Quality of Life in Adult Patients with Growth Hormone Deficiency

_Bridging the gap between clinical evaluation and health economic assessment_

MARIA KOŁTOWSKA-HÄGGSTRÖM
Dissertation presented at Uppsala University to be publicly examined in B21, BMC, Husarg 3, Uppsala, Friday, January 18, 2008 at 13:15 for the degree of Doctor of Philosophy (Faculty of Pharmacy). The examination will be conducted in English.

Abstract

The goals of this thesis are to evaluate quality of life (QoL) in adult patients with growth hormone deficiency (GHD) in relation to population normative data, to construct a preference-weighted index (utility) from a disease-specific QoL measure and to assess it in a clinical context.

The study included samples from the general population and patients with GHD from four European populations: England & Wales, the Netherlands, Spain and Sweden. The country-specific patient cohorts were retrieved from KIMS (Pfizer International Metabolic Database).

A questionnaire was developed that contained items from existing QoL questionnaires including, among others, Quality of Life Assessment in Growth Hormone Deficiency in Adults (QoL-AGHDA) and the EQ-5D. The QoL-AGHDA is a disease-specific measure for use in adults with GHD. The EQ-5D is a generic instrument which describes health states for which country-specific preference-based weights are available. Thus, it was possible to generate preference-weighted indices (utilities) based on data generated by both instruments.

This thesis reports QoL-AGHDA normative values for the populations of England & Wales, the Netherlands, Spain and Sweden, and confirms the extent of QoL impairment in patients with GHD in comparison with the general population. Long-term GH replacement resulted in sustained improvements in overall QoL towards normative country-specific values, as well in most of the dimensions that were impaired before treatment.

For use in health economic evaluations, models for generating utilities (QoL-AGHDA utility) from QoL-AGHDA were developed. It is believed that these models may facilitate medical decision making, given that they provide a tool for obtaining utilities in the absence of directly collected preference-weighted indices.

QoL-AGHDA utility effectively monitored treatment effects in patients with GHD. Moreover, this study confirmed a QoL-AGHDA utility deficit before treatment and a gain after starting GH replacement.

The novel aspect of the present approach was to apply preference-weighted indices derived from a disease-specific measure to assess QoL in the clinical context, together with patient demographic and clinical characteristics. The robustness of this analysis is reinforced by the fact that utilities in both general and patient populations were generated using the same methodology.

Keywords: growth hormone deficiency in adults, quality of life, cost-utility analysis, growth hormone replacement, normative data, QoL-AGHDA

Maria Kołtowska-Häggström, Department of Pharmacy, Box 580, Uppsala University, SE-75123 Uppsala, Sweden

© Maria Kołtowska-Häggström 2007

ISSN 1651-6192
ISBN 978-91-554-7052-4
urn:nbn:se:uu:diva-8353 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8353)
To Mama, Tato, Łukasz, Kaśka, Kuba and Bo
I don’t regret
Agnieszka Osiecka

That you didn’t give me
Green-eyed dreams
No, I don’t regret, Mama
That I didn’t know treasure
Or laced words
I don’t regret
That you didn’t tell me how to sneak happiness from under the counter
That you didn’t teach me the masquerade of life
The grey caress of the tormented days
No regret, no regret

I don’t regret
Quite the reverse, thank you very much, my Dear
That you let me go
To live as I have done

That in this Country I have lived
These hard years
I don’t regret
And that eventually I will learn
That - it’s just the way it is
I don’t regret
That They don’t organize a holiday from humiliation
And They will not return my smile
The grey caress of the tormented days
No regret, no regret

No, I don’t regret
Quite the reverse, thank you very much, my Country
For any day of the week
And for a suitcase full of hope

No, I don’t regret
Quite the reverse, thank you very much
That you are my Country
That you are my paradise and my underworld

Translation Paul Kind and Maria Koltowska-Häggström
Translated and reproduced with permission Copyright © Agata Passent
Nie, nie żaluję
Agnieszka Osiecka

Że nie dałaś mi, Mamo
Zielonookich snów
Nie, nie żaluję
Że nie znalazłam klejnotów
Ni koronkowych słów
Nie, nie żaluję
Że nie mówiłaś mi jak szczęście kraść spod lady
I nie uczyłaś mnie życiowej maskarady
Pieszczoty szarej tych umęczonych dni
Nie żal mi, nie żal mi

Nie, nie żaluję
Przeciwnie bardzo ci dziękuję, Kochana
Żeś mi odejście pozwoliła
Po to bym żyła tak jak żyłam

Że w tym Kraju przeżyłam
Tych trudnych parę lat
Nie, nie żaluję
Że na koniec się dowiem
Ot, tak się toczy świat
Nie, nie żaluję
Że nie załatwił mi urlopu od pogardy
I że nie zwrócę mi uśmiechu jak kokardy
Pieszczoty szarej tych udręczonych dni
Nie żal mi, nie żal mi

Nie, ja nie żaluję
Przeciwnie bardzo ci dziękuję, mój Kraju
Za jakiś czwartek jakiś piątek jakiś wtorek
I za nadziei cały worek

Nie, nie żaluję
Przeciwnie bardzo ci dziękuję
Za to, że jesteś moim Krajem
Że jesteś piekłem mzym i rajem

Reproduced with permission Copyright © Agata Passent
List of publications

This thesis is based on the following papers, which will be referred to by their roman numerals in the text.

I. Kołtowska-Häggström M, Hennessy S, Mattsson AF, Monson JP, Kind P, Quality of Life Assessment of Growth Hormone Deficiency in Adults (QoL-AGHDA): comparison of normative reference data for the general population of England and Wales with results for adult hypopituitary patients with growth hormone deficiency, Hormone Research, 2005, 64:46-54


III. Kołtowska-Häggström 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 in Health 2007, 10 (1):73-81

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>15</td>
</tr>
<tr>
<td>Background</td>
<td>17</td>
</tr>
<tr>
<td>Growth hormone deficiency</td>
<td>17</td>
</tr>
<tr>
<td>Growth hormone/insulin-like growth factor axis</td>
<td>17</td>
</tr>
<tr>
<td>Growth hormone deficiency (GHD) in adult patients – overview</td>
<td>18</td>
</tr>
<tr>
<td>Clinical characteristics</td>
<td>18</td>
</tr>
<tr>
<td>Epidemiological data</td>
<td>18</td>
</tr>
<tr>
<td>Primary aetiology</td>
<td>19</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>19</td>
</tr>
<tr>
<td>GH dosing</td>
<td>19</td>
</tr>
<tr>
<td>Response to treatment</td>
<td>19</td>
</tr>
<tr>
<td>Mortality</td>
<td>20</td>
</tr>
<tr>
<td>Quality of life</td>
<td>20</td>
</tr>
<tr>
<td>Definition and basic concepts</td>
<td>20</td>
</tr>
<tr>
<td>QoL measurement</td>
<td>24</td>
</tr>
<tr>
<td>Application of QoL measurements</td>
<td>28</td>
</tr>
<tr>
<td>QoL data in pharmacoeconomic evaluations</td>
<td>28</td>
</tr>
<tr>
<td>General remarks</td>
<td>28</td>
</tr>
<tr>
<td>Methodological issues in eliciting utilities</td>
<td>29</td>
</tr>
<tr>
<td>Deriving utilities from condition-specific measures</td>
<td>30</td>
</tr>
<tr>
<td>QoL and pharmacoeconomic evaluations in adult GHD</td>
<td>31</td>
</tr>
<tr>
<td>QoL instruments used in adult GHD</td>
<td>31</td>
</tr>
<tr>
<td>QoL in untreated patients</td>
<td>32</td>
</tr>
<tr>
<td>Effects of GH treatment on QoL</td>
<td>35</td>
</tr>
<tr>
<td>Pharmacoeconomic evaluations</td>
<td>36</td>
</tr>
<tr>
<td>Cost of illness/burden of illness studies</td>
<td>36</td>
</tr>
<tr>
<td>Aims</td>
<td>39</td>
</tr>
<tr>
<td>Study populations and methods</td>
<td>40</td>
</tr>
<tr>
<td>Study design</td>
<td>40</td>
</tr>
<tr>
<td>Questionnaires</td>
<td>40</td>
</tr>
<tr>
<td>Ethical considerations</td>
<td>42</td>
</tr>
<tr>
<td>Study populations</td>
<td>43</td>
</tr>
<tr>
<td>General populations</td>
<td>43</td>
</tr>
<tr>
<td>Patients with GHD</td>
<td>46</td>
</tr>
</tbody>
</table>
Handling of missing data .................................................................................................48
Analytical procedures ....................................................................................................50
Deriving population normative data for QoL-AGHDA scores (papers I, II and III) .................................................................50
Modelling to assess QoL impairment and evaluate response to treatment for both general scores and dimensions (paper II) ........50
Deriving QoL-AGHDA_{utility} for Sweden and England & Wales (papers III and IV) ..................................................................53
Computation of QoL-AGHDA_{utility} at baseline, total gain and gain per year and comparison with general population values (paper IV) ......54

Results ..................................................................................................................................56
Population normative values for total QoL-AGHDA scores (papers I, II and III) ........................................................................56
The effects of GH replacement therapy on QoL in relation to normative values (papers I and II) .................................................................57
Comparison with untreated patients in England & Wales (paper I) ...57
Treatment effects – four-country analysis (paper II) .........................................................58
Preference-weighted index based on the QoL-AGHDA – QoL-AGHDA_{utility} (papers III and IV) .................................................................60
The Swedish population (paper III) ..................................................................................60
The English and Welsh population (paper IV) ....................................................................62
The effects of GH as measured by QoL-AGHDA_{utility} compared with population values (paper IV) .................................................................62
Patient characteristics .....................................................................................................62
QoL-AGHDA_{utility} ...........................................................................................................64
Patient subgroups .............................................................................................................66

Discussion .............................................................................................................................69
Methodological issues .........................................................................................................69
Restriction to country-specific data ..................................................................................69
Sources of data ..................................................................................................................70
From a disease-specific measure to a preference-weighted index (utility) ........................................................................................................72
QoL-AGHDA_{utility} as a treatment outcome in the clinical setting ......76
Main findings .......................................................................................................................76
Population normative data for QoL-AGHDA (papers I and III) ......76
The effects of GH replacement therapy on QoL (paper II) ..........................78
A preference-weighted index (QoL-AGHDA_{utility}) (papers III and IV) .................................................................80
Final remarks ....................................................................................................................83

Acknowledgements ............................................................................................................85
References .............................................................................................................................89
List of abbreviations

**General**

AACE  American Association of Clinical Endocrinologists
ACh    acetylcholine
AO-GHD adult-onset GHD
BMD    bone mineral density
BMI    body mass index
CO-GHD childhood-onset GHD
E&W    England & Wales
GH     growth hormone
GHD    growth hormone deficiency
GHRH   GH-releasing hormone
GHS    GH secretagogue
GRS    Growth Hormone Research Society
HR-QoL health-related quality of life
IGF-I  insulin-like growth factor-I
ISPOR  International Society for Pharmacoeconomics and Outcomes Research
ITT    insulin tolerance test
KIGS   Pfizer International Growth Database
KIMS   Pfizer International Metabolic Database
LOCF   the last observation carried forward technique
NFPA   non-functioning pituitary adenoma
NICE   National Institute for Health and Clinical Excellence
PAI-1  plasminogen activator inhibitor type I
PRO    patient-reported outcomes
QALY   quality adjusted life year
QoL    quality of life
QoL-AGHDA utility estimated preference-based index (utility) based on QoL-AGHDA scores (utility-weighted QoL-AGHDA)
R²     coefficient of multiple determination
RCT    randomized, placebo-controlled, double-blind clinical trials
RTB    registry of population permanently living in Sweden
SCB    Swedish National Statistic Office
SD     standard deviation
SDS standard deviation score
SMR standardized mortality rate
TBI traumatic brain injury
TNS NIPO the Dutch Institute for Public Opinion and Market Research
t-PA tissue-type plasminogen activator
WHO World Health Organization

**QoL measures**
BAS Brief Anxiety Scale
BSI Brief Symptom Inventory
BDI Beck Depression Inventory
CMI Cornell Medical Index
CPRS Comprehensive Psychopathological Rating Scale
DIS Disease Impact Scale
DSQ Disease Specific Questionnaire
EWI Experimental World Inventory
FACT-L Functional Assessment of Cancer Therapy-Lung
FLZM Fragen zur Lebenszufriedenheit Module
GHIQ Growth Hormone Injection Questionnaire
GHQ (-28,-60) General Health Questionnaire (28-, 60-items)
GWBI (S) General Well-Being Index (Schedule)
HADS Hospital Anxiety and Depression Scale
HDQoL Hormone Deficiency-Specific QoL Questionnaire
HDRS Hamilton Depression Rating Scale
HSCL-56 Hopkins Symptom Checklist
KSQ Kellner Symptom Questionnaire
KIMS PLSF KIMS Patient Life Situation Form
LFS Life Fulfilment Scale
MADRS Montgomery Asberg Depression Rating Scale
MFQ Mental Fatigue Questionnaire
MMPI-2 Minnesota Multiphasic Personality Inventory-2
NHP Nottingham Health Profile
OCS Obsessive Compulsive Scale
PGI Patient-Generated Index
PGWB Psychological General Well-Being Schedule
POMS Profile of Mood States
QLS-H Questions on Life Satisfaction–Hypopituitarism
QoL-AGHDA Quality of Life Assessment of Growth Hormone Deficiency in Adults
SAS-SR Social Adjustment Scale – Self Report
SCAN Schedule for Clinical Assessment in Neuropsychiatry
SCL-90 Symptom Checklist
SES Self-Esteem Scale
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-36</td>
<td>Short Form 36 Health Survey</td>
</tr>
<tr>
<td>SG</td>
<td>Standard Gamble</td>
</tr>
<tr>
<td>SIP</td>
<td>Sickness Impact Profile</td>
</tr>
<tr>
<td>SRS</td>
<td>Social Relationship Scale</td>
</tr>
<tr>
<td>STAI</td>
<td>State-Trait Anxiety Inventory</td>
</tr>
<tr>
<td>TTO</td>
<td>Time Trade-Off</td>
</tr>
<tr>
<td>VSP-AM</td>
<td>Vécu et Santé Perçue de l’Adolescent – Malade</td>
</tr>
<tr>
<td>W-BQ (12)</td>
<td>Well-Being Questionnaire (12 items)</td>
</tr>
<tr>
<td>WHOQOL</td>
<td>World Health Organization Quality of Life</td>
</tr>
</tbody>
</table>
The question is not only about actual living but also living a good life. Such a desire is inherent in human nature and applies to any circumstance in life: being rich or being poor; being born in a highly industrialized world, or in a remote village far away from civilization; being young or being old; being highly or poorly educated; or, finally, being in full health or suffering from incurable disease. Everybody wants to live, to be happy, and to have a good life (or at least a better life).

Historically, medical treatment focused primarily on the quantity of life i.e. on preventing premature mortality or extending survival, but often neglected its quality. Progress in medicine has resulted in an increasing number of successful interventions, particularly in patients with fatal disorders, thus raising a problem of quality, in other words – not only: “to live or not to live” but also “how to live”. At the same time, patients’ voices are becoming more widely acknowledged and their subjective well-being is turning into a recognized outcome of medical interventions. Therefore, nowadays, any successful medical outcome refers not only to “saving the patient’s life” but also to “saving the patient’s good life”. Thus, patients hope not only that their lives will be saved, but also that treatment will make them feel good.

Life could be viewed as having two relatively independent dimensions: quantity and quality. This thesis focuses on the latter. Although current medical practice considers both the quantity and quality of life, a problem arises when other key players come into the picture who have different standpoints with specific values and expectations. One perspective is clinically oriented and represents an individual approach, with care for a single patient to achieve the best effects of treatment often irrespective of costs. The other comes from the health policy makers. They are responsible for providing medical care to society at large and by definition do not look at individual patients. They act for society to secure the most efficient (optimal effect at minimal costs) health service to the whole population. Clearly such
contradictory interests must lead to misunderstandings, conflicts and common frustration.

Another reason for such a situation can be the lack of a clear and generally approved definition of the term “quality of life” (QoL) (65, 77, 84), and consequentially a large variation in the conceptual basis for practical applications, resulting in huge differences in methodological approaches. Thus, pharmacoeconomic evaluation, commonly used in medical decision making, requires that health status is expressed as a preference-based single summary score (a health status index), which is capable of identifying and quantifying differences across diseases as well as aggregate changes in health status over time (175). By contrast, clinical applications usually require a measure that captures specific changes within a certain disease, in patient populations (in clinical trials) and in individual patients (in daily clinical practice) (66).

Despite the differences between clinicians and health economists – or more widely medical decision makers – there is a strong interdependence between them and one cannot exist in a professional setting without the other. Clinicians need funding for treatment, and funding without an executive structure is worthless. There is clearly a need for speaking the same language and for mutual understanding of information needs. But how does one make it happen?

One way worth trying, would be to provide common tools, or if this is not possible to prepare the means for inter-translations. So far such tools for measuring QoL have been largely developed for individual needs, meant to operate within a narrow and specific context. On the one hand, unfortunately, as these tools serve exclusively their purposes and are limited to restricted applications, the information they collect seems not to be accessible to other users. On the other hand, fortunately, this problem is being recognized and research into the interrelationship between preference-based instruments and clinically oriented measures has been initiated. For example, studies in obesity and gastroenterology have been published (29, 38), but there are many other medical disciplines where similar work needs to be done. It should be emphasized that well-refined means of converting treatment effects into units applicable to resource allocation are critical, particularly when expensive procedures are involved.

This thesis is about quality of life. It deals with adult hypopituitary patients who require life-lasting hormonal replacement – growth hormone – which in the long run might save their lives, and gradually restores their energy, vitality and life drive, in short – their will to live. The treatment, for all that, is relatively expensive.

This thesis, then, aims at helping bridge the gap between clinicians and health economists; “to measure what is measurable, and to make measurable what is not so” (Galileo Galileo 1564-1642).
Background

Growth hormone deficiency

Growth hormone/insulin-like growth factor axis

Human growth hormone (GH) is a polypeptide of 191 amino acids, produced in the anterior pituitary gland. It is secreted mostly during sleep in a pulsatile manner and acts either directly or via insulin-like growth factor-I (IGF-I). GH secretion is regulated centrally at different levels and peripherally by feedback mechanisms with IGF-I and glucose playing crucial roles (Figure 1) (152). GH is released throughout the lifespan in a variable mode in different phases of life i.e. higher secretion is observed during childhood and diminishing in adulthood, being the lowest in the elderly. Women have higher twenty-four-hour integrated GH serum concentrations than men (152).

Figure 1. Central and peripheral regulation of growth hormone production and secretion. ACh – acetylcholine; GH – growth hormone; GHRH – GH-releasing hormone; GHS – GH secretagogue; IGF-I – insulin-like growth factor

Modified figure from Rees and Scanlon, The physiology of the growth hormone/insulin-like growth factor axis In: Growth Hormone Deficiency in Adults: 10 Years of KIMS, Abs R, Feldt-Rasmussen U, (eds) Oxford PharmaGenesis™ pp 15-28; Reproduced with permission.
Despite its misleading name, GH is not only an essential stimulator of post-natal growth and development but is also primarily a potent anabolic hormone that modulates a large variety of physiological functions e.g. energy balance (152), lipid (2) and protein metabolism, body composition, body fluid control (130), bone growth and mineralization (152), cardiovascular function (71), and, finally mood, cognition, memory and learning (109) as well as general well-being and QoL (86).

Growth hormone deficiency (GHD) in adult patients – overview

Clinical characteristics

Given the broad range of GH actions, the widespread diversity in phenotype of patients with GHD is not at all surprising; furthermore, depending on the patients’ age, different disease presentations should be expected and, indeed, this is the case. In children, the predominant presentation is growth retardation, whilst after completion of linear growth, adverse metabolic changes, increased cardiovascular risk, abnormal body composition, reduced bone mineral density and impaired QoL are the major features (134). Overall, the adverse effects of GHD in adults lead to an increased cardiovascular risk and have a negative impact on daily life (3). In adults the disease can persist from childhood – childhood-onset GHD (CO-GHD), or arise in adulthood – adult-onset GHD (AO-GHD). GHD may occur as an isolated hormonal deficit or be part of multiple pituitary hormone deficiency (panhypopituitarism).

Epidemiological data

According to the Society for Endocrinology’s estimate, the prevalence of GHD is 3 in 10,000 of the adult population, with one-third developing it during adult life (173). Nevertheless, conversely, KIMS (Pfizer International Metabolic Database) indicates that most (78%) adult patients had AO-GHD (133). Surprisingly, the reported prevalence of GHD varies greatly between countries. Thus, it is in numbers (given per one million): in the UK the prevalence is 100 – 200, in north-western Spain – 455, and in France – it is 46 cases (80). The most thorough evaluation of incidence rates was carried out by Stockholm for the entire Danish population. The highest incidence rate was identified for CO-GHD in men (2.58 per 100,000 per year) and the lowest was the incidence of AO-GHD in women (1.42 per 100,000 per year). Interestingly, an increase in incidence over time was observed for all subgroups, with the exception of women with AO-GHD (166).
Primary aetiology

GHD in adults most frequently results from tumours in the hypothalamic-pituitary region or as a consequence of their treatment e.g. surgery or radiotherapy. In the KIMS database, pituitary adenomas account for almost 50% of cases and half of these are non-functioning pituitary adenomas (NFPA). The second most frequent group are patients with idiopathic disease (17.4%), followed by craniopharyngioma (11.4%) (80). Other underlying conditions – for example, cranial tumours distant from the hypothalamic-pituitary region, and traumatic brain injury (TBI) – are listed in the KIMS classification list (Appendix A). It is worth emphasizing that cranial irradiation – for example, for leukaemia or other malignancies outside the cranium – is very likely to result in hypopituitarism (63).

Diagnosis

As the symptoms and signs in adult GHD are non-specific, the diagnosis should be based on a combination of all clinical features, i.e. the medical history, phenotype and biochemical examinations. The most commonly recommended provocative tests of GH secretion are the insulin tolerance test (ITT), using a combination of arginine and GHRH, arginine alone or glucagon. The cut-off point for the ITT regarded as sensitive enough to diagnose severe GHD by the Growth Hormone Research Society is 3 μg/L and the cut-off points for other tests are as described by the producers (82). Conversely, the corresponding cut-off recommended by the American Association for Clinical Endocrinology is 5 μg/L for all provocative tests (1).

GH dosing

In adults, the dose should be tailored from a low dose (0.15 – 0.30 mg/day) individually on the basis of clinical evaluation, IGF-I levels and tolerance of treatment. The objective is to obtain the best clinical response and safety profile as well as to maintain IGF-I levels (as IGF-I is a sensitive marker of GH action) within the upper range of age- and gender-normative values, with the lowest effective dosage (74). GH is administered as a daily subcutaneous injection in the evening (82). It is worth noting that women require higher doses than men (74).

