University of Sheffield

ScHARR

School of Health and Related Research

CLINICAL AND COST EFFECTIVENESS OF RECOMBINANT HUMAN GROWTH HORMONE (SOMATROPIN) IN ADULTS

PREPARED ON BEHALF OF THE NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE

Nick Bansback
John Brazier
Jim Chilcott
Eva Kaltenthaler
Myfanwy Lloyd-Jones
Suzy Paisley
Stephen Walters

10 April 2002

Please note:
Pharmacia & Upjohn submitted some information to the National Institute for Clinical Excellence in confidence and references to this information have been removed from the report. However, it should be noted that the Institute's Appraisal Committee had access to the full report when drawing up their guidance on the use of recombinant human growth hormone (somatropin) in adults.

CONTENTS

FIGURES

EXECUTIVE SUMMARY

This report addresses the issues identified by the NICE Appraisal Committee on the use of recombinant human growth hormone (somatropin) in adults with growth hormone (GH) deficiency, namely:

- inconsistencies in the evidence on short-term quality of life benefit
- the validity of the economic model submitted by Pharmacia & Upjohn (PU) predicting long term outcomes from Framingham study.

Many of the studies of growth hormone replacement therapy are of poor quality. The randomised controlled trials (RCTs) are generally small and of short duration and their reporting quality is poor. The quality of the uncontrolled studies has not been formally assessed, but again is poor, with very few providing any information relating to withdrawals or adverse events.

The search found no randomised controlled trials of somatropin replacement therapy which reported quality of life outcomes and which were not included in the Southampton review, but the current review has been able to use additional data from those trials. There has been little change to the overall results from the RCTs. For the most commonly used measure of quality of life, the Nottingham Health Profile (NHP), only the social isolation dimension was found to be statistically significantly more improved in the intervention group. Whilst energy and emotional reaction also show small differences in favour of the intervention, pain, sleep and physical mobility scores tend to favour placebo. However, these RCTs may not provide a good indication of the effectiveness of somatropin because of their poor methodological quality and because, in many of the trials, the comparatively good quality of life of patients at baseline meant there was little scope for improvement.

Two RCTs which have only been published in abstract form used the Quality of Life Adult Growth Hormone Deficiency Assessment (QoL-AGHDA) questionnaire. This is a disease-specific measure of quality of life (QoL), and may therefore be more sensitive to response to treatment than generic measures. It is a 25-point scale in which a higher score corresponds to worse QoL.

The full results of the McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

The observational, prospective, uncontrolled studies found large before-and-after improvements in quality of life scores. The measure most commonly used in these studies was the QoL-AGHDA which, after 12 months' follow-up, showed mean changes of 3.7 points from the baseline prior to starting treatment. The magnitude of the gain varied between studies, ranging from 2.8 to 7.2 points. It has been suggested that the size of the change is associated with the baseline scores. However, these studies may suffer from bias due to the placebo effect, regression to the mean and patient selection at entry.

The observational evidence, along with the anecdotal evidence and the high adherence rates to somatropin, suggests that this intervention does improve the quality of life of some patients. The observational data probably over-estimate the extent of the gain, but there is no means of quantifying the size of any bias.

The economic submissions by the industry estimated the short-term quality of life impact of somatropin therapy as:

Eli Lilly	0.16 QALYs
Novo Nordisk	0.126 - 0.376 QALYs
PU	0.02 – 0.12 QALYs.

The validity of the assumptions underlying these estimates is discussed further, but all show significant weaknesses and are consistently overestimated.

The benefits of somatropin, in terms of improved lipid levels and bone mineral density (BMD), were investigated for their economic impact. Whilst the evidence shows an improvement in lipid levels as a result of somatropin treatment, the evidence for a reduction in the risk of fractures is weaker. The long-term implications of these improvements are very small and have very little economic impact. This is demonstrated in the ScHARR analyses, and in both the PU and Novo Nordisk models. For instance, any potential long-term mortality benefits constitute less than 1% of the total QALY benefits, and the effect on the incremental cost effectiveness ratio (ICER) of removing these long-term clinical effects is to change the ICER from £51,457 to £51,617.

The time horizon of the analysis has some effect on the ICER. This is primarily due to differential discounting of the short-term clinical benefits and costs. The ScHARR analysis therefore uses an analytical time horizon of four years.

The economic submissions by the industry estimated the ICER for somatropin treatment as:

Lilly	£25,700 - £30,600 per QALY
Novo Nordisk	£13,600 - £22,400 per QALY
PU	£27,500 – £37,600 per QALY.

The most optimistic assumption for the utility gain within the ScHARR analysis gives an incremental cost per QALY in the region of £52,000. The actual utility gain is disputed, but it is clear that the cost effectiveness is very sensitive to this figure. The estimate of £52,000 per QALY is at the floor of the estimate.

The full conclusions on the health economics are not reported here because they were based upon the full results of the McKenna trial (1) which was submitted to NICE in confidence by Pharmacia & Upjohn.

1. BACKGROUND: PROJECT HISTORY

The use of recombinant human growth hormone (somatropin) in growth hormone-deficient adults was identified by the Department of Health as a priority area for appraisal by NICE and was included in the 4th wave of topics announced in November 2000. The National Coordinating Centre for Health Technology Assessments therefore commissioned Southampton University to undertake a Technology Assessment Report focussing on the clinical and cost effectiveness of this intervention. In addition to patient and clinical groups industrial submissions were obtained from:

- Eli Lilly
- Novo Nordisk
- Pharmacia.

The NICE Appraisal Committee first met to consider this topic on 22 November 2001. On the basis of the submissions presented to it, the Committee felt unable to make a Preliminary Appraisal Determination. The following issues were identified by the Committee as remaining unclear:

- inconsistent evidence on short-term quality of life benefit
- validity of the economic model submitted by Pharmacia & Upjohn predicting long term outcomes from Framingham study.

In order to attempt to resolve these issues, the ScHARR Rapid Reviews Group was asked to clarify these issues. An initial proposal was submitted to NICE and the SRRG commenced work on the assessment immediately. The first piece of work undertaken was to analyse the relative impacts of the claimed short- and long-term effects of human growth hormone. This analysis demonstrated immediately that the long-term effects had very little impact on the economics of treatment. This analysis was forwarded to NICE immediately and the SRRG review was refocused on clarifying the remaining issues regarding the short-term impacts of the treatment.

