6.95 (5.44)	14.00	17.65	14.00	17.00	19.4	23.1	
Male	528	6.02 (5.30)	6.43 (5.33)	13.00	15.50	14.00	16.00	17.2	21.2	
18–24	57	7.71 (4.71)	7.73 (4.69)	14.00	16.10	14.00	16.07	5.9	7.2	
25–44	195	6.07 (5.66)	5.93 (5.55)	14.00	16.40	13.00	15.92	15.3	12.8	
45–64	167	6.15 (5.18)	6.35 (5.29)	14.00	16.00	14.00	16.00	21.5	24.9	
65 and older	109	5.69 (5.00)	6.77 (5.22)	13.00	14.00	14.00	15.00	20.7	42.1	
Female	498	7.42 (5.59)	7.49 (5.51)	16.00	19.00	15.47	18.04	21.7	24.8	
18–24	82	8.71 (5.63)	8.28 (5.52)	16.40	20.00	14.24	20.00	15.5	8.4	
25–44	178	7.06 (5.53)	6.70 (5.37)	15.10	18.00	15.00	17.33	21.8	15.0	
45–64	160	7.75 (5.76)	7.94 (5.80)	17.00	19.00	17.00	19.21	25.6	29.1	
65 and older	78	6.20 (5.06)	7.62 (5.07)	12.10	19.00	13.83	19.00	20.3	49.3	
<i>Dutch population</i>										
Overall	1,075	4.94 (5.15)	5.48 (5.48)	13.00	16.00	14.00	17.00	16.7	20.1	
Male	527	4.51 (4.97)	5.14 (5.40)	12.00	15.00	13.00	17.00	13.7	17.9	
18–24 <sup>a</sup>	46	3.13 (3.64)	3.13 (3.64)	8.30	12.65	8.30	12.65	0.0	9.1	
25–44	203	4.48 (5.07)	4.72 (5.31)	12.00	16.00	13.00	16.00	10.8	13.0	
45–64	181	4.92 (5.57)	6.00 (6.22)	14.00	17.00	16.00	18.00	16.6	28.1	
65 and older	97	4.48 (3.92)	5.35 (4.23)	9.20	12.00	11.00	14.00	20.6	36.0	
Female	548	5.36 (5.29)	5.82 (5.54)	13.00	16.00	14.00	17.00	19.5	22.2	
18–24	44	6.27 (6.17)	5.93 (5.87)	14.50	22.25	13.94	21.34	15.9	12.8	
25–44	214	5.25 (4.94)	5.28 (4.96)	13.00	15.00	13.00	15.00	15.4	15.8	
45–64	195	5.18 (5.23)	5.65 (5.45)	13.00	16.00	14.00	16.86	24.1	30.1	
65 and older	95	5.53 (5.76)	7.33 (6.56)	14.00	19.20	19.00	21.81	21.1	48.6	

Weighting based on the responses to the 'general health question' and the statistics of the Belgian National Bureau of Statistics (data from 2005) and Dutch National Bureau of Statistics (data from 2005)

<sup>a</sup> In the age category male, age 15 to 24, weighting does not change scores, as there are no responders with less than perfect health

between the Belgian and Dutch population samples might also contribute to the difference. Most of these explanations provide additional arguments concerning the routine use of national reference values emanating from other countries. The national reference values presented here are in line with that reasoning.

We performed two sensitivity analyses: To illustrate what would happen if we had tried to achieve the most parsimonious model, we removed all non-significant contributions from the Dutch model in a 'backwards elimination analysis'. In this, the  $R^2$  dropped only slightly from 0.482 to 0.477. An alternative to the regression of the individual item scores of the EQ-5D<sub>index</sub> on the individual QoL-AGHDA items is a regression of the EQ-5D<sub>index</sub> on the QoL-AGHDA sum score. Unlike the distinction between a full model and the parsimonious model, this variant makes a large difference: the  $R^2$  of the regression of the QoL-AGHDA sum score on the EQ-5D score is only