Response to treatment

GH replacement therapy in GH-deficient patients reverses the unfavourable consequences of GHD in a sustained manner. With regard to body composition, during GH treatment lean body mass increases by a mean of 2.0–5.5 kg and fat mass decreases by approximately 4–6 kg. Similar beneficial changes
are also observed relating to lipid metabolism. In the bones GH increases bone turnover and during the first 6 months of treatment bone resorption dominates bone formation; however, after 12–24 months of treatment new bone mass is accrued. Although the GH effect on muscle strength is less obvious – with a few studies showing improvements after prolonged GH treatment (at least 12 months) – overall, exercise performance has been demonstrated to become enhanced (42, 101).

Mortality
Premature mortality in adults with hypopituitarism has been documented. As early as 1990, Rosén and Bengtsson reported in the total cohort of patients with hypopituitarism a standardized mortality rate (SMR) of 1.8, with female patients tending to have an even worse prognosis (SMR 2.83) (155). Generally, these findings were confirmed by later studies (16, 39, 169, 172) with the exception of the study by Bates (17). Excess mortality, with higher rates in women, was also observed in a large cohort of patients with pituitary adenoma (n=2279, SMR 2.0) (141). Recently all patients (n=2205) with GHD in the Danish population were identified and evaluated for the risk of premature death. These results were consistent with previously published data and indicated a higher SMR in patients with CO-GHD (8.3 in men; 9.4 in women) than AO-GHD (1.9 in men; 3.4 in women) (167).

There has been considerable uncertainty whether or not this tendency can be reversed by appropriate GH replacement therapy. Only a limited amount of information is available. The first analysis of mortality rates in patients on GH followed in the KIMS database comprised a population of 1903 individuals with 2334 patient-years of observation. The results indicated that GH replacement in adults with GHD was not associated with increased mortality (20), although they should be interpreted with caution (despite a high number of patients, the mean observation time was short and there is a risk of selection bias in this type of observational study). Another study reported normalized overall mortality rates in a cohort of 289 GH-deficient patients receiving GH who were prospectively followed for a mean of almost 5 years (169). This study was, however, limited by the low number of patients, which reduced the statistical power.

Quality of life
Definition and basic concepts
QoL, under different names, such as eudaimonia, well-being, happiness, having a good life or satisfaction with life, has always been a feature of human existence both at the individual and the societal level. Individually, it
often constitutes the main aim of person’s life and for societies it is predominant in political and social sciences, and economics as well as medicine and medical decision making. Although the determinants of QoL are far beyond the scope of this thesis, just to illustrate the complexity, the following issues that are included in the quality-of-life index, developed by The Economist Intelligence Unit (190) are mentioned: material well-being, health, political stability and security, family life, community life, climate and geography, job security, political freedom and gender equality. For the present work, QoL is discussed only for its application in medicine and medical decision making and the term is limited to the health determinant. Although in this light the term health-related QoL (HR-QoL) seems to be more appropriate, it has been decided to keep the general term QoL, as it has been traditionally used in all relevant publications.

HR-QoL is defined by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) as “a broad theoretical construct developed to explain and organize measures concerned with the evaluation of health status, attitudes, values, and perceived levels of satisfaction and general well-being with respect to either specific health conditions or life as a whole from the individual perspective” (22). This definition focuses more on the functional aspects of the concept than on the concept as such i.e. subordinates the content to the aims and functions.

Fayers and Machin (66) listed some forms of QoL concepts – for example, a general evaluation of health with a focus on functional status as well as checklists of symptoms, visual analogue scale (VAS), reintegration to normal living, personal well-being, impact of illness on social, emotional and family domains, existential approach, expectation model and patient-preference measure (to be discussed in detail later). The authors emphasize the underlying premise valid for all approaches, such as subjectivism, with patients being a unique source of information about “the issues that are of fundamental importance to patients’ well-being”. It is, however, debatable whether or not a restrictive view on measuring QoL only in an individual patient should have its exemptions. The problem arises whenever QoL is considered in small children, patients with communication problems or patients with diminished intellectual capacity (106). In these situations, as appropriate, so-called proxy measures can be seen as acceptable.

The importance of QoL in patient management is clearly expressed in the World Health Organization’s (WHO) definition of health: “a state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity” (189), where different aspects of well-being are the principal attributes of full health. In the context of this WHO definition of health, the notions of QoL or HR-QoL and health status can be viewed as indistinguishable or at least substantially overlapping and hence can be used interchangeably.
There is, however, an ongoing debate over which theoretical models are most appropriate, i.e. scientifically relevant and practically applicable.

An early model, offered by Ware (187) (Figure 2), specifies three generic health concepts – namely, physical health, mental health and general health – and depicts them as a specific-generic continuum. The first two, physical and mental health, are separately described at both specific and generic levels. In brief, physical conditions are hypothesized to be closely linked to physical symptoms and these, given that they are severe enough, lead to physical limitations and compromised well-being. The same reasoning is valid for mental conditions: from mental symptoms to psychological distress and well-being. Eventually both modulate perceptions of health in general.


An even more complex model (Figure 3), also based on a continuum, was elaborated by Wilson and Cleary (192). The health outcomes in this proposal are divided into five levels – biological and physiological factors, symptoms, functioning, general health perceptions, and overall quality of life. As described by the authors, moving along this continuum of increasing complexity from the left end to the right, one “moves out from the cell to the individual to the interaction of the individual as a member of the society”. This model captures the reciprocal relationships between all of the components and explains the levels of integration.
Nagi’s model of disablement, based on the International Classification of Impairments, Disabilities and Handicaps, is used to structure limitations experienced by patients and related to both GH deficit and excess (acromegaly) (195). In this concept, impairment is defined as a structural abnormality at any anatomical level (cells, tissues, organs), functional limitation refers to a difficulty in performing activities and disability is categorized into physical, mental, social and emotional disability and covers fulfilment of personal role in life.

It is worth emphasising that in each of these models, particular aspects of human health do not exist as separate entities, being placed in a kind of vacuum, but contribute together to create one integrated organism with multiple interrelations and cross-dependences. Therefore, analysing any of them in the absence of the others may lead to misinterpretations and cause considerable controversy.

Need-based model

Human motivation, with individuals’ needs being the primary motivating factors, laid down the foundation of the need-based theory (94) in QoL research. In this view, fulfilment of human needs secures life satisfaction and consequently life quality depends on the personal capacity to satisfy these needs. It is assumed that the highest QoL is equal to the fulfilment of all needs while the lowest is when only a few needs are being met. There is a long list of such needs, for example food, drink, sleep, activity, sex, warmth, security, love, communication, curiosity, identity, recognition and many others. In this concept, poor health interferes, in most cases adversely, with satisfying these needs, and thus has a negative impact on QoL. Nevertheless, this model assumes that as long as the primary needs are fulfilled – for example, by compensation mechanisms – QoL remains unaltered. In other
words, if because of a disease the usual way of satisfying needs is no longer feasible, but the needs can still be met in another manner, then the disease does not have any impact on QoL. Several instruments, among them the QoL-Assessment of Growth Hormone Deficiency in Adults (QoL-AGHDA), are rooted in this model.

QoL measurement

There is an abundance of QoL instruments. A bibliographic study of patient-assessed health outcome measures detected no fewer than 3921 papers which reported on the development and testing of such measures and an increase of new instruments was observed from 144 in 1990 to 650 in 1999 (75). The identified instruments were categorized as: dimension specific (n=690), disease or population specific (n=1819), generic (n=865), individualized (n=62) or a utility measure (n=409).

Traditionally, only two of these categories are referred to, namely, generic and disease specific. The former are designed for general use, both in diseased people, regardless of the condition they suffer from, and in general populations. Although, many of them are limited to physical functioning, there are also those that encompass psychological issues. The most widely recognized instruments are the Nottingham Health Profile (NHP), Psychological General Well-Being (PGWB) (described elsewhere), Short Form 36 Health Survey (SF-36) and EQ-5D (to be described in detail later). They allow for comparisons across different conditions, which is their major advantage (66). However, they are often regarded as not being sensitive enough to detect changes in certain diseases.

Disease-specific questionnaires sometimes referred to as disease-sensitive or disease-oriented measures, have been developed to meet the increasing demand for monitoring patients with a higher level of precision, which in turn requires tools focusing on selected characteristics, specific for the disorder they address. Obviously, the GHD-specific instruments discussed in this thesis fall into this category. Dimension-specific instruments, as defined by Garratt and colleagues, focus on particular domains, for example anxiety and depression, pain, cognitive functioning or fatigue. Individualized measures allow respondents to include items of their choice, based on their own life experience. Utility measures (preference measures) share the property of being based on social preferences and are thus legitimate for use in pharmacoeconomic evaluations. As utility measures are intended for general use, it is disputable whether they should be treated as a separate category or should be classified as generic.

Item structure, form and scales

The item structure partly determines the form of a measure, either index or profile. Single-item instruments generate an index, represented by a single
number, whereas multi-item instruments can produce either a profile or an index. Profiles generate a set of scores which represent each dimension measured by an instrument. They provide insights into the nature of QoL attenuation and allow detection of its specificity, as well as targeted monitoring of specific features (67). Thus, profiles are most suitable in clinical practice, though at the same time, they are incapable of capturing the magnitude and often even the direction of overall change in QoL. For some profiles, a simple sum of dimension scores is permissible, although its accuracy is arguable given that such a procedure assumes equal importance for each dimension, which often may not be the case. This problem can be overcome by applying weights i.e. relative values for each dimension (or even item). Derived in this way, a single aggregated score is believed to be robust and appropriate (106).

There are different manners of quantifying information collected by QoL instruments, most of them being based on an ordinal scale. The simplest are dichotomous variables with a choice of yes/no answers that can categorize health status (e.g. non diseased/diseased), and are also frequently employed in QoL measurement (e.g. QoL-AGHDA). Such a descriptive classification not only distinguishes between different categories but also orders them hierarchically. The most frequently used scales in measuring QoL are various types of rating scales. By and large they are structured around Likert’s scale (114) (Figure 4a) and the VAS (Figure 4b).
Figure 4. Examples of a) Likert scale, b) Visual analogue scale (VAS), c) “Feeling thermometer” from the EQ-5D
Likert’s scale is constructed as a series of statements with a number of alternatives (choices), typically organized from the lowest level of measurement (not important at all, the worst, completely disagree, etc) to the highest (extremely important, the best, fully agree) with intermediate alternatives in between. Although five choices are used most frequently, the number varies from three to nine. The alternatives are coded as sequential numbers e.g.1–5 for the purpose of analysis but the scale maintains ordinal properties. As it lacks a well-defined unit of measurement, it indicates a direction of change but does not assess its magnitude. It cannot be assumed that for example the distance from #1 ‘not important at all’ to #2 ‘little important’ is the same as between #2 ‘little important’ to #3 ‘important’ – in other words, that the change from #1 to #2 is equal to the change from #2 to #3 (106). Additionally, the choices are made subjectively, i.e. values behind them differ among respondents.

Another form of rating scale is a VAS, depicted by a 10 cm horizontal line on which a respondent is expected to mark his/her evaluation of the relevant problem. A VAS is anchored at both ends by extremes such as ‘the worst possible’ and ‘the best possible’ and the respondent’s answer is computed as the distance measured from the lower or left hand end. In this context, VAS presents a continuous scale (100) and is applicable for measuring many diverse concepts. VAS served as a basis for a “feeling thermometer” (Figure 4c), which is most often used in measuring health-state preferences (see below). The top of the thermometer corresponds to a value of 100 and is described as the best alternative, whereas the bottom of the thermometer represents the worst (21). The “feeling thermometer” is included in the EQ-5D.

Cross cultural translations and validations; country-specific normative data

Along with the increasing globalization and cross-cultural communication, as well as the escalating number of international clinical trials, the need for individual instruments in different language versions becomes obvious for any involved party. Although there is also considerable consensus on the requirements such versions must fulfil, the methodology remains markedly controversial. As for the former, the original questionnaire and all its language versions must be conceptually equivalent i.e. express the same concepts, and not literal meanings; each language version must be culturally relevant and acceptable to the target population and they must also be psychometrically comparable (4). The most commonly recommended translation and validation process is the one reported as “Translation and cultural adaptation of patient reported outcomes measures – principles of good practice” and prepared by the ISPOR task force group (191). The alternative method is referred to as the dual-translation panel (171).
Given the most rigorous technique employed and supposing that the new language version matches the source instrument perfectly in all aspects, the question remains whether true cross-cultural differences exist and, if so, to what extent these may modify the magnitude of impairment in the diseased population. This question is of particular relevance when the data from several countries are being pooled for the final analysis.

In interpreting QoL assessments in relation to normative data not only the country of origin plays a vital role, but also the method of sampling. The general population samples can be selected at random, and hence also include diseased individuals, or be restricted to the healthy population. The importance of age and gender should also be emphasized as both are substantial predictors of QoL perception. Finally, there might be a shift over time, so that ideally both patient and reference data should be collected at the same time (69).

Application of QoL measurements
QoL has emerged as an important construct that has found numerous applications across healthcare-related fields, ranging from randomized controlled trials and pharmacoeconomic evaluations through to daily clinical practice. Each of these applications imposes different requirements on the QoL measures. Clinical applications usually require a measure that captures specific changes within a certain disease, in patient populations (in clinical trials) and in individual patients (in daily clinical practice) (144).

On the other hand, pharmacoeconomic evaluation often requires that health status is expressed as a single summary score (a health status index), which is capable of identifying and quantifying differences between diseases as well as aggregate changes in patients’ health status over time (175).

QoL data in pharmacoeconomic evaluations
General remarks
Generally, there are two types of analysis: cost-benefit, where both costs and treatment outcomes are expressed in money; and cost-effectiveness, where costs valued in money are compared with a single primary outcome. Cost-utility analysis is a special form of cost-effectiveness analysis and requires a primary outcome measured as a gain in quality-adjusted life year (QALY), (48). This unit of measurement combines information on the length of life (quantity) and the quality of life, where the latter is measured by utilities on a scale that has values of 1 and 0 respectively for full health and dead (Figure 5).
Figure 5. Concept of quality-adjusted life years (QALY) shown on the example of English & Welsh general populations and adult patients with growth hormone deficiency. The area under the curve represents QALYs. A (dotted line) depicts general population values, B – a gain during treatment, and C – values for patients without treatment. The value 1 on the y-axis stands for full health and 0 for death. From Koltowska-Häggström M, Kind P, Monson JP and 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 2007 Clinical Endocrinology Published article online: 4-Sep-2007 doi: 10.1111/j.1365-2265.2007.03010.x by Blackwell Publishing Ltd. Copyright © 2007, Official Journal of the Society for Endocrinology. Reproduced with permission.

The unit QALY is therefore defined as 1 year of life with full health (176). In the theoretical basis for QALY provided by Torrance, full health was defined as ‘perfect functioning’ (174). Utilities are individual’s preferences for a certain health state assigned under uncertainty (68).

Methodological issues in eliciting utilities

When a QoL index is used to calculate QALY benefits, health economists also require that the value of health should be estimated in terms of utility weights using preference measurement techniques such as Time Trade-Off (TTO) or Standard Gamble (SG); however, a simple VAS evaluation has also mostly been accepted as valid (60).

In the TTO method, a respondent is asked how many years of life in a chronic condition (serving as a proxy for an examined health state) she/he is willing to trade-off for a life in full health. The time horizon is usually based on the life expectancy of the respondent, but this arrangement is sometimes modified so that a fixed time horizon is used. Where this is the case, then the typical time horizon for a health state in question is 10 years. The expecta-
tion is that the less desirable the health state, the more years of life a respondent will be willing to trade-off. The utility of a given health state is expressed as the ratio of the number computed as ten minus the number of traded-off years. The SG technique offers a choice of two alternatives: 1) treatment with two outcomes: good, with a defined probability \( p = 1 \), and bad, with a probability \( = 1 - p \); 2) health state in question. Respondents define the lowest probability to achieve a good outcome i.e. the highest risk of a bad outcome (e.g. death) they are prepared to take. The point when they chose alternative 2 and are not willing to further decrease probability, defines the utility of the measured health state. VAS as described above is a line with clearly defined end points of full health and death and a respondent is expected to place the rates of a health state in-between. (176).

There is an extensive controversy as to the robustness and appropriateness of these methods, starting with theoretical and ethical considerations and ending with practical issues (142, 146, 162). The accumulating evidence has demonstrated inconsistency in results, both in respect of order of states and in magnitude of utilities generated (142, 154, 177, 178). Although it is believed that SG and TTO usually generate higher values than rating scales, several studies have reported different results, which in some of them varied for various severities of health states (13). Another open question addressing preferences (utilities) that should constitute the basis for reference – patients or general population (179) – was transformed by Dolan to address the issue of the relevance of taking the patients’ adaptation to poor health states into account when assigning values to those states. To facilitate consensus on this issue he provided a conceptual framework, with the weights for different levels of experience and anticipation of illness being incorporated as appropriate (58). In discussing the source of preference weights, the problem of country-specific weights should not be neglected as many studies have applied weights derived from different societies to their own work. The final question to be asked is whether hypothetical or actual health states should be considered and in relation to which anchors (142) these should be evaluated.

Despite the controversy, SG and TTO remain the gold standards for measuring health utilities. Nevertheless, there are also indirect ways of generating preferences by using generic QoL instruments, such as the Health Utilities Index (HUI2 and HUI3), the Short Form 6D (SF-6D) and the EQ-5D (122).

**Deriving utilities from condition-specific measures**

As they are not preference based, condition-specific measures lack legitimacy for direct use in cost-utility analysis. Hence there is a need for transformation to derive utilities, particularly when the generic instruments are incapable of detecting changes or when such QoL data assessed by generic measures, simply do not exist. Although, several studies have reported that
generic measures have good discriminative properties when used to determine disease severity (110, 122), others have questioned this and developed utilities from disease-specific instruments (29, 38, 44, 107, 115).

Finally, the importance of reliable data on mortality for QALY computation cannot be underestimated. To take a wider view, in addition to mortality data, precise epidemiological data in terms of incidence, prevalence and co-morbidities is equally important for medical decision making.

QoL and pharmacoeconomic evaluations in adult GHD

QoL instruments used in adult GHD

Despite the fact that QoL impairment belongs to the key features of adult GHD, there is no consistently recommended methodology for QoL examination, and therefore numerous and various instruments have been used in clinical trials and daily practice. Hence, making comparisons and drawing robust conclusions has become relatively difficult, if not impossible (151). The QoL measures used in the published studies can be categorized as:

1. Generic – measuring general well-being in any population
2. GHD-specific – designed specifically for use in patients with GHD
3. GHD-adapted – modified versions of measures originally designed for use in non-GHD patients
4. Other specific – designed to measure disorders other than GHD or to measure specific areas of well-being
5. Preference measures
6. Informal measures

In a literature search for QoL instruments in adult GHD (108), of the total of 55 papers relating to QoL and preference measures, eighteen QoL and four preference measures were identified (Table 1). Patient-reported outcome (PRO) questionnaires were used in twelve papers. Two generic questionnaires were frequently encountered: the Nottingham Health Profile (NHP) (33 studies) and SF-36 (7 studies). The SF-6D, a version of SF-36, modified for pharmacoeconomic application, was reported in one paper. There were three instruments developed specifically for adult GHD: Growth Hormone Deficiency Questionnaire (GHDQ) (51), QoL-AGHDA (126) and Questions on Life Satisfaction–Hypopitutarism (QLS-H) (87, 89). The QoL-AGHDA was more common than the QLS-H; however, QLS-H was developed a few years after QoL-AGHDA. There were four GHD-adapted measures and nine from different disease areas which were often used to examine QoL in patients with GHD. Among the latter, PGWB (61) or its adaptation for use in the UK (General Well-Being Index/Schedule – GWBI/GWBS)
(125) was the most commonly employed. One study (32) evaluated the response to GH treatment based on the partner’s assessment, using an in-house developed 12-item partner questionnaire [used also by McMillan (127)]. In addition, a number of cognition and psychological tests were traced. These aimed to analyse intellectual properties, personal traits, level of sexual identification, reactions to stress, perception of body image and relation to the environment (6, 54,160).

Recently, a few more measures have been either developed or exploited in QoL assessment in GH-deficient adults. These include: instruments traditionally administered in depression (119), other generic measures of psychological well-being such as the Well-Being Questionnaire (W-BQ) (128), the recently developed Hormone Deficiency-Specific individualised QoL questionnaire (HDQoL) (128), as well as two new questionnaires for adolescents with growth disorders: Growth Hormone Injection Questionnaire (GHIQ) (49) and Vécu et Santé Perçue de l’Adolescent – Malade (VSP-AM) (50). All identified measures, with the exception of pure cognition and psychological tests, are summarized in Table 1.