2. EFFECTIVENESS EVIDENCE: METHODS

2.1 Identification of studies

2.1.1 Search methods

The aim of the search was to provide as comprehensive a retrieval as possible of studies relating to growth hormone deficiency and quality of life. The search strategy was designed to pick up quality of life studies relating to both treated and untreated populations.

2.1.2 Sources searched

Nine bibliographic databases were searched providing coverage of the biomedical, psychology and health economic literature. A list of the databases is given in Table 1, Appendix 1. In addition the Southampton review[1] was handsearched.

2.1.3 Keyword strategies

Sensitive keyword strategies using freetext and, where available, thesaurus terms were developed. Strategies combined terms relating to growth hormone deficiency and quality of life. The quality of life component included general quality of life terms (e.g. quality of life, qol, hrqol), generic quality of life instruments (e.g. SF-36, EQ-5D, Nottingham Health Profile) and condition specific instruments (e.g. AGHDA). The list of quality of life instruments provided in the Southampton review was used to develop the keyword strategies. Keyword strategies for all databases are given in Appendix 1.

2.1.4 Search restrictions

Date and language restrictions were not used. The search retrieval was not limited to specific study designs. Searches were undertaken in January 2002.

2.2 Inclusion and exclusion criteria

Controlled and uncontrolled trials and observational studies were included which reported quality of life, assessed over a period of time using quantitative measures, in adults aged 18 years or over with growth hormone deficiency who were either untreated or were treated with growth hormone in any dose.

Because many individuals with isolated idiopathic GHD in childhood show normal GH status when reassessed in adult life,[2] studies were excluded which did not reassess at study entry the GH status of subjects with childhood-onset GHD.

2.2.1 Study selection

Studies identified by the search strategy were assessed for inclusion as follows. Titles were initially considered for inclusion. If the titles suggested that the studies were relevant, the abstracts were then considered and, if these also appeared relevant, the full texts were then reviewed.

Relevant references from the retrieved articles were also included in the review.

2.2.2 Quality assessment

RCTs were assessed for quality using the Jadad scale.[3]

2.3 Data synthesis

2.3.1 Meta-analysis

The meta-analysis used the same assumptions and methods as the Southampton report. Full details are given in Appendix 1 of that report. The summary statistic generated was a weighted mean difference using a random effects model. Studies were weighted by the inverse of their variance.

The meta-analysis was constructed using STATA v 7.0 software (STATA Corp (2001) STATA Statistical Software: release 7.0, College Station, Tx: Stata Corporation).

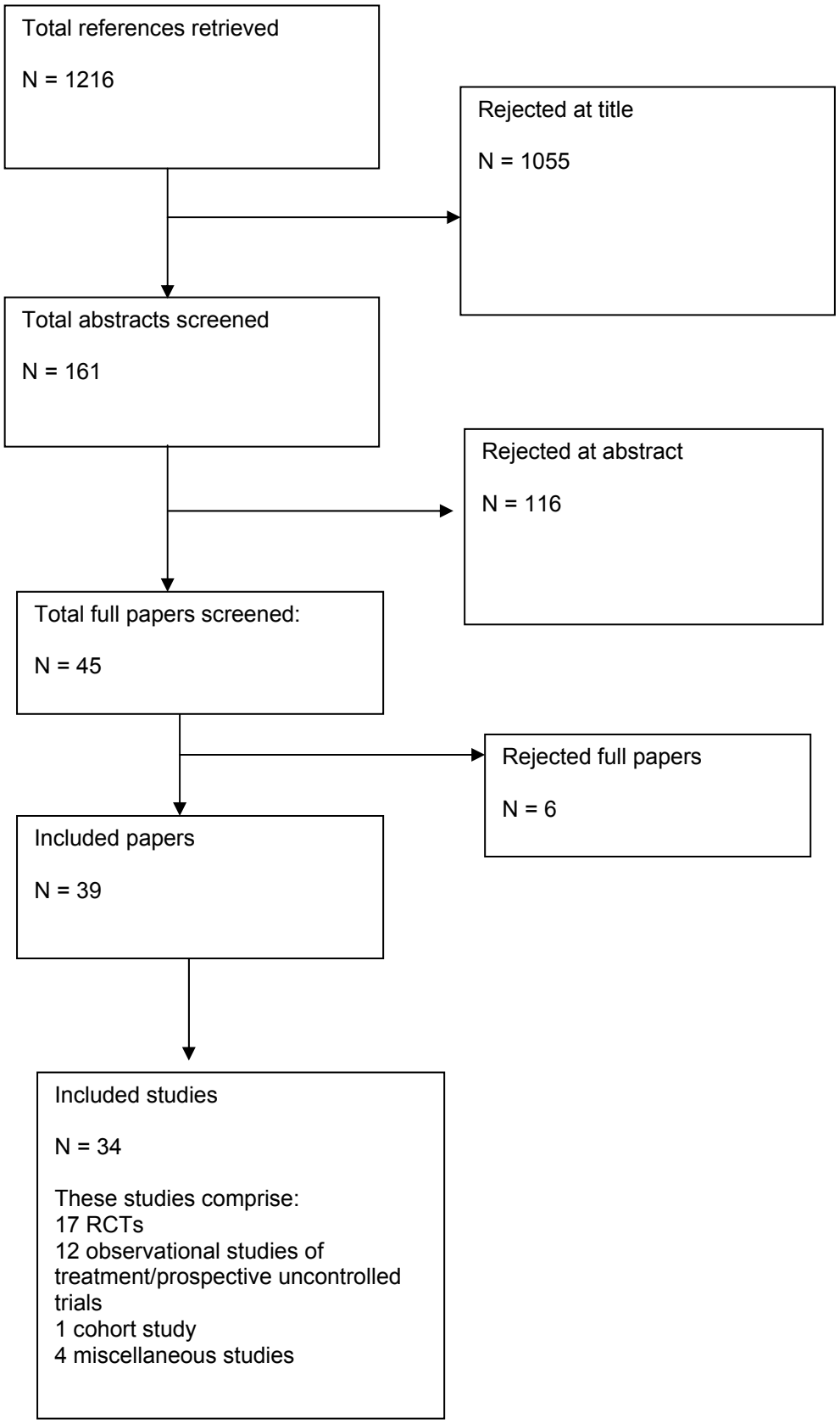
2.4 Results

The electronic literature searches identified 1206 potentially relevant articles. A further 10 potentially relevant articles were identified from citations.[4-13]