0.356 compared to 0.482 for the individual items. Kołtowska-Häggström et al. constructed a model to estimate the EQ-5D<sub>index</sub> score using the QoL-AGHDA sum score, the square of the QoL-AGHDA sum score and several SES variables and interactions [9]. The  $R^2$  of this "full model" was not much higher than our parsimonious model based on the QoL-AGHDA sum score, age and sex (0.38 versus 0.36). In a successive article, Kołtowska-Häggström et al. abandoned this approach, and the new model consisted of the variable age; sex and all 25 individual items of the QoL-AGHDA which all were entered as dummies, as in the present investigation [10]. This model had a  $R^2$  of 0.42, in between the 0.36 of our Belgium data and the 0.48 of our Dutch data. Thus, the results of Kołtowska-Häggström are in line with the results found here, and this suggests that, at least in the relationship between the QoL-AGHDA and the EQ-5D<sub>index</sub>, a saturated model based on the most important SES variables (age and

**Table 2** The regression of the EQ-5D time trade-off index score on the QoL-AGHDA

Variables	The Netherlands	Belgium
Constant	0.94587	0.93756
Men = 1; woman = 2	-0.00632	-0.01563
Younger than or equal to 25 = 1, otherwise 0	0.02394	0.10171
Older than 25 and younger than or equal to 50 = 1, otherwise 0	0.02968	0.02504
Q01 I have to struggle to finish jobs	-0.01966	-0.01086
Q02 I feel a strong need to sleep during the day	-0.00692	-0.01390
Q03 I often feel lonely even when I am with other people	-0.00645	-0.03131
Q04 I have to read things several times before they sink in	-0.00449	0.00052
Q05 It is difficult for me to make friends	-0.00949	0.02414
Q06 It takes a lot of effort for me to do simple tasks	-0.16876	-0.09427
Q07 I have difficulty controlling my emotions	-0.02626	-0.01413
Q08 I often lose track of what I want to say	-0.02450	0.00000
Q09 I lack confidence	0.02437	-0.00299
Q10 I have to push myself to do things	0.00552	-0.01325
Q11 I often feel very tense	-0.03686	-0.04496
Q12 I feel as if I let people down	-0.03184	-0.01642
Q13 I find it hard to mix with people	-0.00804	-0.01081
Q14 I feel worn out even when I've not done anything	-0.04822	-0.06042
Q15 There are times when I feel very low	-0.02474	-0.04879
Q16 I avoid responsibility if possible	0.00680	0.03665
Q17 I avoid mixing with people I don't know well	0.00274	-0.00084
Q18 I feel as if I'm a burden to people	-0.09989	-0.05433
Q19 I often forget what people have said to me	0.03700	-0.02075
Q20 I find it difficult to plan ahead	-0.00941	-0.01288
Q21 I am easily irritated by other people	-0.03853	-0.01329
Q22 I often feel too tired to do the thing I ought to do	-0.05507	-0.01863
Q23 I have to force myself to do all the things that need doing	0.01142	0.00491
Q24 I often have to force myself to stay awake	-0.01832	-0.01233
Q25 My memory lets me down	-0.00874	-0.01183