QoL in untreated patients

As early as 1962, Raben reported improved vigour, well-being and ambition in a 35-year-old hypopituitary patient treated with GH (150). Similar observations have been reported repeatedly (33). The most consistently observed complaints in untreated adult patients with GHD are related to energy levels, vitality, mental fatigue and emotional reactions (24, 124, 185, 193) as well as to social isolation (124, 156) and anxiety (124). Reduced self-confidence has also been reported (24), as has a disturbed sex life (156), decreased physical mobility (24), dissatisfaction with body image (92), poor memory (95), reduced cognitive function and decreased mood (54), as well as attention deficits (180). A higher incidence of mental disorders, more pronouncedly expressed mental distress and relatively frequent cognitive dysfunctions, have also been demonstrated (40). The most common features of QoL impairment in adult GHD are listed in Table 2.
Table 1. QoL instruments used in adult GHD

<table>
<thead>
<tr>
<th>Category</th>
<th>Measure</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic</td>
<td>Cornell Medical Index (CMI)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Nottingham Health Profile (NHP)</td>
<td>Profile</td>
</tr>
<tr>
<td></td>
<td>Short Form 36 Health Survey (SF-36)</td>
<td>Profile</td>
</tr>
<tr>
<td></td>
<td>SF-6D</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Well-Being Questionnaire (W-BQ)</td>
<td>Index</td>
</tr>
<tr>
<td>GHD-specific</td>
<td>Growth Hormone Deficiency Questionnaire (GHDQ)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Growth Hormone Injection Questionnaire (GHIQ)</td>
<td>Index/Profile</td>
</tr>
<tr>
<td></td>
<td>Hormone Deficiency-Specific QoL questionnaire (HDQoL)</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>Quality of Life Assessment of Growth Hormone Deficiency in Adults (QoL-AGHDA)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Questions on Life Satisfaction–Hypopituitarism (QLS-H)</td>
<td>Profile</td>
</tr>
<tr>
<td></td>
<td>Vécu et Santé Perçue de l’Adolescent – Malade (VSP-AM)</td>
<td>Index/Profile</td>
</tr>
<tr>
<td>GHD-adapted</td>
<td>Disease Impact Scale</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Life Fulfilment Scale</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Questions on Life Satisfaction Modules - QLS (M)*</td>
<td>Index</td>
</tr>
<tr>
<td>Other specific</td>
<td>Comprehensive Psychopathological Rating Scale (CPRS)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>General Health Questionnaire (GHQ)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>General Well-Being Index (GWBI)**</td>
<td>Index/Profile</td>
</tr>
<tr>
<td></td>
<td>Hopkins Symptom Checklist (HSCL-56)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Kellner Symptom Questionnaire (KSQ)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Mental Fatigue Questionnaire</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Minnesota Multiphasic Personality Inventory-2 (MMPI-2)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Personality Assessment Schedule</td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td>Profile of Mood States (POMS)</td>
<td>Profile</td>
</tr>
<tr>
<td></td>
<td>Psychological General Well-Being Schedule (PGWB)</td>
<td>Index/Profile</td>
</tr>
<tr>
<td></td>
<td>Self-Esteem Scale</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Sjöberg Mood Questionnaire</td>
<td>Profile</td>
</tr>
<tr>
<td></td>
<td>Symptom Checklist (SCL-90)</td>
<td>Index</td>
</tr>
<tr>
<td>Depression &amp; anxiety measures</td>
<td>Brief Anxiety Scale (BAS)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Brief Symptom Inventory (BSI)</td>
<td>3 Indices</td>
</tr>
<tr>
<td></td>
<td>Corrected-GHQ</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Hamilton Depression Rating Scale (HDRS)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Hospital Anxiety and Depression Scale (HADS)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Montgomery Asberg Depression Rating Scale (MADRS)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Obsessive Compulsive Scale (OCS)</td>
<td>Index</td>
</tr>
<tr>
<td></td>
<td>Schedule for Clinical Assessment in Neuropsychiatry (SCAN)</td>
<td>Profile</td>
</tr>
<tr>
<td></td>
<td>State-Trait Anxiety Inventory (STAI)</td>
<td>Index</td>
</tr>
<tr>
<td>Preference assessments</td>
<td>Conjoint analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Likert Scale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time Trade-Off (TTO)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visual Analogue Scale (VAS)</td>
<td></td>
</tr>
<tr>
<td>Social functioning</td>
<td>KIMS Patient Life Situation Form (KIMS PLSF)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social Adjustment Scale – Self Report (SAS-SR)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social Relationship Scale (SRS)</td>
<td></td>
</tr>
</tbody>
</table>

* Before a module for hypopituitarism was developed. ** British version of PGWB
Table 2. Features of QoL impairment in adult GHD

<table>
<thead>
<tr>
<th>Feature</th>
<th>Exercise physiology</th>
<th>Cognition</th>
<th>QoL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise capacity</td>
<td>↓</td>
<td>↓ Short-term memory</td>
<td>↓ Energy</td>
</tr>
<tr>
<td>Maximal exercise performance</td>
<td>↓</td>
<td>↓ Long-term memory</td>
<td>↓ Sexual life</td>
</tr>
<tr>
<td>Isometric quadriceps force</td>
<td>↓</td>
<td>↓ Iconic memory</td>
<td>↓ Self-esteem</td>
</tr>
<tr>
<td>Cross-sectional muscle area</td>
<td>↓</td>
<td>↓ Concentration</td>
<td>↓ Self-confidence</td>
</tr>
<tr>
<td>Maximum oxygen uptake</td>
<td>↓</td>
<td>↓ Attention</td>
<td>↓ Life fulfilment</td>
</tr>
<tr>
<td>Submaximal aerobic performance</td>
<td>↓</td>
<td>↓ Perceptual-motor performance</td>
<td>↓ Body image</td>
</tr>
<tr>
<td>Myosin heavy chain</td>
<td>↓</td>
<td></td>
<td>↓ Mood</td>
</tr>
<tr>
<td>Physical leisure activities</td>
<td>↓</td>
<td></td>
<td>↓ Sleep</td>
</tr>
<tr>
<td>Physical mobility</td>
<td>↑</td>
<td></td>
<td>↓ Perception of mental health</td>
</tr>
<tr>
<td>Complaints of muscle weakness and fatigue</td>
<td>↑</td>
<td></td>
<td>↓ Stamina</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↓ Life drive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↓ Vigour</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↓ Perception of general health</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Social isolation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Psychological lability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Anxiety</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Depression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Mental fatigue</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Tiredness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Pain</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑ Irritability</td>
</tr>
</tbody>
</table>

↓ - decreased/ impaired/ deteriorated/ reduced
↑ - increased/ improved/ alleviated

Decreased QoL in adults with GHD has also been demonstrated in relation to normative values in a large open study in six European countries (France, Germany, Italy, the Netherlands, Spain and UK) and in the United States. QoL was measured by QLS-H and both raw scores and z-scores were computed (27).

That said, it should be acknowledged that there is great individual variability in the perception of QoL and that not all patients experience QoL
reduction to the same degree (90, 91). Some of the observed variability can be attributed to the time of disease onset. There are data suggesting that patients with CO-GHD experience a reduction in QoL to a lesser degree and thus the rate of response to GH replacement is limited (11). On the other hand, age is not likely to influence the extent of QoL impairment (113, 131, 132).

Another question is whether the observed reduction of QoL is GHD-related or a non-specific consequence of a chronic condition. Compared with diabetic patients, patients with GHD reported significantly more depression and mental fatigue as well as significantly less self-esteem and life fulfilment (118, 186). On the other hand, comparison with patients who had undergone mastoid surgery did not show any significant difference in SF-36 and GWBS scores between the groups (148).

Finally, there are three studies which report longitudinal observations on QoL in untreated patients. In the study of Badia et al (12) the average reduction in QoL compared to normal population data and measured by QoL-AGHDA, was sustained over 12 months of observation, being significantly worse in relation to general population normative data at both points in time. These findings were supported by the results from a 10-year follow-up of eleven untreated patients (76), while Gilchrist et al (78) reported deterioration in QoL in twenty-seven patients who remained off GH for 9 years.

**Effects of GH treatment on QoL**

The impact of GH replacement therapy on QoL in hypopituitary patients is controversial and the results of placebo-controlled clinical trials are at variance.

The majority of double-blind, randomized, placebo-controlled (RCT) studies were of 6-month duration, followed by an open-label period for an additional 6–12 months. Overall, the patient cohorts included individuals of both genders with a wide age range (20–60 years of age) who had severe GHD, mainly of adult onset and usually of mixed origin (idiopathic and organic lesions) with various degrees of hypopituitarism. The most commonly used QoL measures were NHP and PGWB. Many of the studies demonstrated significant improvements, particularly in energy and vitality domains, and also in cognitive functioning, which were further enhanced during the open-label phase (11, 19, 32, 41, 51-53, 79, 119, 124, 139, 145, 183, 185). In almost all of these papers the benefits of GH were revealed by some but not all instruments employed and/or were attributed to certain but not all dimensions measured – while, finally, they seemed to be heavily dependent on the degree of the initial impairment. Parenthetically, the latter observation was confirmed in the study of Murray and colleagues (137, 138). The other studies at least partly refuted these findings (18, 23, 45, 55, 72, 161, 188).
Apart from randomized clinical trials (RCT) there are other clinical study settings which have been used to investigate GH treatment effects. Some of these studies were based on an RCT follow-up in an open design, (10, 54, 55, 76, 78, 193) or were just open-label studies (135). Most of them confirmed the positive impact of long-term GH substitution. Moreover, studies seeking GH dose optimization demonstrated an enhanced QoL (5, 59, 137).

Other types of studies exploited in research on GH effects in hypopituitary patients have been the so-called “withdrawal” studies (127, 160) and surveillance studies (based on large observational research databases). There are several publications based on analyses from such studies showing improvement in QoL in the total cohorts (88, 157, 158, 170) and, in some, but not all, specific patient subgroups such as in Cushing syndrome and acromegaly (70), patients with TBI (43), elderly patients (131, 132), patients with craniopharyngioma (184), Sheehan’s syndrome (103), or those who were irradiated (120).

An overview of the studies that investigated effect of GH replacement on QoL is presented in Appendix B.

Pharmacoeconomic evaluations

Cost of illness/burden of illness studies

The first assessment of the burden of GHD was undertaken in a Belgian cohort of adults with untreated GHD (n=129). The study showed reduced QoL (measured by SF-36) and a higher rate of unemployment due to health problems (11% of patients vs. 4.8% of the Belgian population), accompanied by a higher annual number of sick leave days in those who were employed (twice as high as the population average). Healthcare utilization, expressed as an annual number of visits to general practitioners and specialists as well as hospitalization days, was also higher than in the general population. Overall, the estimated annual healthcare costs and costs of diminished productivity per hypopituitary patient approximately doubled the mean annual costs per person for the Belgian population (85).

The results of large Spanish longitudinal survey in untreated GH-deficient patients (n=356) reflected a similar tendency (159). The detailed Swedish analysis of total societal costs of hypopituitarism when conventional hormonal replacement for all pituitary axes except for GH was administered, yielded concordant results (62).

Several studies also considered information about healthcare utilization such as the number of doctor visits, hospital days, and sick leave days before and during treatment (32, 51, 88, 104, 158, 170, 183, 184)
Cost-effectiveness/cost-utility analyses

The costs of GHD treatment are predominately driven by the usage of drug and thus are relatively easy to estimate. In the USA, for example, the annual cost of treatment would vary from a maximum of $17,185 to a minimum of $4,599 per patient (151).

The first full expert reports for evaluating the value of GH treatment for GH-deficient adults were prepared in 1995 and 1997 by the Wessex Institute for Health Research and Development (7, 8). Benefits from GH replacement were listed as follows: increased exercise capacity, near normalization of body composition, improved cardiac structure and function and increased BMD. Reduced cardiovascular mortality and morbidity and reduced fracture risk were recognized as potential benefits. The response of QoL to GH was based only on Burman et al ‘s study (32), and the report concluded that there were no benefits in administering GH over administering placebo. The total annual cost of treatment was estimated as between £3,400 and £7,400 per patient. The cost-utility of GH substitution in adults was regarded as uncertain and the final recommendation for treatment categorized as borderline.

Another systematic review and economic evaluation of GH replacement in adults was done in 2002 (also in the UK) for use by NICE (National Institute for Clinical Excellence) (31). The benefits in this evaluation were restricted to QoL and judged overall as modest. On the other hand, in the meta-analysis the authors found statistically significant improvements in social isolation (measured on the NHP) in patients treated with GH, whilst analyses from individual trials indicated improvements for pain, emotional reactions and sleep. The annual cost was estimated as £3,424 per patient based on an average maintenance dose.

Finally, an analysis of clinical and cost effectiveness of GH was prepared on behalf of NICE (14). The model was based on 34 studies and initially included as benefits of treatment: improvement in QoL, lipid levels and bone mineral density. However, it concluded that the long-term implications of lipid lowering and reduction of fracture risk have very little economic impact. Therefore, the model mainly focused on QoL. Meta-analysis of the NHP scores in the studies included indicated that only social isolation was significantly improved in the GH-treated group; energy and emotional reactions also showed small benefits, whilst pain, sleep and physical mobility tended to favour placebo. In addition to the NHP analysis, the estimate of a change in QoL measured by the QoL-AGHDA across seven studies was an improvement of 3.7 points in QoL-AGHDA score. To derive utility values for the QoL changes, two-step mapping was undertaken. First, NHP scores were mapped on to the SF-6D, which generated NHP-based utilities and these were subsequently mapped on to QoL-AGHDA scores (56). Based on this model, NICE recommended GH replacement in adults with multiple pituitary hormone deficits, who have a GH peak response of less than 9
mU/litre in the ITT or a similarly low result in another reliable test and a score on the QoL-AGHDA of at least 11 with an improvement by 7 points after 9 months of GH replacement therapy (140).
Aims

The overall aim of this thesis is to gain clinically oriented insights into QoL in adult GHD in relation to normative data and to construct a preference-weighted index applicable to health economic evaluations.

Therefore, this thesis comprises the following steps:

1. Deriving population normative data for the applied QoL measure (QoL-AGHDA)
2. Investigating the QoL deficit in patients with adult GHD, defined by the normative values
3. Constructing a preference-weighted index (QoL-AGHDAutility) using general population data
4. Examining the applicability of QoL-AGHDAutility in the clinical setting

The aims of the studies comprising the thesis are as follows:

1. To evaluate QoL-AGHDA detailed normative data for the English & Welsh (E&W) (paper I) and the Swedish (paper III) populations and compare the former with patient data.
2. To evaluate the effects of GH replacement therapy on QoL, measured by the QoL-AGHDA in relation to normative data (E&W, the Netherlands, Spain and Sweden) (paper II).
3. To evaluate specificity in QoL impairment and response to GH with regard to QoL dimensions (paper II).
4. To construct a preference-weighted index (utility) based on QoL-AGHDA items (QoL-AGHDAutility) for clinical monitoring that could be used in cost-utility analyses for Swedish and E&W populations (papers III and IV).
5. To evaluate the effects of GH, measured by QoL-AGHDAutility in relation to population values and examine potential specificity in different patient subgroups (paper IV) – E&W population.
Study populations and methods

Study design
The present research included the general population and adult patients with GHD from four European populations: England & Wales, the Netherlands, Spain and Sweden. Population data were collected between 2003 and 2005 by cross-sectional surveys (postal or electronic) in all countries except for Spain, where the data were acquired earlier by trained interviewers (12). The country-specific patient cohorts were retrieved from KIMS (Pfizer International Metabolic Database) (83).

Questionnaires
For the purpose of these studies the following information from a specifically constructed questionnaire (Appendix C) was used:

1. QoL-AGHDA (126)
2. EQ-5D (a generic measure of HR-QoL) (64)
3. Questions from the KIMS Patient Life Situation Form (KIMS PLSF) about an individual’s general situation and social functioning (88)
4. A standard five-point rating scale of self-reported health status (In general, how would you say your health has been? – excellent, very good, good, fair, poor) (73).

The constructed questionnaire was administered to the general population samples, but not to the patients. The patient data on QoL-AGHDA and KIMS PLSF were retrieved from the KIMS database and, as KIMS does not collect information on EQ-5D and self-reported general health (83), these were not available in patients.

The recipients were asked to answer the questionnaire, if not specified otherwise, based on their current health state, which was explained as: “This questionnaire asks you some general questions about your health. There are no right or wrong answers; we are just interested in how you are feeling. (…)

40
For each group of statements please indicate which one best describes your health today”.

**QoL-AGHDA**

QoL-AGHDA is a need-based measure that was constructed based on in-depth interviews with adult patients with GHD (n=35; 14 men; age range 20–59 years) attending the Christie Hospital in Manchester, UK.

Almost all patients were dissatisfied with their body image and complained of lack of energy (94% and 91%, respectively), 83% had problems with memory and concentration, 71% described themselves as being short-tempered and easily irritated, 66% suffered from lack of strength and stamina, 63% experienced reduced physical and mental drive and 57% had difficulties coping with stressful situations and avoided external stimulation (92).

The pool of items was prepared based on the interviews and finally the measure was constructed of 25 items that evoke yes/no answers, acknowledging or denying certain problems. The QoL-AGHDA score is computed by summing a number of recognized problems i.e. each “yes” answer is assigned a score of 1, and therefore a high numerical QoL-AGHDA score denotes poor QoL. It is recommended that incomplete questionnaires should be excluded from the analysis (126).

Five language versions (English for use in the UK, Swedish, German for Germany, Italian and Spanish for Spain) were developed simultaneously using dual translation panels (171). The language versions for the US, France (111), Belgium, the Netherlands, Denmark, Norway and Iceland were developed later. Recently, the Japanese version has been published; however, it needs to be emphasized that it was produced with a different methodology (back/forward translation) (168).

Before the initiation of the present research, population normative data were available for only two countries: Spain (12) and Sweden (194). Both sets of data confirmed the profound QoL deficit in adult patients with GHD compared with the general population. For the first time the country-specific QoL-AGHDA data were presented in 1999 (by John Monson at the International Symposium on Growth and Growth Factors in Endocrinology and Metabolism, Boston 1999). It is noteworthy that a few published studies (US, UK) assessed QoL in GH-deficient adults and healthy controls using QoL-AGHDA (15, 121). However, the number of individuals was rather small and these cannot be viewed as representative of the population values.

**EQ-5D**

The EQ-5D, designed as a generic instrument to measure health, independent of diagnosis or disease severity (64), defines five dimensions of health: mobility, self-care, usual activities, pain or discomfort and anxiety or depression. Each dimension is categorized into three levels of burden: 1) no problem, 2) a moderate problem or 3) an extreme problem. The combination of
three categories for five dimensions generates 243 health states with dead and unconscious, in addition. The respondents first indicate the level of burden that is best applicable and then record their perception of their general health state on the feeling thermometer (VAS) (Figure 4c). The feeling thermometer is anchored at 0 and 100, corresponding to the worst and the best imaginable health states, respectively (105). The descriptive part of EQ-5D produces a series of five-digit numbers that code for different health states and that can be matched to the relevant score from the VAS scale. For example, full health is depicted by 11111 and the poorest health by 33333. The EQ-5D is available in more than 50 languages and is being used in many clinical and economic studies as well as population surveys all over the world. Over the years, population surveys in many countries have been conducted to collect country-specific preference weights for health states described by the EQ-5D and a few years ago an attempt to derive a single matrix for Europe was undertaken (81).

PGWB

For external validation data from the PGWB were used. The PGWB was produced as an index that was thought to measure self-representation of intrapersonal emotional states to reflect a sense of subjective well-being or distress (61). It contains six states (dimensions) – namely anxiety, depressed mood, positive well-being, self-control, general health and vitality- with three to five items assessed for each dimension. The items are rated on a scale of 0–5 for their intensity or frequency; 0 being the worst and 5 – the best score. The sum of item scores constitutes an overall PGWB index score which ranges from 0–110. A high score denotes good QoL. McKenna & Hunt adapted the PGWB for use in the UK and Europe and this modified version is often referred to as the General Well-Being Index (GWBI or AGWBI) (125).

Ethical considerations

The study complied with the Declaration of Helsinki (153).

Samples from the general population

As the participants of the survey in E&W and the Netherlands belonged to regular survey panels, they had given general written informed consent to be re-contacted for the purpose of health surveys. The study protocol for the Spanish survey was approved by the Spanish Health Authority and all individuals consented to participation in the study (12). The Swedish survey contained a letter of informed consent, which explained to participants how personal data would be handled and described the background, objectives and design of the study. It also stated that by returning questionnaires, respondents were consenting to participate in the study. The study complied
with the Swedish confidentiality law and was approved by the Statistics Sweden Ethics Board.

**Patient cohorts**

All patients were followed in the KIMS database. It is a condition of entry to KIMS that each centre obtains approval from its local ethics committee and that patients give informed consent, either verbally or in writing, depending on local legal requirements. Additionally, enrolment of participants follows good clinical practice guidelines (83).

**Study populations**

**General populations**

**England & Wales (E&W) (papers I, II and IV)**

*Survey*

The constructed questionnaire was sent out in the first half of 2003 to 1190 members of a general population survey panel maintained by the Outcomes Research Group in York. The panel comprises members of the general public over 18 years of age in E&W who have been recruited over a period of the 3 preceding years and had taken part in at least one survey related to the assessment of health status. A reminder letter was sent out to non-responders after the initial mailing.

*Population sample*

The responders comprised 1007 individuals – 435 men (43%) and 572 women (57%). There were 183 non-responders, of whom 103 were men and 80 were women. The response rates were 81% for men and 87% for women.

Because there were very few respondents (n=6) aged 18–19, this age group was excluded from the analyses. The complete QoL-AGHDA was received from 921 respondents (56% women). The mean age of the participants was 54.4 years (56.9 years for men and 42.5 years for women). The characteristics of the population sample are presented in Table 3.
Table 3. The characteristics of the population samples (n=number of respondents)

<table>
<thead>
<tr>
<th></th>
<th>E&amp;W (n=1001)</th>
<th>Netherlands (n=1075)</th>
<th>Spain (n=963)</th>
<th>Sweden (n=1945)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of men (age in years)</td>
<td>43 (57)</td>
<td>49 (47)</td>
<td>57 (46)</td>
<td>48 (50)</td>
</tr>
<tr>
<td>% of women (age in years)</td>
<td>57 (43)</td>
<td>51 (47)</td>
<td>43 (46)</td>
<td>52 (49)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>54</td>
<td>47</td>
<td>46</td>
<td>50</td>
</tr>
<tr>
<td>Range</td>
<td>20–90</td>
<td>18-84</td>
<td>18-91</td>
<td>18-85</td>
</tr>
<tr>
<td>% of respondents aged</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–39</td>
<td>17</td>
<td>37</td>
<td>42</td>
<td>33</td>
</tr>
<tr>
<td>40–59</td>
<td>46</td>
<td>38</td>
<td>31</td>
<td>37</td>
</tr>
<tr>
<td>60+</td>
<td>37</td>
<td>25</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>**Employment status/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>main activity (%)</td>
<td>In paid employment</td>
<td>48</td>
<td>56</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Retired</td>
<td>33</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Housework/childcare</td>
<td>10</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>**Personal situation (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td>17</td>
<td>21</td>
<td>Not available</td>
<td>22</td>
</tr>
<tr>
<td>Living with children</td>
<td>34</td>
<td>28</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td><strong>Assistance with daily activities (%)</strong></td>
<td>12</td>
<td>11</td>
<td>Not available</td>
<td>6</td>
</tr>
<tr>
<td>(% requiring assistance)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% reporting any problems on each dimension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EQ-5D dimensions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>19</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Mobility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-care</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Usual activities</td>
<td>23</td>
<td>20</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Pain/discomfort</td>
<td>47</td>
<td>37</td>
<td>37</td>
<td>46</td>
</tr>
<tr>
<td>Anxiety/depression</td>
<td>29</td>
<td>15</td>
<td>24</td>
<td>30</td>
</tr>
<tr>
<td><strong>EQ-5D visual analogue scale rating of health state</strong></td>
<td>79</td>
<td>77</td>
<td>73</td>
<td>80</td>
</tr>
<tr>
<td>Mean</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Netherlands (paper II)

Survey
The data for the Dutch study were sampled from the TNS NIPO (The Dutch Institute for Public Opinion and Market Research) database which is a panel of over 200,000 respondents interviewed regularly through the Internet. In July 2005, a nationally representative sample of the Dutch population aged over 18 was drawn. Of the invited 1400 respondents of the panel, 1075 responded (response rate – 77%).

Population sample
Complete QoL-AGHDA data were available from 1038 individuals (mean age 47.2 years, SD 16.1), with almost equal gender distribution. The mean age (years) of men (n=510) was 47.3 (SD 16.5) and of women (n=528) 47.2 (SD 14.9)

The details are shown in Table 3.

Spain (paper II)

Survey
The Spanish normative values for QoL-AGHDA published in 1998 by Badia et al (12) were used in our study. A survey was performed in a random sample (n=1930) of the Barcelona general population that matched the Spanish population as reported in the 1991 Spanish census for age and gender. A total of 963 individuals agreed to participate (response rate almost 50%); however, the data for QoL-AGHDA were available for 868 individuals (57% women). The questionnaires were partly administered by trained interviewers and partly self-completed by respondents.