From their titles, 161 of these articles appeared potentially relevant; when their abstracts were read, this figure was reduced to 45; and 39 articles were retained when the full text had been reviewed. These 39 articles related to 34 relevant studies (see figure 1) – 12 conventional RCTs,[9;10;14-23] five cross-over RCTs,[4;7;12;24;25] eight prospective uncontrolled studies,[26-33] four observational studies of treatment,[34-37] one cohort study in untreated patients,[38] and four miscellaneous studies.[8;33;39;40]

Details of all these studies are provided in Appendix 2.

FIGURE 1 Quality of life: summary of study selection and exclusion



In a number of these studies, the quality of life data were incomplete. Attempts were made to contact 15 authors in relation to 16 studies.[4;5;8-10;12;14;15;17;20;22-24;30;37;39] Four authors responded and three provided additional relevant data.

Excluded studies are listed in Appendix 3, with reasons for exclusion.

3. EFFECTIVENESS EVIDENCE: RESULTS

3.1 Quantity and quality of trials of somatropin in adult GH deficiency

No new RCTs have been published since the Southampton review was undertaken. Details of the available trials are presented in Tables 2 and 3, with the crossover trials listed separately. The quality of these RCTs was reviewed in detail in the Southampton review. This section summarises their key features.

Table 2: RCTs included in the assessment of effectiveness of somatropin in GH-deficient adults.

Included Studies	Number of Patients	Outcomes (QoL scales)				
		GHQ	HDS	NHP	PGWB	Other
Attanasio 1997[14]	173			✓		
Baum 1998[15]	40	✓		✓	✓	MMPI-2
Beshyah 1995[16]	40	✓				CPRS
Cuneo 1998[17]	163			✓		GHDQ
Deijen 1998[18]	48					HSCL, POMS, STAI
Giusti 1998[19]	26		✓			KSQ
McGauley 1989[20]	24	✓		✓	✓	
McKenna 1997(2)[9]	30			✓		QoL-AGHDA
McKenna 1997(1)[10]	69			✓		QoL-AGHDA
Soares 1999[21]	10		✓			SADS, BDI
Verhelst 1997 [22]	148			✓		
Wallymahmed 1997[23]	35			✓		LFS, HAD, SES, MFS

Table 3. Crossover RCTs included in the assessment of effectiveness of somatropin treatment in GH-deficient adults.

Included Studies	Number of Patients	Outcomes (QoL scales)				
		GHQ	HDS	NHP	PGWB	Other
Bengtsson 1993[4]	10					CPRS, SCL90 psychiatric interview
Burman 1995[24]	36			✓	✓	HSCL-56, spouses questionnaire
Degerblad 1990[7]	6					POMS, SMQ, psychiatric interview, finger tapping
Florkowski 1998[25]	20					DSQ, SCL-90, SAS
Whitehead 1992[12]	14			✓	✓	

The reporting quality of many RCTs, as assessed using the Jadad scale, was poor. Few reported the method of randomisation or double-blinding, and some failed to provide sufficient information relating to withdrawals. Differences between the Jadad scores given by the Sheffield and Southampton teams are tabulated below. The crossover trials were also generally inadequately reported in that data from all patients who received active treatment, whether as the first or second intervention, were pooled rather than reported separately, and similarly for placebo.

Table 4: Southampton and Sheffield Jadad scores, with reasons for discrepancies (if any)

Included Studies	Sheffield Jadad score	Southampton Jadad score	Reason for discrepancy
Attanasio 1997	3/5	2/5	Description of withdrawals felt by Sheffield to be adequate
Baum 1998	5/5	5/5	-
Bengtsson 1993	5/5	4/5	Description of method of randomisation felt by Sheffield to be adequate
Beshyah 1995	4/5	3/5	Description of method of double-blinding felt by Sheffield to be adequate
Burman 1995	2/5	2/5	-
Cuneo 1998	4/5	4/5	-
Degerblad 1990	3/5	3/5	-
Deijen 1998	1/5	2/5	Description of withdrawals felt by Sheffield to be inadequate
Florkowski 1998	3/5	1/5	Description of method of double-blinding and of withdrawals felt by Sheffield to be adequate
Giusti 1998	1/5	2/5	Description of withdrawals felt by Sheffield to be inadequate
McGauley 1989	5/5	2/5	Description of method of randomisation and double-blinding and of withdrawals felt by Sheffield to be adequate. Some of this information was provided in a supplementary paper [11]
McKenna 1997(2)	2/5	?1/5	Statement of double-blinding available to Sheffield team
McKenna	2/5	1/5	Statement of double-

1997(1)			blinding available to Sheffield team
Soares 1999	3/5	3/5	-
Verhelst 1997	2/5	3/5	Description of withdrawals felt by Sheffield to be inadequate
Wallymahme 1997	3/5	3/5	
Whitehead 1992	3/5	2/5	Description of method of double-blinding felt by Sheffield to be adequate

Many of the studies are small, the majority of trials having between 21-40 participants. The duration is typically 6 months, and only one study (in men only) was longer. Participants included a mix of patients with adult- and childhood-onset GH deficiency. Whilst the dosing varied between the studies, a key feature is that the dosage of somatropin replacement was determined by the patient's weight (although in the study by Baum *et al.* the dose was subsequently adjusted according to serum IGF-1). Current practice is to use a more sensitive method based on titration of the somatropin dosage according to response. The earlier method resulted in obese patients being over-dosed and others being under-dosed. This has potentially important implications for adverse events and continuation with therapy.