**Table 3** QoL-AGHDA scores in GHD patients without GH treatment and EQ-5D scores derived from those QoL-AGHDA scores

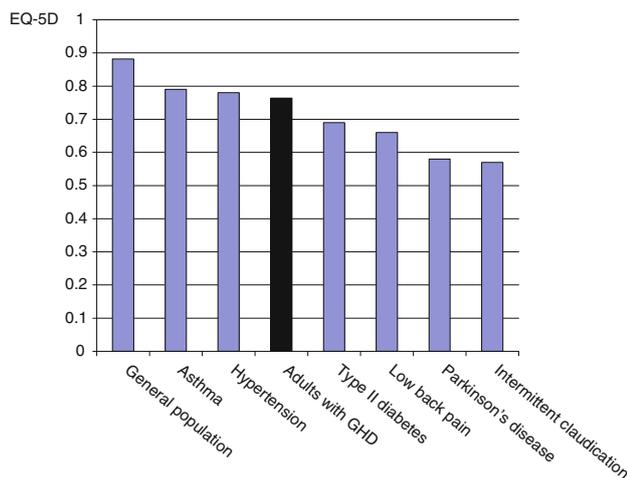
	Sum score QoL-AGHDA								QoL-AGHDA EQ-5Dscore		
	N	Mean	SD	20%	40%	Median	60%	80%	Mean	Med	SD
Belgium											
Male	161	9.52	6.22	3.00	7.00	9.00	11.00	15.00	0.7827	0.8153	0.1510
Female	138	12.21	6.58	5.80	10.00	13.00	15.00	18.00	0.7129	0.7348	0.1666
The Netherlands											
Male	116	8.40	6.67	2.00	5.80	7.00	10.00	15.00	0.8091	0.8312	0.1186
Female	118	11.28	6.75	4.00	10.00	11.00	14.00	18.00	0.7516	0.7707	0.1269

Mean age, 43.81; SD 15.27

sex) and all items of the disease-specific instrument seems to be the most favourable model.

A number of items have positive signs, which is remarkable as all items are conceptually negatively related

to QoL: an additional problem should result in a lower QoL, and thus a lower EQ-5D index. The most likely explanation is that the positive signs are reflections of interactions between the variables. Given the high number



**Fig. 1** Quality of life in terms of EQ-5D<sub>index</sub> values in adults with GHD compared to other illnesses. Values from the general population are based on the Dutch data of the present study. See text for other references

of variables we did not include interaction in the analyses, and therefore any interaction is ‘forced’ into the sign of the variables. For this reason the positive signs found in this study are not relevant in that respect, and the conceptual interpretation of the sign should be made cautiously.

## Implications

### *GHD burden in adults*

We have shown that the burden of disease in adults with GHD is considerable and comparable with other patients whose burden is undisputed. This confirms suggestions made by others [1, 16, 28], but it is the first time that such a conclusion can be underpinned with a uni-dimensional generic QoL questionnaire comparing QoL of GHD patients with other patients. Former attempts made use of multidimensional questionnaires and disease-specific questionnaires, neither of which are not suitable for making a decisive judgment [16, 28]. Using a multidimensional questionnaire, it is not clear which of the dimensions are the most important. Disease-specific questionnaires do not allow for comparisons between patient groups, as important other, i.e., non-disease-related (side) effects might be overlooked. Therefore, disease-specific questionnaires might not provide the full picture of the burden of disease. But by using both a generic instrument and a validated disease-specific instrument, we have provided new evidence of a significant burden in adults with GHD.

The evidence of a high burden of diseases found in this study is obviously not a sufficient condition either to encourage treatment for GHD or to defend the reimbursement of treatment. To justify treatment, it is necessary to

provide evidence that the burden of disease is reduced in a clinically significant way. To justify reimbursement of treatment, it is further necessary that the treatment is reasonably cost effective. Despite these limitations, burden of disease plays an important role in reimbursement, because burden of disease is an important factor in the interpretation of the cost effectiveness of treatment. If the burden of disease is high, society is more inclined to be generous in the interpretation of the cost effectiveness than when the burden of disease is low [11, 12]. This study suggests that the cost effectiveness of any GHD treatment should indeed be interpreted generously and not restricted as if it was a disease with a low burden.

## Limitations of the study

### *Response rate*

The response rate of the Dutch general population sample was much higher (74%) than that of the Belgian sample (15%). This was related to the way the samples were recruited: Dutch respondents received a relatively high compensation for their efforts, which contributed to a higher response rate. In this respect it should be noted that “representativeness is more important than response rate in survey research” [29]. It could be argued for instance that the Dutch sample was less representative of the general population as it consisted of ‘professional responders’, while the Belgian sample was closer to the classical objective of population research of achieving a ‘naïve sample’. The line of reason is that we should define “what the sample should be representative of”. In our case we wanted the samples to be representative of QoL, because we wanted to have norm values for the QoL-AGHDA. Therefore, factors that were associated with QoL were more important than factors that had no such association. As we know that age, gender and obviously ‘self-reported health’ are associated with health-related QoL [30], we weighted the samples for these factors using the figures of the National Bureaus of Statistics. In this way the samples used here represent the population of both countries with respect to the most important determinants of health-related QoL.