Population sample
The mean age (years) of the total sample was 46.3 (SD 18.2) [men, 46.2 (SD 18.3) and women 46.5 (SD 18.1)]. Other features of the Spanish sample are shown in Table 3.

Sweden (papers II and III)

Survey
The questionnaire was sent out in 2004 by the Swedish National Statistics Office (Statistics Sweden – SCB) to a random sample (n=3005) drawn from the population aged 18–85 years and permanently registered in Sweden (RTB). Two reminder letters were sent out to non-responders. In total, 1945 responses were received (65% response rate).

Comparison between responders and non-responders showed that there were more women than men in the former group and a greater percentage of responders were married rather than unmarried, whereas the opposite was
true for non-responders. A greater proportion of non-responders were immigrants i.e. not born in Sweden, or were not Swedish citizens. All of the questionnaires that were returned by responders were scanned for mistakes, and information on birth year and gender was matched with data from the RTB. Only valid answers were included in the analysis. The sampling method accounted for the observed differences between responders and non-responders.

Population sample
In total, 1945 responses were received (65% response rate). A QoL-AGHDA score was available for 1752 respondents. The gender distribution was almost equal. The mean age (years) of respondents was 49.5 (SD 17.4), with men being on average slightly older [49.8 (SD 17.0)] than women [49.2 (SD 17.7)]. Table 3 presents more characteristics.

Patients with GHD

KIMS database
KIMS (Pfizer International Metabolic Database), an international pharmacoepidemiological survey (83), was launched in 1994 at the request of endocrinologists and healthcare decision makers to monitor the outcomes and safety of long-term GH replacement therapy (Genotropin®) in hypopituitary adults with GHD being treated in a conventional clinical setting. The study to date contains data on more than 12,000 patients from 31 countries. The other aims of KIMS are to improve understanding of the consequences of GHD in adult hypopituitarism and to contribute to optimization of GH replacement.

KIMS is run according to the Survey Guidelines, the standardized guidance to all personnel involved in the clinics and at Pfizer (both nationally and globally). Data are collected electronically or through case report forms or questionnaires completed by the investigators or patients, respectively, and monitored at both national and central levels (in the Stockholm centralized database).

According to the “By-laws of KIGS and KIMS” patient data belong to the participating investigators and the database is scientifically governed at different levels (national and global) by advisory boards which consist of representatives of the KIMS investigators. Pfizer owns the operating systems and provides financial support (83).

Patients with GHD (papers I, II and IV)
The subsets of patient data from respective countries were retrieved from the KIMS database for each study separately; from E&W in all three papers and the Dutch, Spanish and Swedish data only in paper II.
All patients had severe GHD confirmed by appropriate GH stimulation tests and had not received GH replacement for at least 6 months prior to entry into the database. The main patient characteristics are presented in Table 4. In each country, the majority of patients had multiple pituitary hormone deficiencies. The most common combination was GH and three additional pituitary hormone deficits (in 37% of patients from E&W, in 46% from The Netherlands and 47% and 41.5% for the Spanish and Swedish patients, respectively).

**Table 4. Clinical characteristics of patients in KIMS (n=number of patients).**

<table>
<thead>
<tr>
<th></th>
<th>E &amp; W n=758</th>
<th>Netherlands n=247</th>
<th>Spain n=197</th>
<th>Sweden n=484</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender (n; %)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>363 (48)</td>
<td>123 (50)</td>
<td>75 (38)</td>
<td>247 (51)</td>
</tr>
<tr>
<td>Women</td>
<td>395 (52)</td>
<td>124 (50)</td>
<td>122 (62)</td>
<td>237 (49)</td>
</tr>
<tr>
<td><strong>Mean age at entry into KIMS (years; SD)</strong></td>
<td>48 (12.6)</td>
<td>48 (13.3)</td>
<td>45 (11.1)</td>
<td>51 (13.0)</td>
</tr>
<tr>
<td><strong>Disease onset (n; %)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childhood-onset</td>
<td>54 (7)</td>
<td>25 (10)</td>
<td>21 (11)</td>
<td>31 (6)</td>
</tr>
<tr>
<td>Adult-onset</td>
<td>704 (93)</td>
<td>222 (90)</td>
<td>176 (89)</td>
<td>453 (94)</td>
</tr>
<tr>
<td><strong>Primary cause of GHD * (n; %)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pituitary adenoma</td>
<td>510 (67)</td>
<td>156 (63)</td>
<td>81 (41)</td>
<td>314 (65)</td>
</tr>
<tr>
<td>Craniopharyngioma</td>
<td>62 (8)</td>
<td>22 (9)</td>
<td>19 (9.5)</td>
<td>33 (7)</td>
</tr>
<tr>
<td>Other pituitary/hypothalamic tumours</td>
<td>44 (6)</td>
<td>13 (5)</td>
<td>9 (5)</td>
<td>21 (4)</td>
</tr>
<tr>
<td>Cranial tumours distant from pituitary/hypothalamus</td>
<td>24 (3)</td>
<td>7 (3)</td>
<td>4 (2)</td>
<td>13 (3)</td>
</tr>
<tr>
<td>Treatment for malignancy outside cranium</td>
<td>10 (1.5)</td>
<td>5 (2)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other causes of acquired GHD</td>
<td>85 (11.5)</td>
<td>35 (14)</td>
<td>69 (35)</td>
<td>65 (13)</td>
</tr>
<tr>
<td>Idiopathic GHD</td>
<td>23 (3)</td>
<td>9 (4)</td>
<td>15 (7.5)</td>
<td>37 (8)</td>
</tr>
<tr>
<td><strong>Extent of hypopituitarism (n; %)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panhypopituitarism</td>
<td>133 (17)</td>
<td>40 (16)</td>
<td>44 (22)</td>
<td>61 (13)</td>
</tr>
<tr>
<td>Isolated GHD</td>
<td>77 (10)</td>
<td>27 (11)</td>
<td>1 (0.5)</td>
<td>46 (10)</td>
</tr>
<tr>
<td>**Irradiation (n; %) **</td>
<td>396 (52)</td>
<td>109 (45)</td>
<td>49 (25)</td>
<td>175 (36)</td>
</tr>
</tbody>
</table>

* according to the KIMS classification list (appendix A)

** either alone or in combination with surgery


All patients had received appropriate hormonal replacement for other pituitary hormone deficits. Surgery either alone or in combination with radiother-
apy was most commonly performed in patients from E&W (73%), The Netherlands (72%), followed by the Swedish (66%) and Spanish (57%) cohorts. The most common co-morbidities reported at entry into KIMS were hypertension, arthritis and diabetes mellitus. Overall, 26% of patients had a history of fractures.

Differences between countries (paper II)
Owing to the nature of KIMS, which is an open, observational, non-interventional study, with ongoing enrolment of new patients, the duration of follow-up in this study varied between the countries, from a maximum of 4 years in Spain to 8 years in Sweden. The Dutch patients were followed for up to 6 years and the patients from E&W for up to 7 years. As a result, the number of patients at each time point decreased. QoL-AGHDA scores were obtained at baseline, and then annually (Figure 6). The mean duration of patient follow-up (years) was 3.2 for patients from E&W, 3.4 for the Dutch cohort, 2.5 for the Spanish and 4.5 for the Swedish.

Differences between patient cohorts from E&W (papers I, II and IV)
The patient data for each of these studies were prepared at different points in time and therefore the numbers of patients included differed; even so, they met similar inclusion criteria, for example the patient numbers in E&W were 836, 758 and 894 in papers I, II and IV, respectively. Additionally, patients in paper I were assessed only at baseline, whereas those included in the two other studies had at least one follow-up visit reported to the database. Despite the variation in patient numbers, the demographic and clinical characteristics did not demonstrate substantial discrepancies between the groups (Table 4).

Handling of missing data
The QoL-AGHDA score was not computed when data were missing on one or more items. Whenever the analysis required a full dataset from the whole, constructed for these studies questionnaire i.e. QoL-AGHDA, EQ-5D, both the descriptive part and EQ-5DVAS, only individuals with complete data were analysed; otherwise those with full relevant information were included. This approach resulted in slight variations in the number of participants for certain analyses.
Country-specific differences in QoL-AGHDA scores between general population and KIMS patients during GH replacement.

The studies with patient follow-up either treated missing data at certain time points as missing at random (paper II), or used the last observation carried forward (LOCF) technique (paper IV). In paper II subsets of patients from E&W (77 patients) and Sweden (121 patients) with 5-year longitudinal follow-up and complete (non-missing) QoL-AGHDA data at each time point were prepared for analysis.

Analytical procedures

Deriving population normative data for QoL-AGHDA scores (papers I, II and III)
The population normative data for QoL-AGHDA were developed in detail for E&W (paper I) and for Sweden (paper III), whereas for the Netherlands and Spain (paper II) only population means standardized for gender and an age of 50 years were estimated. For E&W and Sweden the results obtained from the postal surveys were weighted to reflect the general population age and gender profiles. Weighted responses were used to compute mean values for QoL-AGHDA scores, relevant demographics and EQ-5D values. Statistical tests and analyses were based on unweighted responses.

The raw QoL-AGHDA scores, problem rates within EQ-5D dimensions and self-rated EQ-5D VAS were treated as cardinal variables. Analysis of variance (ANOVA) and t-tests were used to examine subgroup differences. Data are given as means ± SD unless otherwise indicated.

Modelling to assess QoL impairment and evaluate response to treatment for both general scores and dimensions (paper II)
For each country, and for each follow-up visit in a cross-sectional analysis, regression models were fitted to estimate the differences between the general population and the KIMS patient population in:

1. mean total QoL-AGHDA score
2. mean scores within the dimensions of QoL-AGHDA

Total QoL-AGHDA score

Differences between the general populations and KIMS patients
As the first step, a linear regression model was developed using total QoL-AGHDA scores for the general population and patients at entry into KIMS. The independent variables in the model were:

1. type of a study cohort (general population = 1, KIMS patients = 2)
2. age at visit (in years)
3. gender (0=male, 1=female)
4. interaction terms between these variables

A significance level was set at 10% for entry to the regression model. The final model was:

\[ \text{Mean QoL-AGHDA score} = f(\text{pop, age, gender, pop*gender}). \]

Age, gender and the interaction term can be perceived as adjustment factors.

For each country and for each yearly visit the differences in mean QoL-AGHDA scores between the general populations and KIMS patients were estimated using this model. The estimated differences were relatively constant over age (for interaction p>0.10), despite the fact that the absolute mean QoL-AGHDA scores differed by age. Results were averaged for gender.

**Analyses of patient subgroups**

Analyses were also performed to identify any specificity within the patient subgroups. The pattern of response was checked for gender and age but also for disease onset (childhood vs. adult-onset of pituitary disorder) and aetiology (NFPA vs. craniopharyngioma). These analyses were performed for each country separately as well as for the pooled data.

Special regression analyses were performed to check whether the calendar-year of entry into the study or exit from follow-up might influence the results. The underlying hypothesis was that the inclusion criteria associated with the studied outcome might have changed over time and that those patients who exited might not have responded in a similar way as those who continued treatment.

For estimating the yearly change in a longitudinal analysis of mean QoL-AGHDA scores in the patient population, repeated measurement regression was applied. Adjustment was made for age and gender.

**QoL-AGHDA scores within dimensions**

**Clustering**

Potential dimensions for clustering individual QoL-AGHDA items were indicated based on impairments in QoL that are specific for GHD. Responses to individual items within the QoL-AGHDA were coded as 0/1 for no/yes, respectively. Each of the 5 visual analogue scales, defined and collected in the parallel study, was separately designated as the dependent variable in regression analysis taking the expected set of items as the independent variables. This approach yielded dimensions for all QoL-AGHDA items, as follows: problems with memory and concentration, tiredness, tenseness, social isolation and problems with self-confidence (Table 5).
Table 5 QoL-AGHDA items categorized in 5 dimensions

<table>
<thead>
<tr>
<th>Dimension</th>
<th>QoL-AGHDA item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problems with memory and concentration</td>
<td>I have to struggle to finish jobs</td>
</tr>
<tr>
<td></td>
<td>I have to read things several times before they sink in</td>
</tr>
<tr>
<td></td>
<td>I often lose track of what I want to say</td>
</tr>
<tr>
<td></td>
<td>I often forget what people have said to me</td>
</tr>
<tr>
<td></td>
<td>I find it difficult to plan ahead</td>
</tr>
<tr>
<td></td>
<td>My memory lets me down</td>
</tr>
<tr>
<td>Tiredness</td>
<td>I feel a strong need to sleep during the day</td>
</tr>
<tr>
<td></td>
<td>It takes a lot of effort for me to do simple tasks</td>
</tr>
<tr>
<td></td>
<td>I have to push myself to do things</td>
</tr>
<tr>
<td></td>
<td>I feel worn out even when I’ve not done anything</td>
</tr>
<tr>
<td></td>
<td>I often feel too tired to do the things I ought to do</td>
</tr>
<tr>
<td></td>
<td>I have to force myself to do all the things that need doing</td>
</tr>
<tr>
<td></td>
<td>I often have to force myself to stay awake</td>
</tr>
<tr>
<td>Tenseness</td>
<td>I have difficulty controlling my emotions</td>
</tr>
<tr>
<td></td>
<td>I often feel very tense</td>
</tr>
<tr>
<td></td>
<td>There are times when I feel very low</td>
</tr>
<tr>
<td>Social isolation</td>
<td>I often feel lonely even when I am with other people</td>
</tr>
<tr>
<td></td>
<td>It is difficult for me to make friends</td>
</tr>
<tr>
<td></td>
<td>I find it hard to mix with people</td>
</tr>
<tr>
<td></td>
<td>I avoid mixing with people I don’t know well</td>
</tr>
<tr>
<td></td>
<td>I am easily irritated by other people*</td>
</tr>
<tr>
<td>Problems with self-confidence</td>
<td>I lack confidence</td>
</tr>
<tr>
<td></td>
<td>I feel as if I let people down</td>
</tr>
<tr>
<td></td>
<td>I avoid responsibilities if possible</td>
</tr>
<tr>
<td></td>
<td>I feel as if I’m a burden to people</td>
</tr>
</tbody>
</table>

* The item: “I am easily irritated by other people” fits marginally better into social isolation, although it could also be categorized within the dimension tenseness.


**Differences between the general populations and KIMS patients**

As the within-patient dimension scores were expected to be correlated, mixed-linear regression was used to analyse the mean difference by QoL-AGHDA dimensions between the general population and KIMS patients.
The number of items differed by dimension, so that standardization was necessary and this was achieved by computing the percentage of items within a dimension that a subject expressed problems with. Adjustments were made for age and sex. Analyses were performed at yearly visits (cross-sectionally).

As these statistics for dimensions had relatively lower precision and lower stability – especially at the later stages when there were few patients – the presented time-series was shown for rather fewer years of follow-up compared with the overall QoL-AGHDA score.

**Deriving QoL-AGHDA utility for Sweden and England & Wales (papers III and IV)**

As the first step to derive utilities for QoL-AGHDA, the EQ-5D<sub>index</sub> was computed. Preference-based weights for the health states described by the EQ-5D profile do not exist for the Swedish population and therefore, in paper III it was decided to use the European weights as published by Greiner and colleagues (81). For the E&W population (paper IV) the weights available for the UK population (57) were used.

The European set of weights was estimated using a multi-level regression model based on EQ-5D<sub>VAS</sub> evaluations in the pooled data from 11 studies (Finland, Germany, the Netherlands, Spain, Sweden and the UK) that in total encompassed almost 83,000 VAS evaluations on forty-four EQ-5D health states elicited from 6,870 respondents. The EQ-5D<sub>VAS</sub> scale used in these studies was anchored as the best and the worst imaginable health, and hence it did not comply with the definition of QALY, where the index must be anchored as ‘full health’ = 1 and ‘death’ = 0. Therefore the European values were rescaled, using median values for death as recommended (81). The aggregated coefficients from the European pooled data were applied to the present Swedish data. In this way, estimates of the EQ-5D<sub>index</sub> for the health states reported in the Swedish sample were achieved.

A similar approach was taken for computing the EQ-5D<sub>index</sub> in the E&W cohort (paper IV), where the coefficients published by Dolan (56) were employed. The difference, however, was that Dolan used TTO, rather than a VAS scale and therefore rescaling for death was not necessary. Nevertheless, both methods used general population data for hypothetical health states.

Subsequently, in multiple regression analyses (ordinary least squares) EQ-5D<sub>index</sub> was used as the dependent variable. For the Swedish population (paper III), a simple model was used with QoL-AGHDA score together with age and gender as independent variables, whereas for E&W (paper IV), responses for each QoL-AGHDA item (entered as dummy variables) were used together with age. The decision to use different profiles of QoL-AGHDA scores was dictated by the characteristics of the original data sets.
which resulted in diverse model fits (using single responses in the E&W sample yielded a slightly higher $R^2$).

A full model for the Swedish sample was also developed using demographic variables together with QoL-AGHDA scores as independent variables. The demographic variables were age and sex, and the dummy variables: Swedish citizenship, unmarried, divorced, widow/widower, more than 13 years of education, paid employment and living alone.

The simple model with only age as a covariate was used for the E&W population because gender in this set of data did not appear as a significant predictor.

The obtained coefficients for the Swedish data were internally validated using Jack-knife and bootstrap methods. To check the external validity, the computed QoL-AGHDA-based utilities were compared with the general health state rates using ANOVA and, additionally, were correlated with the EQ-5Dv.as. Patients’ estimated QoL-AGHDA$_{utility}$ scores based on the E&W model were analyzed for correlation with PGWB scores.

The E&W model demonstrated an adjusted $R^2$ of 0.42. Each regression coefficient, $b_i$, represented the utility weight for the corresponding QoL-AGHDA item and when aggregated across all 25 items this yielded an estimate of the utility-weighted QoL-AGHDA referred to as QoL-AGHDA$_{utility}$. The Swedish full model reached an adjusted $R^2$ of 0.38 and the simple model an adjusted $R^2$ of 0.36.

**Computation of QoL-AGHDA$_{utility}$ at baseline, total gain and gain per year and comparison with general population values (paper IV)**

Patients had been followed in the KIMS database for a varying number of years. The total QoL-AGHDA$_{utility}$ for each patient was calculated using the trapezoid (Figure 5) formula as follows:

$$\Sigma (u_{i-1} - 2*u_0 + u_i)/2 \quad (i=1, t)$$

where $t =$ total duration of patient follow-up in KIMS

and $u_i =$ QoL-AGHDA$_{utility}$ at year $i$.

The average change in QALYs over the entire time period was also computed as gain/year. The calculation was performed conservatively assuming that QoL in untreated patients, as measured by the QoL-AGHDA, would remain the same as at baseline. The patient QALY deficit was calculated as the difference between the QoL-AGHDA$_{utility}$ observed in patients and the corresponding value computed for age- and gender-matched individuals in the general population sample.
A high QoL-AGHDA_{utility} score denotes a better QoL, which is contrary to the interpretation of a QoL-AGHDA raw score, where a high value indicates a poor QoL.

**Patient subgroups**
Finally, QoL-AGHDA_{utility} at baseline and following GH treatment was evaluated with respect to age, gender, primary aetiology, onset of pituitary disease (childhood vs. adulthood), extent of hypopituitarism and medical history.
Results

Population normative values for total QoL-AGHDA scores (papers I, II and III)

The set of population-based, age- and gender-specific reference values for QoL-AGHDA total scores was derived for E&W and Sweden. The weighted mean score was higher in E&W (6.7 SD 5.8) than in Sweden (3.9 SD 4.8). This effect was the same in men 6.2 (SD 5.6) vs. 3.6 (SD 4.7) and women 7.1 (SD 6.1) vs. 4.3 (SD 5.0). Despite this, in both countries women on average scored numerically higher (indicating poorer QoL) than men, although the difference was significant only in the Swedish population (p<0.003). The median values in the E&W population were 4.0 in men and 5.0 in women (in the pooled data 5.0) whereas in Sweden they were the same in both genders (2.0).

In the E&W populations age did not explain variability in QoL-AGHDA scores, contrary to the Swedish data where QoL measured by QoL-AGHDA improved significantly with age (r= – 0.073, p<0.003). QoL-AGHDA scores categorized by gender and age are presented in Table 6.

The standardized means for gender and age of 50 years for the Dutch and Spanish cohort were very similar 4.9 (SD 4.9) and 5.0 (SD 4.6), respectively.

Nearly 10% of the general population scored above the patient mean on the QoL-AGHDA. Analysis of respondents in this subgroup revealed that they belonged to the 40–59 year age-group, were twice as likely to suffer from a chronic condition, and reported a higher level of problems on any of the 5 EQ-5D dimensions. Their self-rated health state was worse and they more often needed help with their daily activities than those who scored below the patient mean QoL-AGHDA score.
Table 6. Age- and gender-specific total QoL-AGHDA scores in the English & Welsh and Swedish general populations

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age group</th>
<th>England &amp; Wales</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;29*</td>
<td>12</td>
<td>5.8</td>
<td>4.3</td>
</tr>
<tr>
<td>30–39</td>
<td>38</td>
<td>6.9</td>
<td>5.7</td>
</tr>
<tr>
<td>40–49</td>
<td>78</td>
<td>5.8</td>
<td>6.0</td>
</tr>
<tr>
<td>50–59</td>
<td>98</td>
<td>6.7</td>
<td>6.1</td>
</tr>
<tr>
<td>60–69</td>
<td>106</td>
<td>5.6</td>
<td>5.0</td>
</tr>
<tr>
<td>70–79</td>
<td>55</td>
<td>6.9</td>
<td>5.2</td>
</tr>
<tr>
<td>&gt;80**</td>
<td>16</td>
<td>6.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Total</td>
<td>403</td>
<td>6.2</td>
<td>5.6</td>
</tr>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;29*</td>
<td>25</td>
<td>6.6</td>
<td>5.7</td>
</tr>
<tr>
<td>30–39</td>
<td>82</td>
<td>6.8</td>
<td>5.7</td>
</tr>
<tr>
<td>40–49</td>
<td>124</td>
<td>7.4</td>
<td>6.7</td>
</tr>
<tr>
<td>50–59</td>
<td>135</td>
<td>6.7</td>
<td>6.0</td>
</tr>
<tr>
<td>60–69</td>
<td>94</td>
<td>6.5</td>
<td>5.5</td>
</tr>
<tr>
<td>70–79</td>
<td>45</td>
<td>7.7</td>
<td>6.8</td>
</tr>
<tr>
<td>&gt;80**</td>
<td>13</td>
<td>6.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Total</td>
<td>518</td>
<td>7.1</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>921</td>
<td>6.7</td>
<td>5.8</td>
</tr>
</tbody>
</table>

* The lower cut-off point for age was 20 years in E&W and 18 years in Sweden
** The upper cut-off point for age was 90 years in E&W and 85 years in Sweden

The effects of GH replacement therapy on QoL in relation to normative values (papers I and II)

Comparison with untreated patients in England & Wales (paper I)

The total QoL-AGHDA scores in E&W patients before they started GH treatment were evaluated in detail in relation to normative values. Both male and female patients in each age category demonstrated a significant reduction in QoL (p<0.001 in most categories). When controlling for age, younger patients with CO-GHD reported better QoL than those with a disease of
adult onset. Patients who reported one or more co-morbidities recorded worse QoL than those without additional health problems [mean QoL-AGHDA scores 14.6 (SD 6.3); 16.0 (SD 6.3); 13.9 (SD 6.6)]. A similar tendency was observed in the general population sample. Individuals with no other chronic disease had better QoL with lower QoL-AGHDA scores (mean: 4.9, SD 5.1) than those with only one (mean: 6.2, SD 5.6) or more (mean: 8.9, SD 8.0) reported long-lasting conditions. A similar tendency was found when the mean population QoL-AGHDA scores were computed for each self-rated health state (excellent, very good, good, fair and poor), showing increasing QoL-AGHDA scores as self-rated health status worsened (ANOVA: F = 88.7, d.f. = 4, p < 0.001). Furthermore, a strong negative correlation was observed between EQ-5D VAS and QoL-AGHDA scores (r = –0.53, p<0.001), indicating that the fewer the number of problems recognized by the QoL-AGHDA (a lower QoL-AGHDA score), the better the general health status reported on the EQ-5D VAS (a higher EQ-5D VAS score).