Quality of life was assessed in these trials using self-completed questionnaires (see the Southampton Report). Whilst this approach can be subject to biases, self-report is recognised as the best quantitative approach to measuring quality of life where it is feasible.[41] 23 different types of instrument were used across the trials. The most frequently used was the Nottingham Health Profile, followed by the General Health questionnaire (GHQ), whilst other measures included the QoL-AGHDA, HAD, and Psychological General Well-Being Schedule (PGWB). The remainder were used in one study. Only the QoL-AGHDA and the GHDQ were developed for patients with GH-deficiency.

3.2 Quantity and quality of observational and prospective uncontrolled studies of somatropin replacement in patients with GH-deficiency

For reasons discussed earlier, it was decided to review the published observational prospective uncontrolled studies of somatropin replacement. Thirteen studies met the inclusion criteria, and these are listed in Table 5. It is clear from the publication dates of these studies that they were undertaken after the trials reported above. They vary widely in size with between 10 to 972 patients. The period of follow-up is broadly comparable with the trials, at between 3 to 12 months, except for one study that is a follow-up at 10 years of patients entered into an earlier somatropin trial.[35]

Table 5: Observational and prospective uncontrolled studies

Study	Number of patients	NHP	QoL-AGHDA	PGWB	STAI	KIMS data base
Ahmad 2001[26]	46		√			
Bengtsson 1999[34]	665		√			√
Bulow & Erfurth1999[27]	10		√			
Davies 2000[28]	39		√			
Drake 1998[29]	50		√			
Gibney 1999[35]	21	√				
Hayes 1999[30]	12	√				
Hernberg-Stahl 2001[36]	304		√			√
Monson 2000[37]	972		√			√
Murray 1999[31]	65		√	√		
Riva 1993,[32] Sartorio 1995[42]	8				√	
Wiren 1998[33]	71	√		√		

Although some studies only reported on patients with either adult-onset or childhood-onset GH deficiency, the majority reported on a mix of adult and childhood onset. The method of determining the somatropin dosage varies between studies, with some allowing considerable discretion.

These studies used just four quality of life scales compared to the 23 used in the RCTs. The NHP was used in four studies, the Psychological General Well-Being Schedule (PGWB) in two and the State-Trait Anxiety Inventory (STAI) in one. The most widely-used was the condition specific QoL-AGHDA, which was used in eight of the 12 studies.

The quality of these studies as observational studies has not been formally assessed. However, they are poorly reported. Very few provide information relating to withdrawals or adverse events. Information is also lacking on methods of recruitment, take-up of somatropin replacement therapy and compliance with treatment. The dosages used also vary considerably.

3.3 Assessment of the clinical effectiveness of somatropin in adults as measured using quality of life in trials

This section is limited to reporting new results over and above those already available in the Southampton report. For the RCTs it only presents a re-analysis of the NHP data to incorporate newly available data, and an analysis of the QoL-AGHDA trial results which have been made available to the ScHARR team by Pharmacia and Upjohn. For an analysis of the remaining instruments, the reader should see the Southampton report. This section also presents the results of the review of observational studies.

3.3.1 Nottingham Health Profile

The Nottingham Profile (NHP) is a general measure of health-related quality of life.[43] It assesses six dimensions of health: physical mobility, pain, emotional reaction, energy, sleep, and social isolation. Each dimension contains a set of items each with dichotomous yes/no responses. A score is generated for each dimension between zero and 100 using weights produced by the developers, where a high score indicates a worse quality of life.

The NHP has been used on a wide range of conditions with varying success.[41] It has been shown to be reliable and valid in terms of sensitivity to between-group differences for many different conditions. The NHP has shown significant differences between GH-deficient patients and general population controls.[20;44-46] However, there is concern regarding its sensitivity in populations with mild to moderate health problems. One defect of the NHP is that large proportions of respondents score zero across its dimensions, indicating no health problem. It has been shown that many people with zero scores on the NHP have health problems according to other instruments and indicators of health.[47] According to the NHP, patients scoring zero cannot get any better, which has important implications for the responsiveness of the NHP to improvements.

The original Southampton review only reported on the results from four of the ten published trials which used the NHP.[15;17;23;24] Five studies were excluded because they did not present all the NHP data. Two of these were trials that have only been reported in conference abstracts and did not show analysis by treatment.[9;10] The data from one of these trials[9] have been made available to the Sheffield team. Whitehead has sent additional data and these are incorporated into the analysis. Finally, McGauley[20] reported the results for two of the six NHP dimensions and these have been included in the analysis reported below. This leaves McKenna[10] and Attanasio[14] who have been approached for the data needed for the meta-analysis reported below.

The other modifications to the Southampton analysis resulted from two corrections; one was the use of standardised errors around the mean rather than standard deviations from the Baum study[15] and the other was the use of 18 and not 36 per arm in the Burman trial.[24]

The results of the meta-analyses are presented on figures 1-7 and described below.

The meta analysis does not include the results of McKenna (1) trial because it was submitted to NICE in confidence by Pharmacia & Upjohn.