### *Criterion validity*

The ICC between the estimated EQ-5D<sub>index</sub> on the basis of the QoL-AGHDA and the ‘standard’ EQ-5D<sub>index</sub> was only moderate (ICC = 0.4070), but comparable with the ICC when the EQ-5D was compared with the SF-6D in patients with ankylosing spondylitis (0.45) and in patients with knee osteoarthritis (0.47) [31, 32]. The moderate ICC suggests considerable differences at the individual level,

but small differences at the group level. Indeed we found no statistically significant difference between the estimated EQ-5D<sub>index</sub> on the basis of the QoL-AGHDA and the 'standard' EQ-5D<sub>index</sub>. This means that our estimations of the EQ-5D<sub>index</sub> on the basis of the QoL-AGHDA should not be used at an individual patient level, but only when sufficient patient numbers are involved. When this methodology was employed in GHD research, sufficient patient numbers were indeed used as the data came from the large international KIMS database [9, 10].

## Conclusions

In this article we have described Belgian and Dutch general population reference values for the QoL-AGHDA and provided a transformation function for the estimation of EQ-5D<sub>index</sub> scores on the basis of QoL-AGHDA data. By obtaining QoL-AGHDA reference values for the Netherlands and Belgium, the needs of patients with GHD can now be understood more clearly. As we were able to link the scores of the QoL-AGHDA to QALYs, we can now evaluate the health outcomes of interventions better in growth hormone deficiency. Based on these outcomes of the QoL-AGHDA, we can also confirm that the burden of adults with GHD is considerable.