Treatment effects – four-country analysis (paper II)
The demographic and clinical characteristics of patients in each country are displayed in Table 4.

Total QoL-AGHDA score
All patients, regardless of country of origin showed, attenuated QoL in relation to the respective national normative data, as shown by significantly higher total QoL-AGHDA scores. The greatest deviation was observed in patients from E&W, followed by the Spanish, Dutch and Swedish patients. These differences were diminished by GH replacement over the follow up and values returned towards the respective general population means (Figure 6). The patterns of such improvement were notably similar, independently of magnitude of the primary QoL reduction. The most dramatic improvement occurred during the first year of treatment. Longitudinal analysis of change (trend) over time in the pooled as well as in country-specific data confirmed these findings. Although the estimates of change per year between the first year and the subsequent third, fifth and seventh year were of similar magnitude, the significant improvement observed during the first year of treatment was maintained. Hence, the cumulative effect over time brought patients’ QoL to the level presented by the general population.

Patient subgroups
Gender-specific cohorts analysed separately for each country demonstrated similar patterns. The gender differences estimated in the pooled data reached statistical significance in the patient cohort (mean QoL-AGHDA score in men was 11.0; 95% CI 10.6–11.4; and in women 13.2 95% CI 12.6–13.6 p < 0.0001), which was not the case in the general population [mean (95% CI),
5.3 (5.1–5.6) and 5.3 (5.0–5.5), respectively p=NS). Moreover, before GH treatment the difference between adults with GHD and the general population was greater in women than in men. This variation disappeared during follow-up, as women responded better to treatment than men and at later visits the distance to normative values was similar in both genders.

Interestingly, the disease onset (childhood- vs. adult-onset) or primary aetiology (craniopharyngioma vs. NFPA) did not impact on the beneficial pattern of response to GH. In order to detect possible confounders the same modelling was performed in patients who continued follow-up in KIMS vs. those who did not, as well as in the series of patients from E&W and Sweden in whom longitudinal QoL-AGHDA data with no missing values at any time point, were available. Both analyses yielded results consistent with those obtained previously. Additionally, no statistically significant influence of the calendar-year of entry was shown on the results, with the exception of the Swedish patients entering KIMS between 1998 and 2000, in whom QoL appeared to improve less rapidly during the first year.

**QoL dimensions measured by QoL-AGHDA**

Before commencing GH, all QoL dimensions measured by QoL-AGHDA were significantly compromised in comparison with the respective population data. Again, striking similarities between the countries were apparent, as the profiles of dysfunction and patterns of recovery were the same throughout all studied cohorts. Problems with memory and tiredness were greater than those reported in other dimensions before GH treatment started. Despite the significant improvement during the first year of therapy, by the end of follow-up, memory and tiredness remained reduced compared with the general population data. Social isolation was the first dimension to reach the population values, followed by tenseness and problems with self-confidence. There was no gender-specificity in this pattern of response and it is worth emphasising that this pattern was consistent in all countries (Figure 7).

These findings rank the dimensions by severity of impairment (memory and concentration, tiredness, self-confidence, tenseness and social isolation), which at least is partly concurrent with the biological functions of GH. The improvement, however, occurred in the reverse order. It may be hypothesized that the response to treatment is driven inversely by the degree of dysfunction (Figure 7).

Preference-weighted index based on the QoL-AGHDA – QoL-AGHDA utility (papers III and IV)

The Swedish population (paper III)

In the full model the following variables negatively influenced the EQ-5D\textsubscript{index} in a significant way: age, age\textsuperscript{3}, QoL-AGHDA score and being divorced, whereas age\textsuperscript{2}, QoL-AGHDA score\textsuperscript{2}, more than 13 years of education, paid employment and Swedish citizenship had a positive impact. The R\textsuperscript{2} value was 0.38.

The simple model included age and gender, yielding the transformation algorithm:

\[
\text{QoL-AGHDA utility} \rightarrow \text{ED-5D}_{\text{index}} = 1.05 - 0.0189*\text{QoL-AGHDA score} - 0.00238 * \text{age} - 0.0127*\text{gender} \\
\text{(men = 0; women = 1)}
\]
Table 7  Simple model regression estimates and 95% confidence intervals used in the computation of QoL-AGHDA-based utilities using EQ-5D<sub>index</sub> as the dependent variable (R<sup>2</sup> = 0.36),

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>QoL-AGHDA score</th>
<th>Age</th>
<th>Male/female (0/1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.0472</td>
<td>-0.0189</td>
<td>-0.00238</td>
</tr>
<tr>
<td>Lower 95% limit</td>
<td>1.0264</td>
<td>-0.0202</td>
<td>-0.00275</td>
</tr>
<tr>
<td>Upper 95% limit</td>
<td>1.0681</td>
<td>-0.0176</td>
<td>-0.00202</td>
</tr>
</tbody>
</table>

From Kołtowska-Häggström 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) 2007 Value in Health 10(1):73-81 by Blackwell Publishing Ltd. Copyright © 2007, Published on behalf of the International Society for Pharmacoeconomics and Outcomes Research Reproduced by permission

Table 8. Mean (SD) gender- and age-specific QoL-AGHDA<sub>utility</sub> for the English & Welsh and Swedish populations.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>England and Wales</th>
<th>Sweden</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>men</td>
<td>women</td>
<td>men</td>
</tr>
<tr>
<td>18–29</td>
<td>0.90 (0.10)</td>
<td>0.89 (0.12)</td>
<td>0.91 (0.08)</td>
</tr>
<tr>
<td></td>
<td>(n=12)</td>
<td>(n=25)</td>
<td>(n=134)</td>
</tr>
<tr>
<td>30–39</td>
<td>0.86 (0.13)</td>
<td>0.87 (0.12)</td>
<td>0.90 (0.09)</td>
</tr>
<tr>
<td></td>
<td>(n=38)</td>
<td>(n=82)</td>
<td>(n=130)</td>
</tr>
<tr>
<td>40–49</td>
<td>0.87 (0.13)</td>
<td>0.84 (0.13)</td>
<td>0.87 (0.09)</td>
</tr>
<tr>
<td></td>
<td>(n=78)</td>
<td>(n=124)</td>
<td>(n=147)</td>
</tr>
<tr>
<td>50–59</td>
<td>0.83 (0.13)</td>
<td>0.83 (0.14)</td>
<td>0.86 (0.09)</td>
</tr>
<tr>
<td></td>
<td>(n=98)</td>
<td>(n=135)</td>
<td>(n=181)</td>
</tr>
<tr>
<td>60–69</td>
<td>0.82 (0.10)</td>
<td>0.81 (0.13)</td>
<td>0.83 (0.09)</td>
</tr>
<tr>
<td></td>
<td>(n=106)</td>
<td>(n=94)</td>
<td>(n=157)</td>
</tr>
<tr>
<td>70–79</td>
<td>0.77 (0.12)</td>
<td>0.75 (0.18)</td>
<td>0.80 (0.08)</td>
</tr>
<tr>
<td></td>
<td>(n=55)</td>
<td>(n=45)</td>
<td>(n=85)</td>
</tr>
<tr>
<td>80–85</td>
<td>0.76 (0.13)</td>
<td>0.75 (0.13)</td>
<td>0.76 (0.11)</td>
</tr>
<tr>
<td></td>
<td>(n=16)</td>
<td>(n=13)</td>
<td>(n=27)</td>
</tr>
<tr>
<td>Total</td>
<td>0.83 (0.13)</td>
<td>0.83 (0.14)</td>
<td>0.86 (0.10)</td>
</tr>
<tr>
<td></td>
<td>(n=403)</td>
<td>(n=518)</td>
<td>(n=861)</td>
</tr>
</tbody>
</table>

The parameter estimates and 95% confidence intervals for the estimates are presented in Table 7. The R<sup>2</sup> value reached 0.36 and was slightly lower than that for the full model.
The mean of the weighted estimate (QoL-AGHDA<sub>utility</sub>) for the total population (n=1752) was 0.85 (SD 0.164). Again, the estimate for men (n=861; mean 0.86; SD 0.10) was higher (p<0.001) than for women (n=891; mean 0.84; SD 0.10). Age- and gender-specific estimates are shown in Table 8.

Jack-knife and bootstrap analyses, performed for internal validation, yielded very similar results as the regression analysis.

QoL-AGHDA<sub>utility</sub> declined significantly as self-rated health state deteriorated and correlated significantly with EQ-5D<sub>vas</sub> (r=0.60, p<0.001).

The English and Welsh population (paper IV)

A similar technique with small modifications as described in the method section was employed to construct QoL-AGHDA<sub>utility</sub> for the E&W population. The following transformation algorithm was used:

\[
\text{QoL-AGHDA}_{\text{utility}} \rightarrow \text{ED-5D}_{\text{index}} = b_0 + c*\text{age} + \sum b_i x_i + e_i,
\]

where \(x_{i, i=1-25}\) correspond to the 25 dichotomous items (coded as 0 – no or 1 – yes) that are summed to form the QoL-AGHDA score, \(b_i\) is the regression coefficient estimates, and \(e_i\) corresponds to error terms.

The model demonstrated an adjusted \(R^2\) of 0.42.

The mean QoL-AGHDA<sub>utility</sub> score for the total sample was 0.83 (SD 0.136). Contrary to the results for the Swedish data (paper III), men [0.83 (SD 0.127)], and women [0.83 (SD 0.141)] scored similarly.

The effects of GH as measured by QoL-AGHDA<sub>utility</sub> compared with population values (paper IV)

Patient characteristics

The patients (n=894; 47% men) originated from E&W and the mean follow up was 3.4 (SD 1.74) years with a range of 1 – 6 years. Most of the patients had developed their disease during adulthood, and only 21.6% had CO-GHD; the mean age at diagnosis of GHD was 40 years (SD 16.5) and at entry into KIMS 45 years (SD 14.3). Men were slightly older than women at both time points: 41 (SD 17.1) vs. 40 (SD 15.9) years at diagnosis and 45 (SD 14.7) vs. 44 (SD: 13.9) years at entry into KIMS. The profile of primary aetiology in this cohort is presented in Table 9. Close to 90% had multiple pituitary hormone dysfunctions with approximately one-third of the patients having GHD and three other pituitary hormones deficits, whereas 17.4% had panhypopituitarism. At entry and during the follow-up, all pituitary hormone deficits other than GH were routinely replaced.
Table 9. QoL-AGHDA utilities scores (absolute and change) at baseline and at the last reported visit by primary aetiology for hypopituitarism*. Data shown as mean (SD).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
<th>Baseline visit</th>
<th>Last reported</th>
<th>Total gain</th>
<th>Gain/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-functioning pituitary adenoma</td>
<td>201</td>
<td>22.5</td>
<td>0.64 (0.169)</td>
<td>0.76 (0.172)</td>
<td>0.36 (0.537)</td>
<td>0.1 (0.121)</td>
</tr>
<tr>
<td>Secreting pituitary adenoma</td>
<td>311</td>
<td>34.8</td>
<td>0.64 (0.165)</td>
<td>0.76 (0.166)</td>
<td>0.36 (0.592)</td>
<td>0.09 (0.124)</td>
</tr>
<tr>
<td>Other sellar tumour</td>
<td>64</td>
<td>7.2</td>
<td>0.68 (0.182)</td>
<td>0.78 (0.183)</td>
<td>0.3 (0.498)</td>
<td>0.09 (0.124)</td>
</tr>
<tr>
<td>Craniopharyngioma</td>
<td>91</td>
<td>10.2</td>
<td>0.71 (0.172)</td>
<td>0.8 (0.172)</td>
<td>0.31 (0.602)</td>
<td>0.07 (0.123)</td>
</tr>
<tr>
<td>Extra cellular tumour</td>
<td>55</td>
<td>6.2</td>
<td>0.69 (0.163)</td>
<td>0.76 (0.194)</td>
<td>0.18 (0.379)</td>
<td>0.06 (0.101)</td>
</tr>
<tr>
<td>Idiopathic GHD</td>
<td>58</td>
<td>6.5</td>
<td>0.75 (0.171)</td>
<td>0.82 (0.165)</td>
<td>0.28 (0.563)</td>
<td>0.07 (0.119)</td>
</tr>
<tr>
<td>Treatment for malignancy outside the cranium</td>
<td>20</td>
<td>2.2</td>
<td>0.72 (0.21)</td>
<td>0.82 (0.174)</td>
<td>0.08 (0.654)</td>
<td>0.07 (0.17)</td>
</tr>
<tr>
<td>Other causes of acquired GHD</td>
<td>94</td>
<td>10.5</td>
<td>0.68 (0.179)</td>
<td>0.8 (0.154)</td>
<td>0.24 (0.422)</td>
<td>0.07 (0.118)</td>
</tr>
<tr>
<td>Total</td>
<td>894</td>
<td>100</td>
<td>0.67 (0.174)</td>
<td>0.77 (0.171)</td>
<td>0.32 (0.549)</td>
<td>0.08 (0.122)</td>
</tr>
</tbody>
</table>

* according to the KIMS classification list (appendix A)


Overall, more than half of the patient cohort reported one concomitant disease and 20% more than one reported, with fractures being the most common (40%), followed by hypertension (19%), heart problems (12%), asthma and/or allergy (12%), arthrosis (10%) and diabetes mellitus (6%).

IGF-I SDS on GH replacement was lower in women than in men: 0.05 (SD 1.612) vs. 0.73 (SD 1.612) (p<0.00001), respectively, despite a higher mean maintenance GH dose (defined as the dose at the 1 year visit) administered in women [0.44 (SD 0.220) mg/day] than in men [0.37 (SD 0.185) mg/day] (p<0.0001). However, it is necessary to emphasize that in women pre-treatment mean IGF-I SDS was lower than in men [–2.26 (SD 1.782) vs. –1.40 (SD 1.915) p<0.0001]. The change in IGF-I SDS was similar in men and women which may indicate that women still received a too low GH dose.

At baseline QoL-AGHDA score was higher in women than in men [means 15.9 (SD 6.58) and 13.9 (SD 6.10), respectively], denoting poorer QoL. However, the response to GH as evaluated by a total decrease in QoL-AGHDA score over the study period and the average decrease per year was larger in women (Table 10). This finding is consistent with the observed changes in IGF-I levels both in the absence of GH and with its replacement.
QoL-AGHDA	extsubscript{utility}

The effects of GH replacement therapy were measured by QoL-AGHDA	extsubscript{utility} as a total gain and a gain per year during the study period.

The mean QoL-AGHDA	extsubscript{utility} score at baseline was 0.67 (SD 0.172). Women, as expected, had a worse QoL than men when measured by QoL-AGHDA	extsubscript{utility} [mean 0.63 (SD 0.166) vs. 0.70 (SD 0.174), p<0.001] and responded better to GH. Their total observed QALY gain was higher than in men, [0.38 (SD 0.602), vs. 0.25 (SD 0.473), p<0.001], as was the mean QoL-AGHDA	extsubscript{utility} gain per year [0.07 (SD 0.113) for men and 0.10 (SD 0.129) for women, p<0.001]. As shown in Table 10, all within-group changes were significant (p<0.001).

There were significant (p<0.0001) positive correlations between PGWB scores at baseline, last observation and the change in PGWB score and respective measures of QoL-AGHDA	extsubscript{utility}, r=0.68, r=0.68 and r=0.42, respectively.

QoL measured by QoL-AGHDA	extsubscript{utility} in patients before GH treatment commenced, differed significantly from the expected values calculated from the sample of the general population [0.67, (SD 0.174) vs. 0.85, (SD 0.038) p<0.0001]. The mean deficit in untreated patients was –0.19 (SD 0.168). There was also a significant difference in the mean QoL-AGHDA	extsubscript{utility} deficit between men –0.16 (SD 0.170) and women –0.21 (SD 0.162), (p<0.001). The main improvement occurred during the first year of GH treatment, when the QoL-AGHDA	extsubscript{utility} deficit was reduced to –0.07 (SD 0.163) (p<0.001) in the total cohort and to –0.07 (SD 0.160) (p<0.001) in men and –0.08 (SD 0.170) (p<0.001) in women. From then until the last reported visit, the QoL-AGHDA	extsubscript{utility} deficits in both genders remained indistinguishable as was observed for the QoL-AGHDA scores (paper II). Despite a dramatic improvement during the first year of observation which was maintained during the whole follow-up period, patients' QoL-AGHDA	extsubscript{utility} continued to be significantly different (p<0.001) from those reported by the general population (Line A in Figure 5).
Table 10. QoL-AGHDA scores and QoL-AGHDA_utility (absolute and change) at baseline and at last reported visit by gender and for the total cohort

<table>
<thead>
<tr>
<th></th>
<th>QoL-AGHDA score*</th>
<th>QoL-AGHDA_utility**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline visit</td>
<td>Last reported</td>
</tr>
<tr>
<td><strong>Men</strong> (n=423)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>13.9 (6.58)</td>
<td>8.7 (6.91)</td>
</tr>
<tr>
<td>Median</td>
<td>15.0</td>
<td>8.0</td>
</tr>
<tr>
<td><strong>Women</strong> (n=472)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>15.9 (6.10)</td>
<td>9.45 (6.99)</td>
</tr>
<tr>
<td>Median</td>
<td>16.5</td>
<td>9.0</td>
</tr>
<tr>
<td><strong>Total</strong> (n=895)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>14.9 (6.40)</td>
<td>9.1 (6.96)</td>
</tr>
<tr>
<td>Median</td>
<td>16.0</td>
<td>8.0</td>
</tr>
</tbody>
</table>

All changes within groups (paired t-test) were significant p<0.001, p<0.001 significance of differences between groups (independent t-test).

* A low QoL-AGHDA score indicates a good QoL, meaning that a decrease in score denotes improvement in QoL
** A high QoL-AGHDA_utility indicates a good QoL, meaning that an increase in score denotes improvement in QoL

Patient subgroups

The potential impact of several demographic and clinical features such as age, time of disease onset, primary aetiology, previous treatment, extent of hypopituitarism and co-morbidities on QoL-AGHDA utility at baseline and response to GH treatment was analyzed.

Age

QoL-AGHDA utility was negatively correlated with age both at baseline ($r = -0.23; p<0.0001$) and at the last reported visit ($r = -0.25; p<0.0001$), meaning that QoL-AGHDA utility deteriorated with advancing age (Figure 8). Despite that, as the mean total QoL-AGHDA utility gain and the mean gain per year were similar in all age groups, it can be hypothesized that GH affects patients beneficially irrespective of age.

![Figure 8. Change in QoL-AGHDA utility in GH-deficient adults during GH replacement therapy by age group. The dotted line depicts values at baseline and the continuous line – values at the last reported visit. From Koltowska-Häggström M, Kind P, Monson JP and 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 2007 Clinical Endocrinology Published article online: 4-Sep-2007 doi: 10.1111/j.1365-2265.2007.03010.x by Blackwell Publishing Ltd. Copyright © 2007, Official Journal of the Society for Endocrinology. Reproduced by permission.]

Disease onset

QoL-AGHDA utility scores were higher in patients with CO-GHD than in those with AO-GHD both at baseline [0.75 (SD 0.173) vs. 0.64, (SD 0.166),
p<0.001] and the last reported visit [0.82 (SD 0.167) vs. 0.76, (SD 0.170) p<0.001]. Predictably, both the mean total gain and gain per year were lower in patients with CO-GHD than in those with AO-GHD. The former was 0.18, (SD 0.488) in CO-GHD and 0.35, (SD 0.559) in AO-GHD and the latter 0.05, (SD 0.117) vs. 0.09, (SD 0.123), respectively (Figure 9). This observation was confirmed when it was controlled for age and gender in the multiple regression analysis.

![Graph](image)

**Figure 9.** Change in QoL-AGHDA\(_{utility}\) in GH-deficient adults during GH replacement therapy by disease onset. From Kołtowska-Häggström M, Kind P, Monson JP and 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 2007 *Clinical Endocrinology* Published article online: 4-Sep-2007 doi: 10.1111/j.1365-2265.2007.03010.x by Blackwell Publishing Ltd. Copyright © 2007, Official Journal of the Society for Endocrinology. Reproduced by permission

**Primary disease and its treatment**

As shown in Table 9, primary aetiology modified QoL-AGHDA\(_{utility}\) both at baseline and at the last reported visit. Patients in whom GHD was caused by pituitary adenoma had the lowest QoL-AGHDA\(_{utility}\) at both time points. However, like age, the primary cause of GHD had no influence on the response to treatment as measured by total QoL-AGHDA\(_{utility}\) gain and mean gain per year.

Interestingly, neither previous surgery nor irradiation had an impact on QoL-AGHDA\(_{utility}\) at any time point – and neither variable influenced response to GH.
Extent of hypopituitarism

There was no significant difference in any of the QoL-AGHDA utility parameters between the patients with isolated GHD and multiple pituitary hormone dysfunctions. Furthermore, the number of pituitary hormone deficits additional to GH showed no significant correlation with any of the QoL-AGHDA utility parameters (Figure 10).