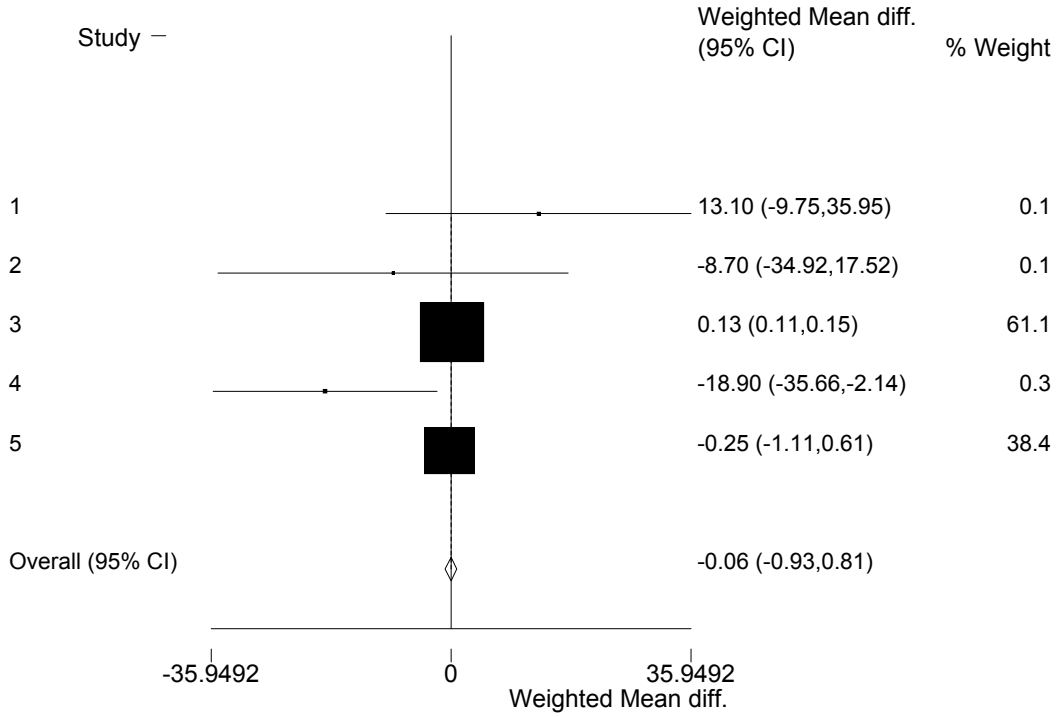
Energy

Five studies reported mean energy scores. The Baum trial found a reduction (*i.e.* improvement) in both arms of the trial, but this did not achieve statistical significance. The Wallymahmed trial found a statistically significant reduction in the intervention group and a non-significant reduction in the control group,

but gave no between-group comparison. The Cuneo trial found a significantly greater reduction in energy score for the control group. Burman showed significant reductions in both arms, though here the difference favoured the intervention group. The additional study by McGauley showed a significantly greater reduction in the intervention than in the control group ($p<0.05$).

The five trials that reported the NHP energy subscale were pooled. There was marked heterogeneity. The overall change estimate score was slightly in favour of the intervention group (whereas the Southampton review had found it to be slightly in favour of the control), but the result was not significant. Somatropin was associated with a very small non-significant gain of -0.06 (95%CI: -0.93 to 0.81).

Figure 2: Meta-analysis of NHP energy dimension score



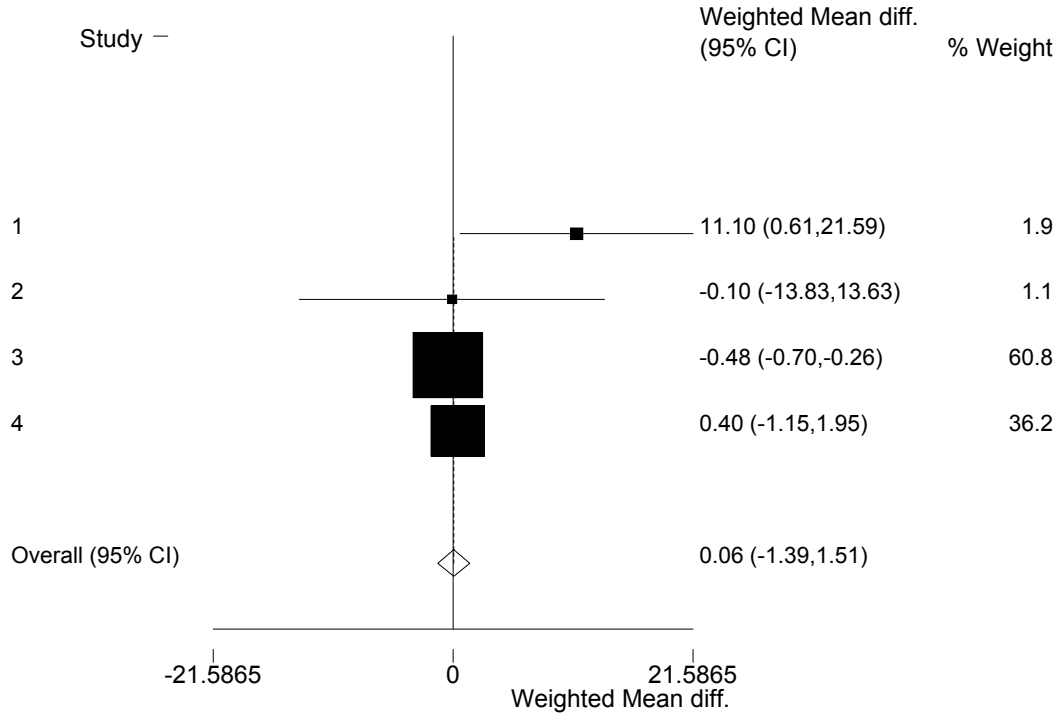
The studies are as follows: 1. Baum, 2. Burman, 3. Cuneo, 4. McGauley, 5. Wallymahmed

Pain

The Baum study found a rise in the pain score (indicating an increase in pain) associated with the intervention and a reduction in the score in the control group. The Southampton report had expressed concern that there were vast differences at baseline which might explain this result. Increases in the pain score were found in both groups in the Wallymahmed trial, but were non-significant. There were no statistically significant changes in the Burman trial. Cuneo found a reduction in the intervention group and a small increase in the control group, but neither change was significant.

The meta-analysis of the pooled scores found a difference in favour of the control group but this is now down to a non-significant difference of 0.06 (95%CI:-1.39,1.51).