**Acknowledgments** This study was sponsored by the local representatives of Pfizer BV in the Netherlands and Pfizer NV/SA in Belgium/Luxemburg. The manuscript was edited by Rosalind Rabin (Euroqol group). The costs were covered by Pfizer in Belgium and the Netherlands. The KIMS database is sponsored by Pfizer Inc. M. Kołtowska-Hägström and W.J. Meerding are current employees of Pfizer. J.J.V. Busschbach, B.H.R. Wolffenbuttel and L. Annemans have no other relevant interest to declare. The authors wish to express their gratitude to the KIMS investigators in Belgium and the Netherlands who provided the data on their patients. The following hospitals participated in the additional collection of EQ-5D patient questionnaires: University Medical Centre Groningen and the University Medical Centre Nijmegen. The research benefited from the efforts and insights of Sophie Marbaix, Marjolein Pompen, Helma Monteban and Netty Dorrestijn.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Woodhouse, L.J., Mukherjee, A., Shalet, S.M., Ezzat, S.: The influence of growth hormone status on physical impairments, functional limitations, and health-related quality of life in adults. *Endocr. Rev.* **27**, 287–317 (2006)
- College voor Zorgverzekeringen. Pakketbeheer in de praktijk. Report 245. CVZ. [http://www.cvz.nl/resources/rpt0612%20pakketbeheer\\_tcm28-22507.pdf](http://www.cvz.nl/resources/rpt0612%20pakketbeheer_tcm28-22507.pdf) (2006). Accessed 1 May 2009
- Cleemput, I., Van Wilder, P., Vrijens, F., Huybrechts, M., Ramaekers, D.: Richtlijnen voor farmaco-economisch onderzoek in België. KCE Reports 78A. Federaal Kenniscentrum voor de gezondheidszorg 10/04/2008. [http://www.kce.fgov.be/index\\_n.aspx?SGREF=5264&CREF=11004](http://www.kce.fgov.be/index_n.aspx?SGREF=5264&CREF=11004) (2008). Accessed 1 May 2009
- Ministerie Van Sociale Zaken, Volksgezondheid en Leefmilieu. Koninklijk besluit van 21 december 2001 tot vaststelling van de procedures, termijnen en voorwaarden inzake de tegemoetkoming van de verplichte verzekering voor geneeskundige verzorging en uitkeringen in de kosten van farmaceutische specialiteiten. Belgisch Staatsblad, 29.12.2001 Ed 3. [www.ejustice.just.fgov.be/cgi/welcome.pl](http://www.ejustice.just.fgov.be/cgi/welcome.pl) (2001). Accessed 1 May 2009
- Bansback, N., Brazier, J., Chilcott, J., Kaltenthaler, E., Lloyd-Jones, M., Paisley, S., Walters, S.: Clinical and cost effectiveness of recombinant human growth hormone (Somatropin) in adults. Report prepared on behalf of the national institute for clinical excellence (NICE), 10 April 2002. School of Health and Related Research (ScHARR), University of Sheffield. Available from: [www.nice.org.uk](http://www.nice.org.uk) (2002). Accessed 1 May 2009
- Kołtowska-Hägström, M.: Quality of life in adult patients with growth hormone deficiency: bridging the gap between clinical evaluation and health economic assessment. PhD-Thesis of the Uppsala University (Sweden), Medicinska vetenskapsområdet, Faculty of Pharmacy, Department of Pharmacy. <http://publications.uu.se/theses> (2008). Accessed 1 May 2009
- Brooks, R., Rabin, R., De Charro, F.: The Measurement and Valuation of Health Status Using Eq-5d: A European Perspective. Kluwer Academic, London (2003)
- Stolk, E.A., Busschbach, J.J.V.: Validity and feasibility of the use of condition-specific outcome measures in economic evaluation. *Qual. Life Res.* **12**, 363–371 (2003)
- Kołtowska-Hägström, M., Jonsson, B., Isacson, D., Bingefors, K.: Using EQ-5D to derive general population-based utilities for the quality of life assessment of growth hormone deficiency in adults (QoL-AGHDA). *Value Health* **10**, 73–81 (2007)
- Kołtowska-Hägström, M., Kind, P., Monson, J.P., Jonsson, B.: Growth hormone (GH) replacement in hypopituitary adults with GH deficiency evaluated by a utility-weighted quality of life index: a precursor to cost-utility analysis. *Clin. Endocrinol.* **68**, 122–129 (2008)
- Stolk, E.A., van Donselaar, G., Brouwer, W.B., Busschbach, J.J.: Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional shortfall. *Pharmacoeconomics* **22**, 1097–1107 (2004)
- Raad voor de Volksgezondheid en Zorg. Zinnige en duurzame zorg. Zoetermeer; RvZ. [http://www.rvz.net/data/download/advies\\_Zinnige\\_zorg.pdf](http://www.rvz.net/data/download/advies_Zinnige_zorg.pdf) (2006). Accessed 1 May 2009
- Hunt, S.M., McKenna, S.P.: The QLDS: a scale for the measurement of quality of life in depression. *Health Policy* **22**, 307–319 (1992)
- McKenna, S.P., Doward, L.C., Alonso, J., Kohlmann, T., Niero, M., Prieto, L., Wíren, L.: The QoL-AGHDA: an instrument for the assessment of quality of life in adults with growth hormone deficiency. *Qual. Life Res.* **8**, 373–383 (1999)
- Wirén, L., Whalley, D., McKenna, S.P., Wilhelmsen, L.: Application of a disease-specific, quality-of-life measure (QoL-AGHDA) in growth hormone-deficient adults and a random population sample in Sweden: validation of the measure by Rasch analysis. *Clin. Endocrinol.* **52**, 143–152 (2000)
- Kołtowska-Haggstrom, M., Mattsson, A.F., Monson, J.P., Kind, P., Badia, X., Casanueva, F.F., Busschbach, J., Koppeschaar, H.P., Johannsson, G.: Does long-term GH replacement therapy in hypopituitary adults with GH deficiency normalise quality of life? *Eur. J. Endocrinol.* **155**, 109–119 (2006)
- National Institute for Health and Clinical Excellence (NICE): Human Growth Homone (somatropin) in adults with growth