![Figure 10](image_url)


Co-morbidities

There was a significant impact of reported co-morbidities on all QoL parameters, paralleling the findings with regard to QoL-AGHDA scores (paper I). Mean QoL-AGHDA utility scores were lower in patients who reported co-morbidities (n=513) than in patients with no additional health problems (n=381) both at baseline [0.63 (SD 0.167) vs. 0.71 (SD 0.172) p<0.001], and at the last reported visits [0.75 (SD: 0.174) vs. 0.81 (SD 0.1599 p>0.001]. At the same time, patients with co-morbidities responded better to GH treatment in terms of QoL-AGHDA utility compared to those with no co-mobidities [mean QoL-AGHDA utility total gain 0.36 (SD 0.565) vs. 0.25 (SD 0.520) p<0.002 and mean gain per year 0.10 (SD 0.124) vs. 0.07 (SD 0.119), p<0.004].
This thesis researched QoL in adult GHD in both clinical and economic settings. As a starting point for defining the deficit in QoL experienced by patients, sets of country-specific population normative data for QoL-AGHDA were derived. Subsequently, QoL deficits in patients in relation to the population normative values were found to be significant. The deficits were significant for both the total QoL-AGHDA scores and individual QoL dimensions. Long-term GH replacement resulted in sustained improvements, which brought QoL towards the normative country-specific values. For potential use in health economic evaluations, the thesis proposed models for generating a preference-weighted index – utility (QoL-AGHDA\text{utility}) based on the disease-specific measure, the QoL-AGHDA. QoL-AGHDA\text{utility} was applied to the clinical setting and found to be useful for monitoring treatment effects in patients with GHD. Moreover, this study confirmed that there was a QoL-AGHDA\text{utility} deficit before treatment and that it was reversed after starting GH replacement in different patient subgroups.

The novel aspects of this thesis are, firstly, the inclusion of normative population data derived from the same sample for both generic and disease-specific instruments; secondly, the assessment of treatment effects on QoL relative to population norms; and, thirdly, the evaluation of the response to treatment using a preference-weighted index derived from a clinically used, disease-specific measure.

The studies were done for England and Wales (E&W) (papers I, II and IV), the Netherlands (paper II), Spain (paper II) and Sweden (papers II and III), using as the main tools the QoL-AGHDA and EQ-5D. The population data were collected by separate surveys and the patient data for the respective countries were retrieved from the KIMS database.

Methodological issues

Restriction to country-specific data

Despite stringent methods applied to the translation and cultural validation of different language versions of instruments, there is substantial inter-country variation in QoL, regardless of the QoL measure employed, as has been observed previously (27, 96, 97). Obviously, it cannot be excluded that
the observed variations are due to differences between the characteristics of studied samples. However, population values are usually derived from large samples and the sampling procedures are carried out so that they mirror the population profile. Moreover, the study of Inglehart and Rabier (97) – which was based on a detailed evaluation of multi-language country data (Switzerland) in relation to other countries speaking the same language (e.g. Germany, Austria) – confirmed that these differences were stable nation-wise and not language-wise. Therefore, it can be hypothesized that the observed variation is predominately caused by the existence of true differences in mentality, culture and the perception of well-being. Different instruments, however, do not yield consistent population variations. As a rule of thumb, it may be predicted that instruments that represent constructs that are easier to define and less complex, as well as requiring lower cognitive capacities, generate more consistent results across countries (37).

Conversely, measures that tend to assess psychological dimensions are more likely to be affected by cultural differences (27). To avoid inaccuracies related to cultural heterogeneity, it was decided to conduct the present research in country-specific settings. Hence the patient data were compared to population values originating from the respective country.

Sources of data

Population samples
The critical issues for the population samples are whether they are both representative and comparable. Here, the population data were collected in a similar manner in all countries, except Spain – where they were obtained earlier and limited to one geographical region (Barcelona). The remaining samples were retrieved from the total country population by professional agencies specializing in such services and were weighted for age and gender to reflect general population demographic parameters. These data were collected over similar time periods (2003–2005), using either postal mailing (E&W and Sweden) or an Internet-based survey (the Netherlands). The other difference was that the Swedish respondents were randomly selected, whereas the Dutch and English & Welsh originated from a panel of people who had agreed to participate in health surveys. This fact might explain the higher response rates in the Dutch and E&W surveys (71% and 82%, respectively) than in the Swedish sample (65%) but at the same time it partly limits the generalizability of the results obtained in the former countries (163). The Swedish sample, though, being randomly selected is believed to mirror more closely the Swedish society at large. The results were adjusted as appropriate for age and gender to reflect demographic profiles. Overall, however, it is believed that the comparability and representativeness of the population data were adequate for this research.
**Patient data**

Patient data were retrieved from the KIMS database, which is by far the largest database of hypopituitary adults with GHD with the longest observation time (over 40,000 patient-years of follow-up). However, it is an observational study, which by definition reflects routine clinical practice. This type of study entails no rigorous inclusion and exclusion criteria or restrictions on the number of patients enrolled and duration of follow-up (26, 102). Obviously, there are certain strengths and limitations inherent in such studies. First of all, the majority of patients treated with GH were included but, as the penetration of such treatment and eligibility of patients for treatment show great variability across the countries, this is likely to result in selection bias. For example, in the UK the major criterion for GH replacement is impaired QoL and probably this is the reason why patients from E&W had a worse QoL relative to patients from other countries and most likely worse than the entire population of GH-deficient adults in the UK. On the other hand, in Sweden almost all patients with GHD routinely receive treatment, meaning that the Swedish patient population was much more representative. Finally, because of the existing contraindications for therapy, patients receiving GH therapy were less likely to have a profound impairment in general health than would be expected in the wider population of patients with hypopituitarism. These speculations suggest a potential selection bias.

The other major problems inherent in any observational study are the completeness of data and the unequal durations of follow-up. An implicit characteristic of the KIMS database is a constant enrolment of patients. Hence, different lengths of patient follow-up result, not only from drop-outs, but also from the continuous inflow of new patients. Therefore, the observation time was calculated from the date of the entry visit until that of the last reported visit for each individual patient. The problem of missing data was handled either by including only patients with complete data or by treating data as missing at random or by using the last observation carried forward technique. The latter technique was applied only if observations were missing between two reported time points.

These limitations are outweighed to certain degree by the large number and great variety of patients from a wide geographical area, as well as the prolonged period of data collection (since 1994) (3). All of these features allow additional analyses to validate the primary results – for example, checking for the impact of year of enrolment, confirming the trends in different patient subgroups (aetiology, age, gender, countries etc), comparing patients who continue the follow-up versus those who do not, and patients with missing data versus those with complete data for each time point.

In the present research, these validation methods were employed and the conclusions were drawn bearing in mind the major limitations.
From a disease-specific measure to a preference-weighted index (utility)

The need to translate results from clinically used QoL measures so they meet health economic requirements is driven by several factors. First of all, most, if not all, disease-specific QoL instruments widely used in clinical practice and in clinical trials are not preference-based and, therefore, collected data are not legitimate for use in cost-utility analysis (29).

This shortcoming is becoming critical and needs to be addressed as there is now a widespread call for assessing incremental costs per QALY of various medical interventions. Furthermore, most clinical studies have not included preference-based measures and have restricted QoL measurement to disease-specific instruments (24). Hence, for many conditions where directly collected utilities are unavailable, modelling could be an acceptable approach to obtain them. This would entail considering five issues:

1. Disease-specific and preference-based QoL measures
2. A set of health-state values serving as a matrix for the model
3. The type of model to use
4. Imposing coefficients estimated in one population onto another
5. Applicability of the methodology to other diseases

QoL measures

The choice of using QoL-AGHDA was made for the following reasons. Firstly, it is the most widely used tool among the disease-specific instruments for adult GHD (108). Secondly, it is used in the KIMS database, and, finally, the NICE recommendations are based on QoL-AGHDA scores (140).

The selection criteria used to choose a preference-based measure included: meeting requirements for generating utilities, the availability of the European preference values and characteristics that allowed for self-completion such as simplicity and brevity. The preference-based techniques considered as gold standards (TTO and SG) are conceptually difficult, and to avoid erroneous answers they should be applied in a face-to-face interview giving the opportunity for additional explanations (177). Therefore, it was decided to make a choice from the generic instruments that had been developed using multi-attribute utility theory: EQ-5D, SF-6D, HU12 or HU13 (122). Among these, EQ-5D was the first choice because of its widespread use in many countries and the availability of population-based preference values (81).

The values to compute $QoL-AGHDA_{utility}$

In estimating utilities, the ultimate results are modified by the applied set of values for different health states. Three main issues are involved in the
choice of such values. First of all, the technique used for their construction – namely TTO, SG or rating scales (VAS) (154, 182), secondly, the reference population (patients or general population) (58, 179) and, finally, whether the current or hypothetical health states are evaluated (175). According to health economic recommendations, the first choice is TTO or SG, although VAS is also acceptable, and the hypothetical health states should be evaluated by the general public (29). Additionally, some studies have shown that differences in preference values exist among countries and therefore have suggested that country-specific preference values should be used (98, 99). Conversely, others have concluded that these differences had no impact on the assessment of treatment benefits (93).

For the Swedish study (paper III), the model proposed by Greiner et al (81) was adopted and a single set of European societal preference values for health states described by the EQ-5D was applied. This decision was driven by the lack of published TTO or SG values for EQ-5D health states in the Swedish population. Although the model derived from the work of Björk and Norinder (25) is the closest to the needs of the present study, the model was published without coefficients and therefore could not be applied. Another option was to use preference values elicited by other models – namely, the VAS scale – to describe personal current health status and TTO for valuing them (117). The final alternative was to use non-Swedish values based on the EQ-5D health state descriptive system and VAS evaluation (81). This had been done in population samples from six European countries (among them Sweden) and the respondents were presented with the hypothetical health states. Given all of the pros and cons, the European VAS values based on the EQ-5D health state descriptions, were chosen as the best possible option.

For the E&W utilities, the original UK preference values were employed. They were derived, using the TTO technique, from the evaluation of hypothetical health states by the general population sample (57). These values have been used in many studies, including Swedish ones (34). From a purely methodological point of view, the choice for E&W seems natural as country-specific estimates were employed and these had been elicited by a technique regarded as a gold standard, with the requisite theoretical basis.

Modelling

The estimates of QoL-AGHDA_{utility} were generated using multiple regression analysis (ordinary least square). The other techniques, such as Jack-knife and bootstrap analyses, were used for internal validation. For the multiple regression analysis it is necessary to decide not only which variables should be entered as independent variables and covariates but also which model should be chosen. For the former, the choice comprised not only different variants of the QoL-AGHDA score (summary score, all individual QoL-AGHDA items or only those that were identified as significant in step-wise forward
regression analysis) but also demographic variables. As a matter of principle, it was assumed that the best model fit, as described by an adjusted R-square, should be a main criterion. Thus, for the Swedish population, the QoL-AGHDA summary score was entered into the model together with either all available demographic information (full model) or only age and gender (simple model). Although, the full model yielded slightly higher \( R^2 \) than the simple one (0.38 vs. 0.36), for practical reasons the latter was adopted. Similar reasoning lay behind the choice of model for E&W where, out of the demographic variables, only age had a significant impact on the utility estimates and individual QoL-AGHDA items showed a better fit than the QoL-AGHDA summary score or items identified in the step-wise regression. Therefore, the final model included age and twenty-five dichotomous QoL-AGHDA items.

The approach of testing several models with different levels of complexity of QoL scores and additional background variables was also undertaken by Brazier et al. who aimed to estimate utilities for the Impact of Weight on Quality of Life – Lite (IWQOL-Lite) instrument from the SF-6D (29). Their results indicated that the more extensive the information entered into the model, the greater the increase in the explanatory power of the model. Similar regression techniques to predict utility from demographic disease-specific QoL data have been used in patients with angina pectoris (115).

One of the first attempts at generating health-state preference values from disease-specific QoL data was undertaken in oncology (44). Two ways of mapping were evaluated. The first was to map cancer-specific QoL data in their original form onto the EQ-5D and HUI (Mark III), the second was to map only a subset of condition-specific data, primarily selected by factor analysis. Although, the second method generated slightly better results, the authors found both to be unsuitable, and therefore recommended that health state preference values should be collected directly in clinical trials. The divergence in domains described by disease-specific and generic measures is believed to be the main reason for poor mapping. In other words, disease-specific measures focus on measuring selected domains in-depth, and thus do not cover all of those included in generic single-index scales, which results in a lack of common dimensions, making mapping difficult.

Kind and Macran (107) proposed a model for converting a standard condition-specific measure used in lung cancer (Functional Assessment of Cancer Therapy-Lung – FACT-L) into a preference-based index. First they constructed a compact system of descriptive health states based on the reduced set of FACT-L items. Then, using the population preference values of hypothetical FACT-L health states in a regression analysis, they estimated item weights, finally computing a preference-based index for FACT-L.

Another study aiming to provide mapping algorithms to generate utilities from disease-specific instruments was performed in patients with Crohn’s disease (38). The modelling was done in a paired setting on two sets of pa-
tient data that included two disease-specific measures for Crohn’s disease together with the SF-6D and EQ-5D. The authors also examined different models for both SF-6D and EQ-5D estimates, including demographic and clinical information as covariates. Consistent with the present results, additional demographic information did not substantially improve model precision.

Interestingly, models for deriving utilities from disease-specific measures that were based on the SF-6D yielded higher adjusted $R^2$ values than those based on the EQ-5D. The former were employed in the studies of Brazier (29) and Buxton (38). The reported $R^2$ values in these studies ranged from 0.37–0.69, whereas EQ-5D-based models in both Buxton’s study (38) and the present study yielded $R^2$ between 0.29 and 0.46. The possible explanation, as suggested by Chancellor and colleagues (44), may lie in the descriptive systems of the modelled instruments, given that the overlap between the content of instruments has an important role in the precision of the model (29). For example, Buxton (38) reported that three out of five EQ-5D domains overlapped domains described when using Inflammatory Bowel Disease Questionnaire (IBDQ), whereas the degree of overlap was greater for the SF-6D (four of five domains overlapped). So, with similar reasoning, it could be stated that the overlap between the QoL-AGHDA and EQ-5D is limited to only two EQ-5D dimensions (anxiety/depression and usual activity in terms of social functioning).

Both models developed in the present work were validated internally and externally. Similar regression coefficients were generated by Jack-knife and bootstrap analyses and thus confirmed the stability of the models. The discriminatory power of QoL-AGHDA utility for self-rated health states and the strong positive correlation between them and EQ-5D VAS in the Swedish data, as well as between QoL-AGHDA utility and PGWB scores, additionally confirmed the reliability of the models.

In adult GHD the previous attempt at generating utilities from QoL-AGHDA was undertaken by Dixon (56) who used a two-step model. The results of this method should be viewed with caution as it incorporated indirect mapping between NHP and QoL-AGHDA in different datasets which may lead to inaccuracy. Furthermore, during the initial step, NHP-based utilities were estimated for general practice patients, not for the general population (as requested by health economists). The strength of the current approach is that both QoL-AGHDA and EQ-5D data used for modelling were collected from the same individuals originating from the general population.

The example of two country-specific datasets (Swedish and English & Welsh) modelled in the present studies suggests that differences exist both in terms of the type of model that best fits these data and in terms of the importance of age and gender.
In summary, it is believed that the models for both countries may facilitate medical decision making by providing a tool for obtaining utilities in the absence of directly collected preference-weighted indexes in the respective countries. However, a direct extrapolation of these results to other cultural environments cannot be recommended, as country-specific characteristics and preference values might be incompatible. On the other hand, modelling as presented here is probably a valuable alternative to utility generation that is worth researching in other diseases.

QoL-AGHDA utility as a treatment outcome in the clinical setting

The following utility outcomes were calculated for clinical evaluation: QoL-AGHDA utility at baseline and at the last reported visit, total QoL-AGHDA utility gain and average QoL-AGHDA utility gain per year of follow-up.

QoL-AGHDA utility was computed for each patient over the whole duration of follow-up, according to the QALY definition of quality and length of life. Such an approach ensures capturing changes in QoL over the study period and incorporating them into the aggregated QoL-AGHDA utility. Nonetheless, owing to the lack of rigid data on survival rates of GH-deficient adults who receive GH, the applied way of QoL-AGHDA utility computation did not account for a difference in the length of life. Therefore, total QoL-AGHDA utility gain represents only a change in utilities, and for the ultimate computation of QALY in this group of patients the values presented here should be used together with mortality data.

Total QoL-AGHDA utility gain was computed as a difference between QoL-AGHDA utility in treated and untreated patients. As data were not available on untreated patients, for the study the values at baseline (before treatment) were taken as representative for the untreated group. Thus, a conservative approach was accommodated assuming that QoL in untreated patients remains at the same level (12, 76), despite some studies reporting deterioration in QoL in untreated patients (78).

In conclusion, the use of estimates of QoL-AGHDA utility response to GH treatment to compute QALY necessitates robust data on mortality in treated patients, as well as reliable information on utilities in patients who do not receive treatment.

Main findings

Population normative data for QoL-AGHDA (papers I and III)

For E&W (paper I) and for Sweden (paper III), the population normative data were developed in detail, whereas for the Netherlands and Spain (paper II) only population means averaged for gender and age of 50 years are pre-
presented. Similarly constructed means were estimated for E&W and Sweden (paper II) and indicated some cross-country differences particularly between the populations of E&W (6.7) and Sweden (3.9). The means for the Dutch and Spanish normative data were very similar (4.9 and 5.0, respectively) but different from the other two countries. The reasons for such variability might be attributed to translation and cultural validation of the measurement instruments as such, to sampling differences or to true cross-country discrepancies.

As already reported, the QoL-AGHDA was developed simultaneously in five language versions, among them English, Spanish and Swedish, and researchers from each country were involved in selecting appropriate items for each concept included in the instrument. The Dutch version was developed separately, but using the same stringent methodology. Therefore, the likelihood that the QoL-AGHDA in these languages expresses the same concepts accounting for the cultural differences is very high.

Obviously, it cannot be excluded that the observed variations are due to differences in sample characteristics. However, population values were derived from large samples and the sampling procedures were carried out at the same time and in a comparable manner for all countries (with the exception of Spain). Additionally, the presented means were adjusted for age and gender, thereby reducing the potential impact of these demographic characteristics. Nevertheless, the question remains open as to whether the mean for Spain would remain the same if these data had been collected almost 10 years later and the sample had covered the whole country.

The normative data for the Swedish population in the present study were consistent with published median QoL-AGHDA scores (194) for men (2.0 in both studies) but were slightly lower than previously reported in women (2.0 vs. 3.0). Such a discrepancy may result from a difference in timing (the data presented here were collected a couple of years after the previous study) and from different sampling methods. Wirén used a sample originating from the Gothenburg region, which thus did not represent the whole country, whereas the current results were based on a random sample of the total Swedish population.

Interestingly, there were also differences in the demographic predictors of the QoL-AGHDA normative values between the data from E&W and Sweden. Age and gender did not significantly explain the variability in QoL-AGHDA scores in the English and Welsh dataset, whereas both factors appeared to have a significant impact in Sweden. In the latter, women tended to have higher scores, indicating poorer QoL and QoL improved with increasing age. The findings in E&W ran counter to the Swedish observation in the present study, as did the normative values reported for QLS-H, another instrument used in adult GHD (27).
Finally, the availability of country-specific normative data for the QoL-AGHDA may increase the precision of evaluation of the reduction in QoL in patients with GHD and hence improve daily patient management.

The effects of GH replacement therapy on QoL (paper II)

**Overall treatment response measured by the total QoL-AGHDA**

The innovative approach applied in this thesis relates to the way in which the treatment results were viewed – not as an absolute value or a change from baseline as measured by a QoL instrument, but as a deficit in QoL in patients before and during treatment versus population normative values. Most previous studies assessed treatment effects in comparison with one of four categories: those observed in a placebo-arm; the pre-treatment results; scores in untreated patients; or scores in non-diseased controls. Additionally, a few RCTs included healthy control groups but restricted comparisons to baseline status, and did not assess post-treatment effects. The only RCT that evaluated treatment effects relative to population norms was done by Mårdh et al. Their conclusions, however, were restricted to the suggestion of normalization of NHP and PGWB scores, without detailed analysis. Furthermore, US normative data were used for the PGWB and Swedish data for the NHP, despite the fact that the patient population originated from twelve European countries.

The results of the present work indicate long-term improvement in overall QoL in GH-deficient patients who receive GH replacement therapy, leading towards normalization when compared with country-specific normative values. These findings do not agree with those reported by Malik and colleagues, who found that, for a minimum of 1 year, treated patients continued to differ in QoL compared with age- and gender-matched controls. Possible explanations for this difference in results may encompass the duration of follow-up and the size of both patient and control cohorts. In this study, despite the dramatic improvement during the first year of treatment, QoL in patients from E&W only reached normal population values after 6–7 years. Thus, normalization may not have been observed in the study by Malik et al because the average duration of treatment was much shorter. Additionally, the patient cohort examined here was larger and a large sample of the general population was used for the comparison and not, as in Malik’s study, a relatively small group of healthy controls. Interestingly, the mean QoL-AGHDA score in their control group was approximately one and a half points lower than reported here, which most likely contributed to the greater deficit between patients and controls. Finally, a striking difference between the two patient cohorts is the proportion of patients with craniopharyngioma...
(15% reported by Malik et al and 8% in the present study) and patients with pituitary adenoma (51% and 67%, respectively) which may have had a substantial impact on the results, as craniopharyngioma is a more severe disorder than pituitary adenoma.

The results presented in this thesis correspond closely to those in the QoL analysis performed in HypoCCS, a similar database on GH-deficient adults receiving GH (Humatrope®) (157). The authors reported that Z-scores for QLS-H in patients were not significantly different from those in age-, gender- and country-matched populations after 4 years of GH replacement therapy. Interestingly they found an identical pattern of response, with a dramatic improvement during the first year, followed by a less rapid but steady enhancement during the following years, as reported here. In addition, similar curves were found for various patient subgroups (gender, disease-onset and completers) in both studies.

The beneficial responses to GH were often similar in different patient subgroups despite observed pre-treatment variations. For example, compared with men, women tended to report a worse QoL at baseline, but this difference diminished during treatment. Patients with adult- and childhood-onset disease and those who had developed GHD due to NFPA or craniopharyngioma also showed a similar magnitude of benefits with the same pattern of response. These findings agree with previous reports (59, 88, 136, 157, 158) but conflict with the results published by Attanasio et al (11), who showed a lower level of distress at baseline and no treatment effects in patients with CO-GHD in contrast to those with AO-GHD. Another study (138) examined GH effects in several patient subgroups using the PGWB and QoL-AGHDA, and reported equal improvement for all subgroups except for men, who demonstrated greater benefits in QoL-AGHDA than women despite indistinguishable pre-treatment scores. The latter finding remains controversial in the light of other reports (5).

Treatment response in individual dimensions measured by the QoL-AGHDA
The magnitudes of the QoL reduction in each dimension were divergent, though the dimensional responses to GH followed the same pattern as the total QoL-AGHDA scores. This finding agrees with the observation by Wirén and colleagues (193) that after the initial improvement (particularly in energy and emotional reactions) most patients experienced continuous benefits for up to 50 months of follow-up.