Figure 3: Meta-analysis of NHP Pain dimension score



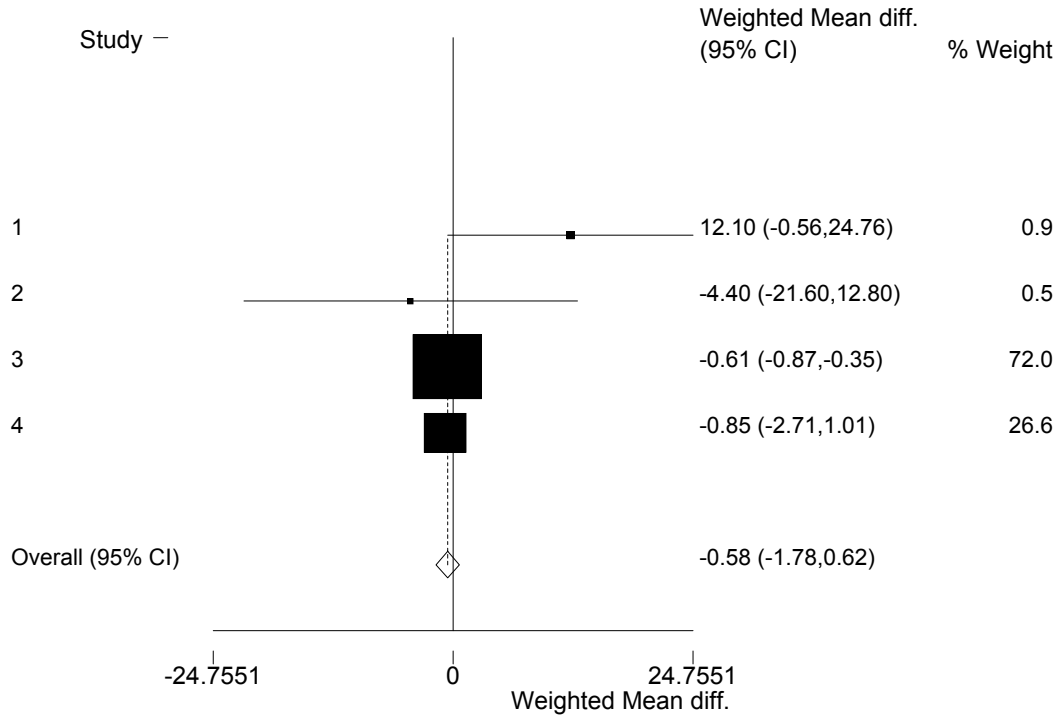
The studies are as follows: 1. Baum, 2. Burman, 3. Cuneo, 4. Wallymahmed

Emotional reaction

Burman found a significant score reduction in the intervention arm and a small non-significant gain for the control. Cuneo and Wallymahmed found small reductions relative to baseline in favour of the intervention. An increase in score was found in the Baum study alongside a small reduction in the control arm.

The pooled analysis gave a small and but not significant difference of 0.58 points in favour of the intervention group (-1.78 to 0.62).

Figure 4: Meta-analysis of NHP emotional reaction dimension score



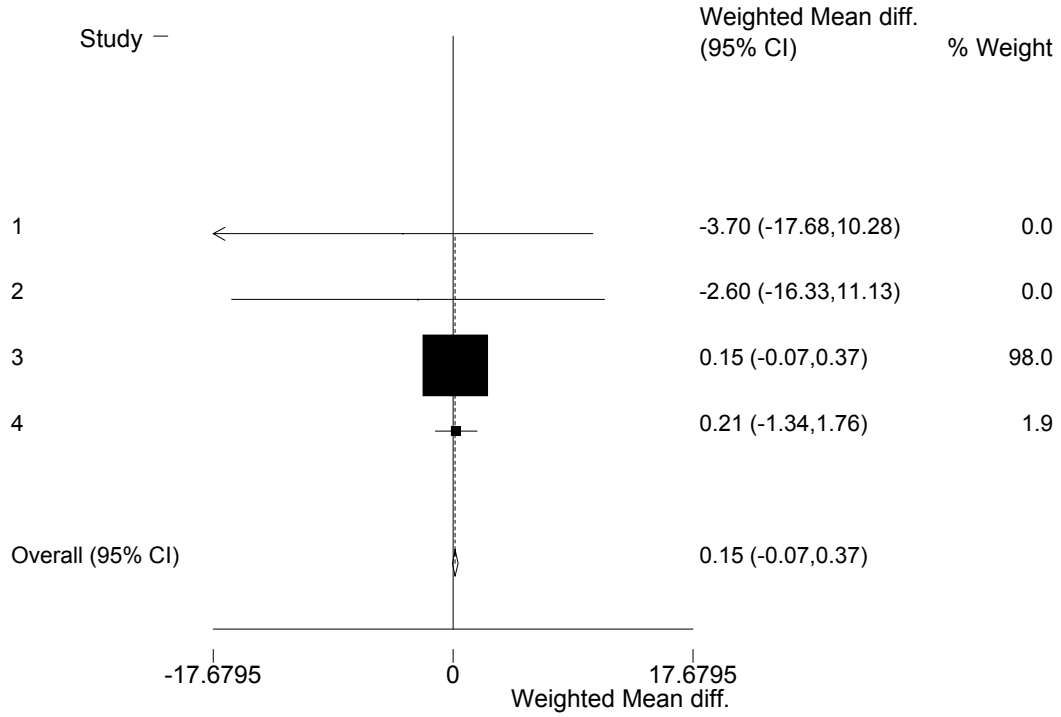
The studies are as follows: 1. Baum, 2. Burman, 3. Cuneo, 4. Wallymahmed

Sleep

There were non-significant improvements in the sleep score in the Baum and Cuneo trials. Wallymahmed found score increases in the intervention group alongside reductions in the control group, while Burman observed a reduction in the intervention group and an increase in score in the control arm, but the differences between the groups were not significant.

The re-analysis has made little change to this result. The summary estimate of the difference in the change in favour of the control group has been reduced to 0.15 (95%CI:-0.07, 0.37).