- hormone deficiency (TA 64). <http://guidance.nice.org.uk/TA64/Guidance/Recommendation> (2003). Accessed Jan 2010
18. Szende, A., Oppe, M., Devlin, N.: EQ-5D Value Sets: Inventory, Comparative Review and User Guide. Springer, Dordrecht (2007)
  19. Lamers, L.M., McDonnell, J., Stalmeier, P.F., Krabbe, P.F., Busschbach, J.J.: The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ.* **15**, 1121–1132 (2006)
  20. Gutiérrez, L.P., Kołtowska-Hägström, M., Jönsson, P.J., Mattsson, A.F., Svensson, D., Westberg, B., Luger, A.: Registries as a tool in evidence-based medicine: example of KIMS (Pfizer International Metabolic Database). *Pharmacoepidemiol. Drug Saf.* **17**, 90–102 (2008)
  21. Busschbach, J.J.V., McDonnell, J., Essink-Bot, M.L., Van Hout, B.A.: Estimating parametric relations between health description and health valuation with an application to the EuroQol instrument. *J. Health Econ.* **18**, 551–571 (1999)
  22. Orfila, F., Ferrer, M., Lamarca, R., et al.: Gender differences in health-related quality of life among the elderly: the role of objective functional capacity and chronic conditions. *Soc. Sci. Med.* **63**(9), 2367–2380 (2006)
  23. Burström, K., Johannesson, M., Diderichsen, F.: Swedish population health-related quality of life results using the EQ-5D. *Qual. Life Res.* **10**, 621–635 (2001)
  24. Koopmanschap, M.: Coping with type II diabetes: the patient's perspective. *Diabetologia* **45**, S18–S22 (2002)
  25. Siderowf, A., Ravina, B., Glick, H.A.: Preferencebased quality-of-life in patients with Parkinson's disease. *Neurology* **59**, 103–108 (2002)
  26. Bosch, J.L., Hunink, M.G.: Comparison of the Health Utilities Index Mark 3 (HUI3) and the EuroQol EQ-5D in patients treated for intermittent claudication. *Qual. Life Res.* **9**, 591–601 (2000)
  27. Craig, B.M., Busschbach, J.J., Salomon, J.A.: Modeling ranking, time trade-off, and visual analog scale values for EQ-5D health states: a review and comparison of methods. *Med. Care* **47**, 634–641 (2009)
  28. Van Rooijen, L., Beckers, A., Stevenaert, A., Busschbach, J.J.V., Van Eijk, W., Rutten, F.F.H.: The burden of hypopituitarism in adults after pituitary surgery. *Endocrinol. Metab.* **4**(Suppl. B), 139–142 (1997)
  29. Cook, C., Heath, F., Thompson, R.L.: A meta-analysis of response rates in web- or internet-based surveys. *Educ. Psychol. Meas.* **60**, 821–836 (2000)
  30. Kind, P., Dolan, P., Gudex, C., Williams, A.: Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* **316**, 736–741 (1998)
  31. Boonen, A., Van der Heijde, D., Landewé, R., Van Tubergen, A., Mielants, H., Dougados, M., Van der Linden, S.: How do the EQ-5D, SF-6D and the well-being rating scale compare in patients with ankylosing spondylitis? *Ann. Rheum. Dis.* **66**, 771–777 (2007)
  32. Xie, F., Li, S., Luo, N., Lo, N., Yeo, S., Yang, K., Fong, K., Thumboo, J.: Comparison of the EuroQol and short form 6D in Singapore multiethnic asian knee osteoarthritis patients scheduled for total knee replacement. *Arthritis Rheum.* **57**, 1043–1049 (2007)