Even though many studies have investigated treatment effects within specific dimensions (11, 32, 51, 136, 185), little is known about dimensional scores in relation to population norms. Among the few studies to examine this, Mårdh et al. (139) showed that after 6 months of GH treatment, energy levels and emotional reactions were comparable to those in the general population as measured by NHP. On the other hand, scores for social isolation...
remained different, irrespective of the initial dramatic improvement. These findings are contradictory to the ones reported here, where socializing was the first dimension to stabilize within the normal range, whereas memory and tiredness did not reach the population values. There may be different reasons for such a discrepancy, one being the source of normative data; Mårdh and colleagues (139) used Swedish norms for all patients, regardless of their country of origin.

The analysis of changes within QoL dimensions during GH replacement therapy allowed them to be ranked according to the severity of impairment before treatment and also enabled the order of improvement to be defined. As in other studies, memory and tiredness were the most impaired (10, 18, 92), followed by problems with self-confidence, tenseness and social isolation. The novel finding, however, was that improvement in response to GH occurred in the reverse order. In other words, the order of improvement was inversely related to the pre-treatment severity. This finding is even more interesting as it contradicts the concept of regression to the mean.

**In conclusion**, the findings indicated that long-term GH replacement in adults with GHD resulted in a sustained QoL improvement occurring in a stable pattern regardless of the initial level of impairment, characteristics of the patients or dimension of QoL. Such alleviation finally led towards the normal range for each country. This pattern of amelioration, which was reproducible across various patient groups, supports the hypothesis that GHD per se may cause the psychological burden in these patients.

A preference-weighted index (QoL-AGHDA\_utility) (papers III and IV)

**QoL-AGHDA\_utility in the general population**

QoL-AGHDA\_utility was computed for the general populations of E&W and Sweden using methodology already discussed.

Four sets of utilities have been published for the Swedish population (34, 35, 36, 117) and, despite methodological variations, the values of all four lie approximately within similar ranges. Two studies by Burström et al (34, 35) estimated utilities based on the EQ-5D health states description system and UK preference weights (57), whilst the study of Lundberg et al (117) as well as the most recent study of Burström et al (36) employed direct TTO assessments for actual population health states in the general population, but using a self-completion method rather than the traditional interviewer-based procedure. Lundberg also reported the values obtained using rating scales for actual health states as utilities, which from a theoretical point of view can be questioned. Additionally, there were sampling differences; one study used a sample from the whole Swedish population (34) and two used regional samples (Stockholm and Uppsala) (35, 117). The mean utility for the whole
samples were comparable in the studies based on EQ-5D: 0.85 in the present study, 0.84 in the study of the Stockholm sample (35), 0.83 in the sample from the whole Swedish population (34). The mean utility was higher in studies using TTO evaluation – 0.90 (117) and 0.92 (36).

The fact that various methods (SG, TTO and rating scales) elicit different values is well known (13, 47, 154, 177, 182). According to Lenert and Kaplan (112), in most studies the utilities obtained by these three methods rank in the following order: from SG generating the highest values, through TTO to the rating scales (VAS) with the lowest numerical values. This observation is consistent with the results discussed here where EQ-5D utilities were lower than those derived directly from TTO. Possible explanations were sought in Behavioural Decision Theory (unwilling to trade off any time at all) and in Prospect Theory (losses are weighted higher than gains and thus a bigger gain is needed to compensate for a loss) by Robinson et al (154). Overall, there are two trends observed in all studies. There is a consistent gender difference, with men scoring higher – thus indicating a better QoL than women – and also a trend for a decline in values with increasing age. However, utilities based on the QoL-AGHDA were slightly higher in the older groups than those derived by other methods. Possibly this was a consequence of methodological differences, but it may also have resulted from the nature of the questionnaire, which is more psychologically oriented than the EQ-5D and therefore captures a different profile of QoL.

The mean QoL-AGHDA utility presented for the English & Welsh general population was lower than that reported by Christensen and colleagues (46). The variation occurred for both men (0.83 vs. 0.87) and women (0.83 vs. 0.85). Contrary to the findings of Christensen et al, the current study found no gender-difference, whereas in their study women scored lower. Nonetheless, the impact of age was consistent in both studies indicating a decrease in utilities with increasing age. Most likely the results obtained by Christensen et al are more representative of the entire UK population than the data presented here, as their analysis was based on a much larger sample (over 14,000 participants) representing the whole country.

QoL-AGHDA utility in patients with GHD and their response to treatment

Evaluation of the response to GH treatment with respect to QoL-AGHDA utility confirmed a profound deficit before starting GH in relation to normative values, with a rapid improvement during the first year of treatment followed by a steady but less rapid enhancement during long-term therapy. This pattern mirrored the one for the QoL change measured by QoL-AGHDA, albeit that the QoL-AGHDA utility scores did not reach the population means by the end of follow-up. Possibly this variation is rooted in the nature of both measures. QoL-AGHDA captures problems directly linked to GHD, is oriented towards diseased people and, at the same time,
covers a narrower spectrum of QoL. Conversely, a utility-weighted index is based on a scoring system that reflects a broader range of health as experienced by the general population. The different nature of both measures most likely affects results in patients and the general population equally.

As there are differences in utilities derived by different measures in populations of patients with other conditions – for example, in those with end-stage liver disease (30), spinal problems (123) or HIV/AIDS (165), it was decided to discuss here only those generated by EQ-5D for comparing patients with GHD with patients with other diseases. The mean QoL-AGHDAutility pre-treatment value was higher than that reported in patients after stroke with a mild level of dependence in activities of daily living, as measured by the Barthel Index (0.67 vs. 0.58, respectively) (181), and also than those in patients with AIDS (0.63) (165). The mean utility in UK patients with multiple sclerosis was 0.49, ranging from 0.87 to below zero, depending on the progress of disease (147). Parenthetically, it should be explained that the negative utility values for certain health states indicate that such health states are regarded as being worse than death. On the other hand, QoL-AGHDAutility in men (0.70) and women (0.63) with GHD were lower than those generated on the basis of UK population data in extremely obese men and women (0.84 and 0.76, respectively) and in men and women with type I diabetes (0.82 and 0.75, respectively) (46).

These utilities discussed above were elicited directly from EQ-5D data, whereas the estimated utilities presented here for patients with GHD were based on the population weights. The impact of such a methodological difference could be studied if EQ-5D data were collected directly in patients and the findings compared with the present ones. The estimates of utilities generated by modelling in patients with lung cancer ranged from 0.70 to 0.11 (107).

Overall utilities estimated for GHD could be placed in the upper range of those reported for other severe diseases. Furthermore, the scores analysed in different patient subgroups showed a narrower range of variation (between 0.60 and 0.80), than that reported for other diseases (28). This observation may indicate that generally patients with GHD represent a more homogeneous group with respect to severity of disease and thus a smaller variety of health states are concerned. On the other hand, a narrow range of scores may result from patient selection, and might not be representative of the entire population of patients with GHD.

During GH treatment there was a statistically significant improvement in QoL in the patient population as measured by a total QoL-AGHDAutility gain (mean, 0.32) and a gain per year (mean, 0.08). The interesting question, however, is whether such a change has any clinical meaning. The statistical assessment of a minimal clinically important change in the analysed dataset falls outside the scope of the present thesis. Nevertheless, the reported mean gain per year exceeded the value for the EQ-5D suggested by other research-
ers as a minimal important difference, which ranged from 0.033 to 0.07 points (116). Moreover, the results of the analysis of the effects of antidepressant treatment (164) in Swedish patients with depression (mean increase 0.23 over 6 months) suggest that the utility gain observed in patients with GHD may be considered as clinically significant.

To summarise the subgroup analyses, it can be inferred that, despite differences at baseline, demographic and clinical characteristics had no impact on the response to treatment. All patients, except those with CO-GHD and those with co-morbidities, experienced similar total and annual QoL-AGHDA utility gain.

Finally, the innovative aspect of the present approach is to apply preference-weighted indices derived from a disease-specific measure to assess QoL in the clinical context together with the patients’ demographic and clinical characteristics. The robustness of this analysis is reinforced by the fact that utilities in both general and patient populations were generated using the same methodology.

Final remarks

It is hoped that this thesis will contribute to bridging the gap between clinicians and health economists and provide a tool that may be used in health economic assessment and hence facilitate medical decision making. However, this work has certain limitations, and, though these have already been discussed, it is worth emphasising them again. The main concern is related to the generalizability and applicability of the results to the general populations of other countries, given that those here were derived from well-defined samples originating from Western European countries. Similarly, the patient data were restricted only to a subset of treated patients and therefore are most likely not representative of the entire population of patients with GHD. Finally, the patient data were retrieved from an observational database, thus, by definition, lacking randomization. The casual interpretation of a change observed in any clinical study, other than double-blind, placebo-controlled trials, is debatable. At the same time, given the ethical considerations of conducting such a placebo-controlled study in patients with approved indications, the next best choice would be an observational study with a large number of patients and long follow-up. In this way the study attempted to compensate for the lack of placebo-controlled data.

It is necessary to emphasise that presented utilities serve as only one component of QALY, i.e. they stand for ‘quality of life’ and information on the ‘quantity of life’ is still missing. Therefore, reliable information on survival rates in GH-deficient patients receiving GH replacement is essential. Only combined information on the impact of treatment on both the quality and quantity of life allows for full assessment of treatment benefits. Fur-
thermore, for the purpose of a full cost-utility analysis, data on treatment costs are also critical. In addition, for the allocation of health-care resources, health policy makers require epidemiological information. This brief and very general walk through the whole process of establishing treatment policies attempts to adequately place the current work in the whole chain of required information.

Notwithstanding, this work is consistent with the current standard practice in health economics, where QALY maximization constitutes the main criterion for allocating health-care resources. However, this thesis would not be complete without acknowledging that several issues remain open to challenge, mainly from the ethical point of view (146). The main argument is that in certain circumstances assessing health priorities based on maximal QALY gain may lead to attaching unjustly higher values to some lives, but not to others (143). This has led to attempts to develop means of capturing public preferences more comprehensively by including, for example, rules of fairness and equity (162). There has also been heavy criticism of the recommended ways of eliciting utilities, Arnesen and Norheim (9) maintaining that the assumptions underpinning TTO, which is regarded as a gold standard for eliciting utilities, are unrealistic and inconsistent in real world. These considerations are obviously a matter for a thorough debate and cannot be solved within the frame of this thesis. There are strong voices that reject the state-of-the-art in health economics and are seeking a method to incorporate a wide set of human values into the systems concerned with one of the most important aspects of life itself: human health.
Acknowledgements

As one of my colleagues, Kim Wittrup-Jensen said: “to write a PhD thesis is rarely a one-man show”. This is very much the case in respect of my thesis as well. Therefore, I would like to take the opportunity of expressing my sincere thanks to all who supported me in many different ways and contributed to this work.

First of all, I want to thank Patrick Wilton, my boss at Pfizer Endocrine Care, without whose approval and supportive acceptance of my crazy ideas nothing would have happened.

Patrick – thank you for standing behind me, believing in me and for your constructive criticism and valuable comments; this goes far beyond my thesis, but it applies as well to our daily work with KIMS.

My gratitude goes to my supervisors Paul Kind from the University of York as well as Chris Bingefors and Dag Isacson from Uppsala University.

Paul – you introduced me to QoL research during a course many years ago in York; you encouraged me to start the work and you have taught me to crawl and walk in this field and I am still hoping to get some lessons on running. Your deep knowledge, thoughtful guidance and continuous support cannot be overestimated, not overlooking the English teaching, of course. All that remains unforgettable!

Chris and Dag – thank you both, first of all for accepting me as your PhD student at the Department of Pharmacy, but also for long hours of fruitful discussions, sharing your knowledge in epidemiology and stimulating independent thinking – and last, but definitely not least, Chris – for guiding me and helping me out with all formalities through the jungle of the Swedish regulations and procedures.

There are a lot of statistics in my work. This would never have been possible without the contribution of Björn Jonsson, who was a key person throughout my work.

Björn – it is quite difficult to find words to describe my gratitude. Tusentack for guiding, explaining, helping, showing unlimited patience and teaching me the complicated issues of understanding the numbers. Not only for that, but also for keeping me going whenever it was needed – my breakdowns happened pretty often!

As far as statistical mysteries are concerned, I would like to thank very much Anders Mattsson, who took care of modelling in two of the papers
constituting this thesis and who, together with Peter Jönsson, was my first teacher in statistics.

**Anders and Peter** – we all know how it looked when I started my job in early 2001; maybe a notion of a mean existed, but a median was already quite confusing. You two showed unbeatable understanding and patience and kept on taking it all over again until I eventually managed to absorb – step by step, number by number and analysis by analysis. Everything I know now originates from your lessons – **tack så hemsk mycket**!

Staying with the numbers and technical issues, Björn Westberg has been always an important part of it.

**Björn** – a big thank for explaining to me how a database functions, how one should communicate with it and how to ask questions and read answers. But also for being there whenever anything “for yesterday” was needed.

I would like to thank the entire Department of KIGS/KIMS/ACROSTUDY Medical Outcomes in Sweden for continuous support and for a feeling of being and doing things together. Also for assisting me from nonexistent to a bit of Swedish and taking on the burden of English in our daily communication.

**Alla mina Vänner** – all your contributions were essential – from stimulating discussions and constructive criticism, ensuring that data collected were of high quality, distributing information among many international contacts, to taking care of the many IT issues.

Additional thank-you goes to Gunilla Wallén, who helped me with many practicalities related to this thesis.

**Gunilla** – for giving me a fantastic feeling of security whatever I am involved in, for looking after every small detail to get everything working, for being such a reliable person and a wonderful companion – **tackar**.

I wouldn’t be a part of this wonderful team but for Annika Wallström.

**Annika** – a very special thank-you for getting me to Stockholm, which made a real change in my life, and also for all you have done to build up the databases.

Further, I am indebted to Ione Kourides, the head of Pfizer Endocrine Care in New York.

**Ione** – I appreciate your constructive criticism and support for our activities.

I would also like to thank the KIMS Strategic Advisory Board, the members of KIMS International Board and all KIMS investigators from all over the world who provided data on their patients, my colleagues working in different countries who are instrumental in collecting these data, but first of all the patients who agreed to be followed in KIMS. Again, without your contribution, this database would not exist.

Particular thanks go to my colleagues Annie Ahlberg from Sweden, Netty Dorrestijn from The Netherlands, Kristin Forssmann from Germany, Carine Claeys from Belgium and Anne Chachuat from France who helped me set up
the studies in their countries. Your enthusiastic help cannot be overestimated – thank you!

I am also grateful to all the participants of population surveys which were carried out.

I want to thank Pfizer for the financial support – it is needless to emphasize, that without it, it would not have been possible to run the studies.

I am especially grateful to the outstanding group of endocrinologists with whom I have been privileged to work during all these years and who had introduced me to the field of pituitary disorders: John Monson, Ulla Feldt-Rasmussen, Roger Abs, and Bengt-Åke Bengtsson. It has been a great experience, not only scientifically but also educationally, learning from your extensive knowledge and rich personalities. Thank you for sharing all that with me!

John – you have been actively involved in almost all of the papers in my thesis and far beyond that with your key contribution to the KIMS database. Thank you very much for spending uncountable hours with me solving my doubts and hesitations, as well as going back and forth through endless versions of manuscripts and hundreds of pages of statistical output. Thank you for personal support when things were getting difficult.

Ulla - for being with me and supporting me for all these years, for building up KIMS and for spending hours reading through my thesis which resulted in many valuable comments and remarks – mange tak!

It is my pleasure to recognize also the co-authors on my papers Susan Hennessy, Gudmundur Johannsson, Felipe Casanueva, Xavier Badia, Jan Busschbach and Hans Koppeschaar whose help by assisting in statistical analysis, contributing with data and thoughtful comments inevitably increased the value of my work.

I would like to acknowledge all the people who helped me in editing and getting it all into proper English - Stephen Lock, and my colleagues from PharmaGenesis, especially Harriet Crofts, David Bennett and David Campbell.

Stephen – my warmest thank-you for your help with putting my Slavonic thoughts into English sentences, for responding to my questions and linguistic problems in such a prompt, helpful and nice way.

Harriet, David and David – All of us working with KIGS, KIMS and ACROSTUDY, have appreciated your reliable and high-quality help with our manuscripts and materials. Now, it is an occasion for me to express it and also to add a special recognition for help with “my” manuscripts, even with the most difficult ones – full of utilities and modelling. Harriet – for unbeatable help in editing and proof-reading it all – many warm thanks and I am so sorry for destroying your weekends!

Charlotte Höybye – for reading it all through and providing me with very good advice – Charlotte – thank you
I would like as well to address Jesper Andersson and the personnel from the Publication and Graphic Services at Uppsala University for efficient help with layout, design and publication.

**Jesper** – for all your help and for finding “my” song, which was just fantastic – **tack så mycket!**

And to all my friends, whom I met during 16 years working in the company, particularly in Poland – for your friendship which is a firm and constant source of strength and re-charges one’s batteries – the word “thank-you” feels far too weak to express it.

Finally, the biggest thank-you goes to the most important people in my life – my Family. I have written this thesis for you who just make my life happen.

**Mamo i Tato** – for being with me and for me for all the years, for bringing me up and making me what I am now; for your unconditional love – **dziekuję z całego serca!**

**Łukulinku, Siuleczku i Kubulinku** – for everything you are in my life, but also for what you are now in your lives; for each and every second we spent together, for the time we have been apart, but still in our hearts tightly together; for teaching me to understand others, patience, tolerance and trust; for our happiness and our sorrow; for love – **najmocniej jak umiem, dziękuję!**

**Bosse** – it would take a whole book to describe who you are in my life… Thank you for your love, for building up self-confidence in me, for serving as a “punch bag” so often, for sharing my life, for all the years – first professionally but then much more personally… **My name is Maria Kołtowska and I have three children…** and of course for standing countless weekends and evenings when I was at home, but not really!
References

3. Abs R, Feldt-Rasmussen U (eds) 2004 Growth Hormone Deficiency in Adults: 10 Years of KIMS, Oxford PharmaGenesis™
onset and adult onset before and during human GH treatment. J Clin Endocrinol Metab 82(1): 82-88


15. Barkan A, 2001 The Quality of Life-Assessment of Growth Hormone Deficiency in Adults questionnaire: can it be used to assess hypopituitarism? J Clin Endocrinol Metab 86(5): 1905-1907


64. EuroQoL Group 1990 EuroQoL- a new facility for the measurement of health-related quality of life. Health Policy 16: 19-208
78. Gilchrist FJ, Murray RD, Shalet SM 2002 The effect of long-term untreated growth hormone deficiency (GHD) and 9 years of GH replacement on quality of life (QoL) of GH-deficient adults. Clin Endocrinol (Oxf) 57: 363-370
83. Gutiérrez LP, Koltowska-Häggström M, Jönsson PJ, Mattsson AF, Svensson D, Westberg B, Luger A 2007 Registries as a tool in evi-
dence-based medicine: example of KIMS (Pfizer International Metabolic Database). Pharmacoepidemiology and Drug Safety doi: 10.1002/pds.1510


90. Holmes SJ, Shalet SM 1995 Characteristics of adults who wish to enter a trial of growth hormone replacement. Clin Endocrinol (Oxf) 42: 613-618


101. Jørgensen JOL, Christiansen JS (eds), Growth Hormone Deficiency in Adults. Front Horm Res. Basel, Karger vol 33


114. Likert R 1932 A technique for the measurement of attitudes. Arch Psychol 140: 1-55


Deficiency in Adults: 10 Years of KIMS, Abs R, Feldt-Rasmussen U, (eds), Oxford PharmaGenesis™ pp: 117-126


140. National Institute for Clinical Excellence (NICE), 2003 Technology Appraisal 64: Human Growth Hormone (Somatotropin) in Adults with Growth Hormone Deficiency. NICE,


143. Nord E 2001 The desirability of a condition versus the well being and worth of a person. Health Econ 10: 579-81
158. Saller B, Mattsson AF, Kann PH, Koppeschaar HP, Svensson J, Pomp- 
en M, Kołtowska-Häggström M 2006 Healthcare utilization, quality of 
life and patient-reported outcomes during two years of growth hormone 
(GH) replacement therapy in GH-deficient adults – comparison be-
tween Sweden, The Netherlands and Germany. Eur J Endocrinol 
154: 843-850

159. Sanmarti A, Lucas A, Hawkins F, Webb SM, Ulied A on behalf of 
the Collaborative ODA (Observational GH Deficiency in Adults) 
Group 1999 Observational study in adult hypopituitary patients with 
untreated growth hormone deficiency (ODA study). Socio-economic 
impact and health status. Eur J Endocrinol 141: 481-489

Growth hormone treatment in adults with childhood onset growth hor-
mone deficiency: effects on psychological capabilities. Horm Res 44: 
6-11

161. Sathiavageeswaran M, Burman P, Lawrence D, Harris AG, Falleti MG, 
Maruff P, Wass J 2007 Effects of GH on cognitive function in elderly 
patients with adult-onset GH deficiency: a placebo-controlled 12-month 
study. Eur J Endocrinol 156: 439-447

162. Schwappach DLB 2002 Resource allocation, social values and the 
QALY: a review of the debate and empirical evidence. Health Expecta-
tions 5: 210-222

163. Smith F 1997 Survey research; (1) Design, samples and response. Inter 
J Pharm Pract Sept: 152-166

164. Sobocki P, Ekman M, Ägren H, Krakau I, Runeson B, Mårtensson B, 
Jönsson B 2007 Health-related quality of life measured with EQ-5D in 
patients treated for depression in primary care. Value in Health 10(2): 
153-160

165. Stavem K, Frøland SS, Hellum KB 2005 Comparison of preference-
based utilities of the 15D, EQ-5D and SF-6D in patients with 
HIV/AIDS. Qual Life Res 14: 971-980

166. Stochholm K, Gravholt CH, Laursen T, Jørgensen JO, Laurberg P, 
Andersen M, Kristensen LO, Feldt-Rasmussen U, Christiansen JS, Fry-
denberg M, Green A 2006 Incidence of GH deficiency – a nationwide 
study. Eur J Endocrinol 155: 61-71

167. Stochholm K, Gravholt CH, Laursen T, Laurberg P, Andersen M, Krist-
tensen LO, Feldt-Rasmussen U, Christiansen JS, Frydenberg M, Green 
157: 9-18

J, Fukuhara S 2006 Validation of the Japanese version of the Quality of 
Life – Assessment of Growth Hormone deficiency in Adults (QoL-


194. Wirén L, Whalley D, McKenna SP, Wilhelmsen L 2000 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 (Oxf) 52: 143-152

Appendices

Appendix A: Aetiology of growth hormone deficiency (KIMS classification list)
KIMS Etiology classification list

Idiopathic
1.1  Idiopathic GHD

Congenital

Genetic cause of GHD
2.1.1.1  GH-gene-defect (Type 1A dominant or recessive)
2.1.1.2  GH-gene-defect (specify)
2.1.1.3  GHRH-gene-defect (specify)
2.1.1.9  Other genetic cause of GHD (specify)