Figure 5: Meta-analysis of the NHP Sleep dimension score

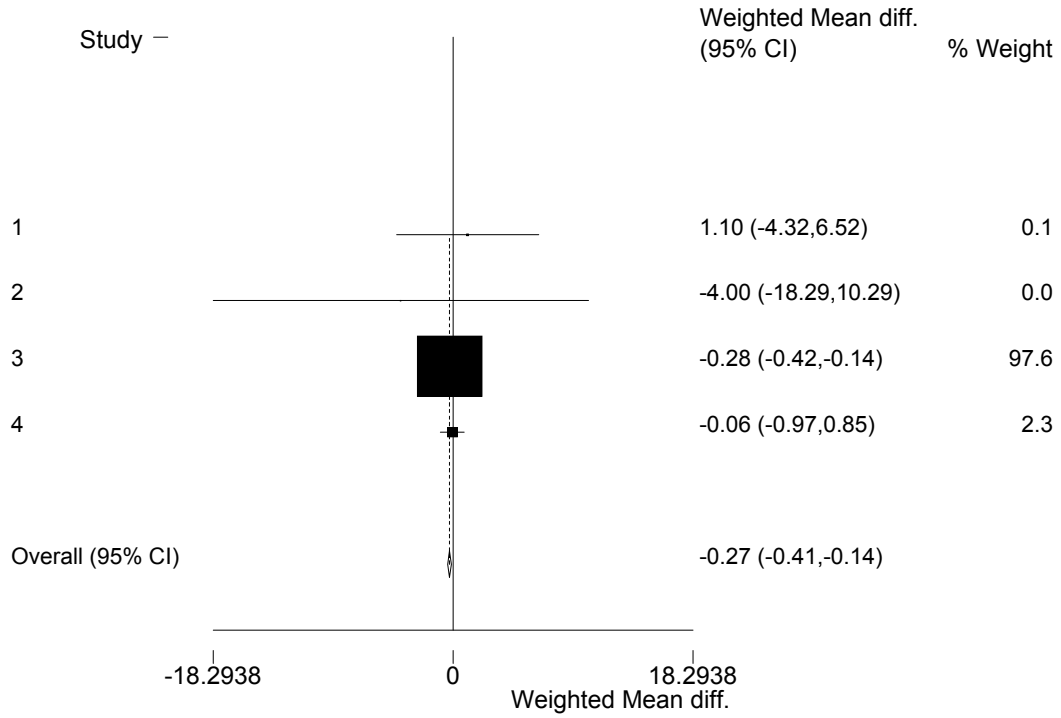


The studies are as follows: 1. Baum, 2. Burman, 3. Cuneo, 4. Wallymahmed

Social Isolation

The re-analysis produced a virtually identical estimate to the Southampton report of the difference in the change of 0.27 in favour of the intervention (95%CI:-0.41,-0.14) that was significant (P<0.05).

Figure 6: Meta-analysis of NHP Social Isolation dimension score



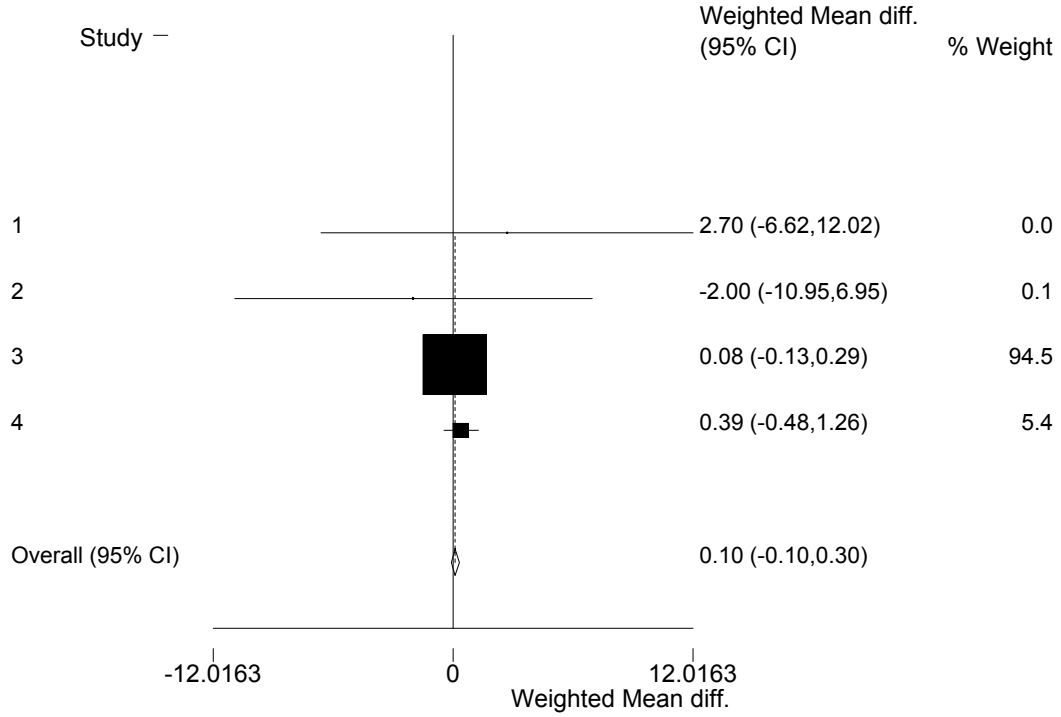
The studies are as follows: 1. Baum, 2. Burman, 3. Cuneo, 4. Wallymahmed

Physical mobility

There were non-significant increases in the physical functioning scores in the Wallymahmed trial in both the intervention and control arms. Baum observed non-significant reductions in both groups. A small increase in score was found in the intervention group of the Cuneo trial and a reduction in the control group. None of these changes were significant.

The re-analysis reduced the difference to a very small and non-significant difference in favour of the control of 0.10 (95%CI:-0.10,0.30).