Central malformation
2.1.2.1  Septo-optic dysplasia
2.1.2.2  Empty sella syndrome (including pituitary aplasia)
2.1.2.3  Solitary central maxillary incisor syndrome
2.1.2.4  Mid-line palatal cleft
2.1.2.5  Arachnoid cyst
2.1.2.6  Congenital hydrocephalus
2.1.2.9  Other central malformation (specify)

Complex syndrome with congenital GHD
2.1.3.1  Fanconi pancytopenia
2.1.3.2  Rieger syndrome
2.1.3.3  EEC syndrome (Ectrodactyly-Ectodermal Dyplasia-Clefting syndrome)
2.1.3.9  Other complex syndrome with congenital GHD (specify)

Prenatal infection
2.1.4.1  Rubella
2.1.4.9  Other than rubella prenatal infection (specify)

Bio-inactive GH syndrome
2.1.5.1  Kowarski type
2.1.5.9  Other than Kowarski type bio-inactive GH syndrome (specify)

Functional GHD
2.1.6.1  GH-receptor defect (Laron Type)
2.1.6.2  GH-receptor/postreceptor defect (specify)
2.1.6.3  IGF resistance (specify)
2.1.6.9  Other functional GHD (specify)

Acquired

Tumors of the pituitary/hypothalamic area
2.2.1.1  Craniopharyngioma
2.2.1.2  Germ cell tumors (specify: dysgerminoma, pinealoma)
2.2.1.3  Hamartoma
2.2.1.4.1  Adenoma
2.2.1.4.2  non-secreting (non-functioning pituitary adenoma – NFPA)
2.2.1.4.3  ACTH (Cushing disease)
2.2.1.4.4  GH (acromegaly)
2.2.1.4.5  Prolactin (prolactinoma)
2.2.1.4.6  Gonadotropin
2.2.1.4.7  TSH
2.2.1.4.8  Co-secreting (specify)
2.2.1.4.9  Other adenoma (specify)
2.2.1.5  Cyst (specify: Rathke's, epidermoid, dermoid)
2.2.1.6  Glioma
2.2.1.7  Meningioma
2.2.1.8  Schwannoma
2.2.1.10 Chordoma
2.2.1.11 Primary pituitary carcinoma
2.2.1.12 Sarcoma
2.2.1.13 Metastatic carcinoma
2.2.1.14 Hematologic metastases
2.2.1.9 Other tumors of pituitary/hypothalamic area (specify)

Cranial tumors distant from the pituitary/hypothalamic area
2.2.2.1  Astrocytoma
2.2.2.2  Ependymoma
2.2.2.3  Glioma
2.2.2.4  Medulloblastoma
2.2.2.5  Nasopharyngeal tumour
2.2.2.9  Other cranial tumors distant from the pituitary/hypothalamic area (specify)

Treatment for malignancy outside the cranium
Leukemia
2.2.3.1.1 Lymphatic leukemia
2.2.3.1.2 Myeloid leukemia
2.2.3.1.3 Aplastic leukemia
2.2.3.1.9 Other leukemia (specify)

Lymphoma
2.2.3.2.1 Hodgkin lymphoma
2.2.3.2.2 Non-Hodgkin lymphoma
2.2.3.2.9 Other lymphoma (specify)

Solid tumor
2.2.3.3 Solid tumor (specify)

Other causes of acquired GHD
Head trauma/injury
2.2.4.1.1 Perinatal head trauma
2.2.4.1.2 Traumatic brain injury
2.2.4.2 CNS infection (specify: meningitis, encephalitis, septic cavernous sinus)
2.2.4.3 Hydrocephalus
2.2.4.4 Granulomatous diseases (specify: sarcoidosis, tuberculosis, syphilis, fungal)
2.2.4.5 Langerhans cell histiocytosis (histiocytosis X, eosinophilic granuloma
Hand-Schüller-Christian disease)
Vascular system
2.2.4.6.1 Infarction (apoplexy)
2.2.4.6.2 Postpartum necrosis (Sheehan syndrome)
2.2.4.6.4 Aneurysm
2.2.4.6.5 Sickle cell anemia
2.2.4.6.6 Thalassemia
2.2.4.6.9 Other vascular (specify)
2.2.4.7 Lymphocytic hypophysitis
2.2.4.8 Hemochromatosis
2.2.4.9 Other (specify)
Appendix B: Overview of the studies investigating QoL in adult GHD
<table>
<thead>
<tr>
<th>Study design</th>
<th>Year</th>
<th>Authors</th>
<th>Patient population</th>
<th>Duration</th>
<th>Questionnaires</th>
<th>Effect on QoL</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1986</td>
<td>Almqvist et al</td>
<td>N=5; age 22-36 yrs</td>
<td>8 weeks</td>
<td>Cognitive tests</td>
<td>↑ Cognitive function</td>
<td>Native human GH and recombinant GH</td>
</tr>
<tr>
<td>Cross-over</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1990</td>
<td>Degerblad et al</td>
<td>N=6; age 20-36 yrs;</td>
<td>12 weeks</td>
<td>Subjective general well-being; POMS, Sjöberg Mood</td>
<td>↑ Subjective general well-being = mood = cognitive function</td>
<td>5 of the patients properly identified the GH period and reported increased mental alertness and vitality and improved physical capacity and muscle strength</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td>n= 5 had CO-GHD</td>
<td></td>
<td>Questionnaire, cognitive tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-over</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1990</td>
<td>McGauley et al</td>
<td>N=24; age 18-55 yrs</td>
<td>6 months</td>
<td>NHP, PGWB, GHQ-60,</td>
<td>↑ Energy ↑ Mood</td>
<td></td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1992</td>
<td>Whitehead et al</td>
<td>N=14, age 19.5-52yrs</td>
<td>13 months</td>
<td>NHP, PGWB</td>
<td>=QoL</td>
<td>in 4 patients IGF-I did not raise indicating lack of response to GH (too low dose? Lack of compliance?)</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-over</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1993</td>
<td>Bengtsson et al</td>
<td>N=10; age 34-58yrs</td>
<td>12 months</td>
<td>Psychiatric evaluation, CPRS, SCL-90</td>
<td>↑ Overall QoL ↓ Psychiatric symptoms</td>
<td>Deterioration during withdrawal in all 4 patients in GH/placebo group</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1994</td>
<td>Mårdh et al</td>
<td>N=233; age 20-60 yrs</td>
<td>6 months</td>
<td>NHP, PGWB</td>
<td>↑ Overall QoL ↑ Vitality ↑ Energy ↑ Emotions ↓ Social isolation</td>
<td>Improvement of QoL led towards normalization when compared with the healthy population</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1995</td>
<td>Beshyah et al</td>
<td>N=40; age 19-67yrs</td>
<td>6 months</td>
<td>GHQ-60, CPRS</td>
<td>=QoL</td>
<td>Improvement on GHQ in the placebo group</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1995</td>
<td>Burman et al</td>
<td>N=36; age 28-57 yrs</td>
<td>21 months</td>
<td>HSCL, PGWB, NHP, partner questionnaire</td>
<td>↑ Energy ↑ Vitality ↓ Anxiety</td>
<td>Placebo effect should be account for</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind</td>
<td>1997</td>
<td>Attanasio et al</td>
<td>N=173; age range- not specified</td>
<td>6 months</td>
<td>NHP</td>
<td>↓ Social isolation ↑ Physical mobility</td>
<td>The effects seen only in AO-GHD</td>
</tr>
<tr>
<td>Randomized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo-controlled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Type</td>
<td>Authors</td>
<td>Sample Size</td>
<td>Duration</td>
<td>Methods</td>
<td>Outcomes</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Carroll et al</td>
<td>N=42; age 18-55yrs</td>
<td>6 months</td>
<td>NHP, PGWB</td>
<td>↑ Overall QoL</td>
<td>Dose finding study; no difference in QoL benefits between the GH dose groups</td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Verhelst et al</td>
<td>N=148; age: 20-60 yrs;</td>
<td>6 months</td>
<td>NHP, social self-reporting questionnaire</td>
<td>↑ Energy ↑ Emotions ↑ Sleep ↓ Sick leave days ↓ Hospitalization days</td>
<td>Changes in NHP during the first 6 months partly due to the placebo effect; however the benefits were maintained in the open phase</td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Wallymahmed et al</td>
<td>N=32; age range-not specified</td>
<td>6 months</td>
<td>LFS, DIS, NHP, HADS, SES, MFS</td>
<td>↑ Self-esteem ↑ Energy ↑ Emotions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Baum et al</td>
<td>N=40; age 24-64 yrs; men with AO-GHD</td>
<td>18 months</td>
<td>Cognitive tests, NHP, PGWB, GHQ, MMPI-2,</td>
<td>=QoL =cognitive function</td>
<td>Placebo effect during the placebo control phase, however improvement during the open phase</td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Cuneo et al</td>
<td>N=166; age: 17-67yrs</td>
<td>6 months</td>
<td>NHP, GHDQ, social history</td>
<td>= QoL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Deijen et al</td>
<td>N=48; men with CO-GHD age range-not specified</td>
<td>6 months</td>
<td>Cognitive function, HSCL, POMS, STAI</td>
<td>↑ Emotions = psychological wellbeing</td>
<td>Anxiety decreased after 2 years of treatment</td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Cross-over Placebo-controlled</td>
<td>Florkowski et al</td>
<td>N=20; age 20-69 yrs</td>
<td>6 months</td>
<td>DSQ, SCL-90, SAS</td>
<td>=QoL</td>
<td>Any observed effect results from placebo effect</td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>Giusti et al</td>
<td>N=26; age 21-74 yrs;</td>
<td>6 months</td>
<td>General psychiatric interview, KSQ, HDS,</td>
<td>↓ Depression</td>
<td>Correlation between a decrease in HDS score and increase in IGF-1 (r² = -0.56; p=0.05)</td>
<td></td>
</tr>
<tr>
<td>Open treatment</td>
<td>Wirén et al</td>
<td>N=161; age 19-76 yrs</td>
<td>Up to 50</td>
<td>NHP, PGWB</td>
<td>↑ Overall QoL</td>
<td>71 patients were followed</td>
<td></td>
</tr>
<tr>
<td>Study Type</td>
<td>Year</td>
<td>Authors</td>
<td>Participants</td>
<td>Duration</td>
<td>Measurements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>---------</td>
<td>--------------</td>
<td>----------</td>
<td>--------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>1999</td>
<td>De Noaves Soares et al</td>
<td>N=9; age 28-52 yrs</td>
<td>6 months</td>
<td>HDS, Beck Depression Inventory (BDI), cognitive tests, measurement of attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↓ Depression, ↑ Attention, ↑ Cognitive efficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>2004</td>
<td>Oertel et al</td>
<td>N=18; age: 21-63 yrs</td>
<td>6 months</td>
<td>Cognitive tests; NHP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ Attention, ↑ Energy, =cognitive function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>2004</td>
<td>Mahajan et al</td>
<td>N=25; age: 18-59 yrs</td>
<td>4 months</td>
<td>Psychiatric interview; NHP, GHQ-28, SCAN, HDS, MADRS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↓ Depression, ↓ Social isolation, ↑ Emotions, ↑ Energy, ↑ Sleep</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>2006</td>
<td>Chihara et al</td>
<td>N=73; age 18-65 yrs</td>
<td>24 weeks</td>
<td>SF-36; QoL-AGHDA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ Overall QoL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-blind Randomized Placebo-controlled</td>
<td>2007</td>
<td>Sathiavageeswaran et al</td>
<td>N=34; age 60-77 yrs; All had AO-GHD</td>
<td>12 months</td>
<td>Neurobehavioral Examination System – cognitive function, POMS - mood</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ Memory, = Mood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Withdrawal studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>After 6 months, no change after 12 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Withdrawal study Double-blind Placebo-controlled</td>
<td>2003</td>
<td>McMillan et al</td>
<td>N=21; age 25-68 yrs</td>
<td>3 months</td>
<td>HDQoL, GWBI, W-BQ12, SF-36, NHP, GHQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↓ Energy, ↑ Tiredness, ↑ Pain, ↑ Irritability, ↑ Depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Withdrawal study Open treatment</td>
<td>1995</td>
<td>Sartorio et al</td>
<td>N=8; age 25-34 yrs; men with CO-GHD</td>
<td>6 months</td>
<td>Cognitive tests, STAI, Experimental World Inventory (EWI), gender identification tests, reaction to stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑ Cognitive function, ↑ Emotions, ↑ Stress handling = body image</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case reports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case report</td>
<td>1962 Raben</td>
<td>One patient</td>
<td>2 months</td>
<td>Observation</td>
<td>↑ Vigor</td>
<td>↑ Ambition</td>
<td>↑ Sense of well-being</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Follow-up studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up study</td>
<td>1999 Gibney et al</td>
<td>N=21; age 21-51 yrs</td>
<td>10 years; continuation of 6 months placebo-controlled study</td>
<td>NHP,</td>
<td>↑ Energy</td>
<td>↑ Emotions</td>
<td>↑ Overall QoL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Comparison between 10 patients who were treated continuously and 11 who were not</td>
</tr>
<tr>
<td>Follow-up study</td>
<td>2002 Gilchrist et al</td>
<td>N=61; age range-not specified</td>
<td>9 years</td>
<td>NHP, PGWB</td>
<td>↑ Overall QoL</td>
<td>↑ Energy</td>
<td>Follow up of 12 months placebo controlled study</td>
</tr>
<tr>
<td>Follow-up study</td>
<td>2005 Artwert et al</td>
<td>N=23; men with CO-GHD; age range-not specified</td>
<td>10 years; continuation of 6 months placebo-controlled study</td>
<td>Education; HSCL, POMS, STAI, memory function</td>
<td>↑ Memory</td>
<td>↑ Vigor</td>
<td>↓ Anxiety</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50 participating in the initial study; of these 36 followed for 39-69 months</td>
</tr>
<tr>
<td><strong>Open treatment studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-sectional analysis</td>
<td>1989 Björk et al</td>
<td>N=36; age range-not specified</td>
<td>2-12 years</td>
<td>NHP, PGWB, tailored questionnaire</td>
<td>↓ Social isolation</td>
<td>↓ Physical mobility</td>
<td>↓ Sleep</td>
</tr>
<tr>
<td>Open treatment</td>
<td>1998 Drake et al</td>
<td>N=50; age 18-69 yrs</td>
<td>Up to 12 months</td>
<td>QoL-AGHDA</td>
<td>↑ Overall QoL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open treatment</td>
<td>1999 Murray et al</td>
<td>N=65; age 17-72 yrs</td>
<td>8 months</td>
<td>QoL-AGHDA, PGWB,</td>
<td>↑ Overall QoL</td>
<td>↑ all PGWB domains with vitality improving most</td>
<td></td>
</tr>
<tr>
<td>Open treatment</td>
<td>2001 Ahmad et al</td>
<td>N=46; age 26-72 yrs</td>
<td>3 months</td>
<td>QoL-AGHDA</td>
<td>↑ Overall QoL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-sectional analysis</td>
<td>2003 Malik et al</td>
<td>N=120; aged &lt;70 yrs</td>
<td>GH &gt; 1 yr</td>
<td>NHP, SF-36, HADS, SES, MFQ, LFS, DIS, QoL-AGHDA, VAS</td>
<td>Maintained post-treatment impairment vs. healthy controls</td>
<td>QoL scores compared with controls and treatment effects have not been assessed</td>
<td></td>
</tr>
</tbody>
</table>
QoL in cancer survivors was compared with that in patients with pituitary pathology and no difference in QoL between these groups was observed.

Benefits in QoL did not correlate with improvements in body composition.

Overall QoL did not improve in all PGWB domains with vitality improving most.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with TBI was compared with patients with Sheehan's disease.

Comparisons with younger patients, similar effects.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

Comparisons with younger patients, similar effects.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.

QoL in patients with Sheehan's disease was compared with females with NFPA, who improved.
<table>
<thead>
<tr>
<th>Institution</th>
<th>Year</th>
<th>Authors</th>
<th>Sample Size</th>
<th>Sample Characteristics</th>
<th>Duration</th>
<th>Measurement</th>
<th>Outcome</th>
<th>Comparison Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIMS</td>
<td>2005</td>
<td>Verhelst et al</td>
<td>N=351; age range-not specified</td>
<td>craniopharyngioma</td>
<td>2 years</td>
<td>QoL-AGHDA</td>
<td>↑ Overall QoL</td>
<td>Comparison with NFPA; similar response</td>
</tr>
<tr>
<td>KIMS</td>
<td>2006</td>
<td>Maiter et al</td>
<td>N=447; irradiated patients age range-not specified</td>
<td></td>
<td>2 years</td>
<td>QoL-AGHDA</td>
<td>↑ Overall QoL</td>
<td>Compared with non-irradiated</td>
</tr>
<tr>
<td>KIMS</td>
<td>2006</td>
<td>Saller et al</td>
<td>N=503 age range-not specified</td>
<td></td>
<td>2 years</td>
<td>QoL-AGHDA, KIMS PLSF</td>
<td>↑ Overall QoL, ↓ Sick leave days, ↓ Hospitalization days, ↓ Doctor visits</td>
<td>Cross country comparisons: Sweden, The Netherlands and Germany</td>
</tr>
</tbody>
</table>
Appendix C: Questionnaire used in the studies constituted the thesis

About Your Health

This questionnaire asks you some general questions about your health. There are no right or wrong answers we are just interested in how you are feeling. How people describe their health sometimes differs according to their personal characteristics (such as age, sex or the type of job they do). It will help us to understand your answers better if you also answer the general background questions that are included with this questionnaire.

For each group of statements please indicate which one best describes your health today. Please tick one box for each group of statements.

Mobility
- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

Self-Care
- I have no problems with self-care
- I have some problems washing or dressing myself
- I am unable to wash or dress myself

Usual Activities
- I have no problems with performing my usual activities (e.g. work, study, housework, family or leisure activities)
- I have some problems with performing my usual activities
- I am unable to perform my usual activities

Pain/Discomfort
- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

Anxiety/Depression
- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed
Think about how good or bad your own health is today.

This scale may help. The best health you can imagine is marked 100 and the worst health you can imagine is marked 0.

Please write in the box below, the number between 0 and 100 that you feel best shows how good your health is today.
Listed below are some statements that people may make about themselves. Read each statement carefully and put a tick in the box marked YES if you think it applies to you.

If you answered YES, then please tick a box to show how much this problem affects your overall quality of life.

**Tick the box marked NO if you think it does not apply to you.**

<table>
<thead>
<tr>
<th>Statement</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have to struggle to finish jobs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel a strong need to sleep during the day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I often feel lonely even when I am with other people</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have to read things several times before they sink in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>It is difficult for me to make friends</td>
<td></td>
<td></td>
</tr>
<tr>
<td>It takes a lot of effort for me to do simple tasks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have difficulty controlling my emotions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I often lose track of what I want to say</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I lack confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have to push myself to do things</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I often feel very tense</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**How much does this problem affect your overall quality of life?**

1. not at all 2. slightly moderately 3. quite 4. extremely
<table>
<thead>
<tr>
<th>How much does this problem affect your overall quality of life?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>I feel as if I let people down</td>
</tr>
<tr>
<td>I find it hard to mix with people</td>
</tr>
<tr>
<td>I feel worn out even when I’ve not done anything</td>
</tr>
<tr>
<td>There are times when I feel very low</td>
</tr>
<tr>
<td>I avoid responsibilities if possible</td>
</tr>
<tr>
<td>I avoid mixing with people I don’t know well</td>
</tr>
<tr>
<td>I feel as if I’m a burden to people</td>
</tr>
<tr>
<td>I often forget what people have said to me</td>
</tr>
<tr>
<td>I find it difficult to plan ahead</td>
</tr>
<tr>
<td>I am easily irritated by other people</td>
</tr>
<tr>
<td>I often feel too tired to do the things I ought to do</td>
</tr>
<tr>
<td>I have to force myself to do all the things that need doing</td>
</tr>
<tr>
<td>I often have to force myself to stay awake</td>
</tr>
<tr>
<td>My memory lets me down</td>
</tr>
</tbody>
</table>
For the following questions put a cross on the line that best marks how you feel. You can put a cross anywhere on the line.

How much would you say problems with memory or concentration affect your day to day life?

No effect on quality of life ———————————————————— Significant effect on quality of life

How much would you say problems with tiredness and lack of energy affect your day to day life?

No effect on quality of life ———————————————————— Significant effect on quality of life

How much would you say problems with feeling tense or worried affect your day to day life?

No effect on quality of life ———————————————————— Significant effect on quality of life

How much would you say problems with socializing and being with other people affect your day to day life?

No effect on quality of life ———————————————————— Significant effect on quality of life

How much would you say problems with self-confidence affect your day to day life?

No effect on quality of life ———————————————————— Significant effect on quality of life
The following section asks some questions about your background

In what year were you born?
(please write in the box)

Are you …

Male ☐ Female ☐

Which of the following best describes your main activity?
(please tick one box)

- employed or self-employed ☐
- retired ☐
- housework ☐
- student ☐
- seeking work ☐
- other (please specify) ____________________________ ☐

Did your education continue after the minimum school leaving age?

Yes ☐ No ☐

Do you have a degree or equivalent professional qualification?

Yes ☐ No ☐

With whom are you presently living?

- Live alone ☐
- Spouse/partner ☐
- children ☐
- Parents ☐
- Other (e.g. friend, sibling, relative) ☐ please specify ……..

Do you have any children living at home with you?

Yes ☐ No ☐

If yes, how many ____________________________
Do you have any long-standing illness, disability or infirmity?

Yes ☐ No ☐

Do you need assistance with any of the following daily life activities?

shopping/errands ☐ grooming/dressing ☐

housekeeping ☐ personal hygiene ☐

other activities ☐

please describe below

In general, how would you say your health has been?

Excellent ☐ Very Good ☐ Good ☐ Fair ☐ Poor ☐

Are you currently receiving treatment for any of the following problems:

Musculo-skeletal problems (such as arthritis, rheumatism) ☐
Respiratory problems (such as asthma or emphysema) ☐
Heart or circulatory problems (such as angina or high blood pressure) ☐
Endocrine problems (such as diabetes or thyroid disorder) ☐
Gastrointestinal or digestive problems (such as stomach ulcer) ☐
Genito-urinary problems (such as kidney or bladder disorder) ☐
Psychological health problems (such as anxiety or depression) ☐
Cancer ☐
Gynaecological or reproductive problems ☐
Blood problems (such as anaemia) ☐
Eye/nose/ear problems ☐
Skin problems (such as eczema) ☐
Other (please specify below) ☐

The space below has been left for any comments you would like to make about the questionnaire.

Did you find filling in this questionnaire …

Very difficult ☐ Fairly difficult ☐ Fairly easy ☐ Very easy ☐

Thank you for your time
Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy 66

Editor: The Dean of the Faculty of Pharmacy

A doctoral dissertation from the Faculty of Pharmacy, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy”.)