Figure 7: Meta-analysis of the Physical Mobility dimension score



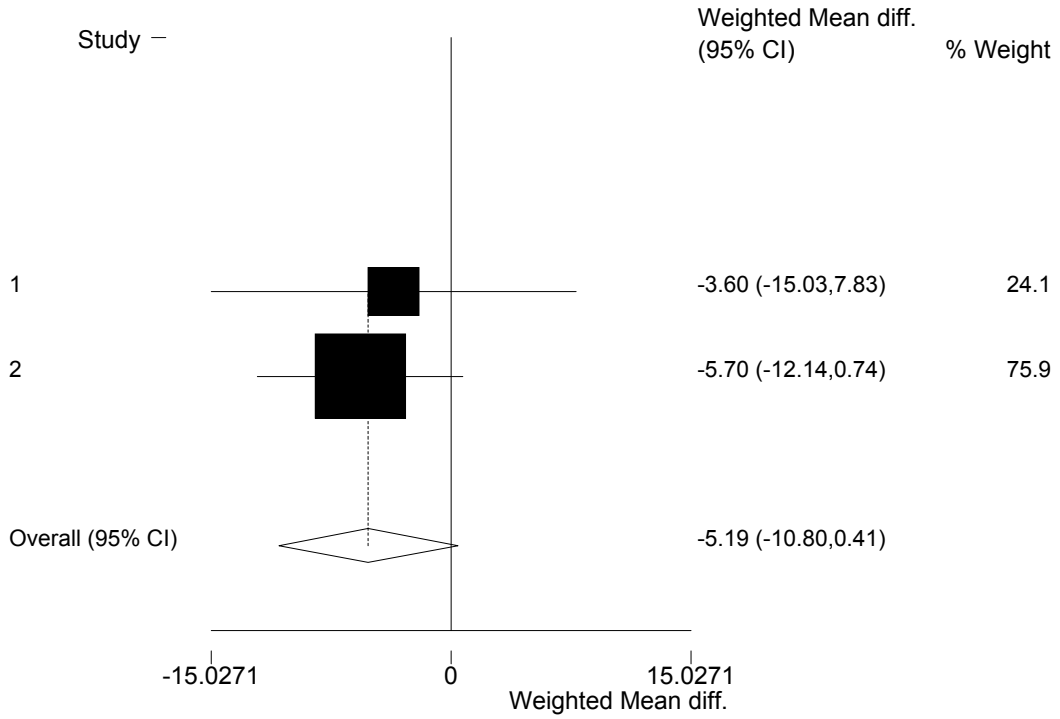
The studies are as follows: 1. Baum, 2. Burman, 3. Cuneo, 4. Wallymahmed

Total NHP score

The NHP was not designed to produce an overall score since the dimensions measure different health domains, and the production of an overall score is therefore not recommended by the authors.[48;49] However, some studies have reported a total score. Two trials, Burman and McGauley, reported mean total scores. Burman showed a significant reduction in the intervention group, but no significant reduction in the control group; the difference between the groups was not significant. McGauley found improvements in both groups, though this was larger in the intervention arm. The summary estimate of the difference in the change for these two studies shows an improvement in

favour of intervention of –5.19 (95%CI: -10.80 to 0.41), which does not quite reach significance (P=0.08).

Figure 8: Meta-analysis of total NHP dimension score



The studies are as follows: 1. Burman, 2. McGauley

3.3.2 QoL-AGHDA

The Quality of Life Assessment of GH Deficiency in Adults (QoL-AGHDA) is a self-completed questionnaire developed to assess the consequences of GHD and somatropin replacement in adults.[50] It was developed from qualitative interviews conducted with GH-deficient patients in the UK[51]. It contains 25 dichotomous yes/no response format items (see Appendix 4). A positive response to each item is assigned a score of one, giving a score range of zero to 25 where a high score indicates a poor QoL.

The measure has been translated into several languages. It has been shown by the developers to have good reliability. Rasch analysis indicated that QoL-AGHDA is unidimensional, with all items contributing to the construct of the quality of life of patients with GH deficiency.[9] It has been shown in a number of studies to discriminate between patients with GH deficiency and control samples of the general population.[38;52] Although one study found it was unable to distinguish between severe GH deficiency and GH excess,[53] this is not problematic in the context of this review. Observational studies have shown that the QoL-AGHDA is responsive to changes following treatment (e.g. Bengtsson *et al.* 1999, Hernberg-Stahl *et al.* 2001 and Monson *et al.* 2000)[34;36;37]

The QoL-ADGHA has been shown to be reliable, valid in terms of between-group comparisons with the general population, and responsive to change following treatment. It would seem to be more appropriate for use in clinical trials of somatropin treatment than the general measures used to date on this patient group. The main problem with the measure is that it does not produce a score that can be used to generate a utility score. However, the Pharmacia and Upjohn submission included a means of mapping from the QoL-AGHDA to utilities that is reviewed below.

Two RCTs used the QoL-ADGHA as an outcome measure. One, undertaken across three centres in Spain, recruited 69 patients,[10] and the other, in the Netherlands, recruited 30 patients.[9] The NHP and QoL-AGHDA were administered at baseline and at 3, 6 and 12 months following the commencement of treatment. Patients were randomised to somatropin (determined by the weight-based dosing method) or placebo for 6 months. After six months, all patients in both groups in each trial were offered somatropin. These studies were both funded by Pharmacia & Upjohn.

Prior to this review, these trials had only been available in the form of conference abstracts.[9;10] These abstracts contained minimal information about the trials and did not present an analysis by treatment group. The sponsor of these studies, Pharmacia & Upjohn, made available to the ScHARR team copies of the documentation relating to these trials and the original QoL-AGHDA and NHP data from the Netherlands study and QoL-AGHDA data from the Spanish study. The results of the analysis of these two data sets are presented below.

The results of the McKenna (1) and (2) trials are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

3.4 Assessment of the clinical effectiveness of somatropin in adults as measured using quality of life in observational studies

3.4.1 Nottingham Health Profile

The study by Gibney *et al*.[35] is a 10 year follow-up of patients recruited into their earlier RCT in which patients in both arms were offered somatropin replacement therapy after the trial was completed. In those who had remained on therapy for seven out of the 10 years, significant improvements were found between baseline and 10 years for the overall score (18.8 to 7.5, P<0.02), energy (43.3 to 5.3, P<0.02) and emotional reactions (19.8 to 3.7, P<0.02) and a non-significant improvement in social isolation. There were small and non-significant deteriorations in sleep (5.2 to 14.5), pain (0.0 to 4.8) and mobility (4.9 to 5.9).

Wiren *et al.*[33] found significant improvements at 6 months in the overall score (11.9 to 7.3, (P<0.01), energy (27.9 to 13.0, 0.001), emotional reactions (10.5 to 7.2, P<0.01) and social isolation (9.1 to 5.9, P<0.01). The remaining two dimensions were improved, but the change was small and not statistically significant. These improvements were maintained at 12 months. The authors

argued that, since they found no significant changes between the first assessment 6 months prior to treatment and the second at baseline, the improvements observed during the administration of somatropin were due to therapy. Hayes presents significant improvements between baseline and 12 months in the overall score (on a scale of 0-600, the change was 119.7 to 48.9, P<0.05).[30] This study also failed to report improvements by dimension.

3.4.2 QoL-AGHDA

Eight observational studies of somatropin reported QoL-AGHDA scores. Three of these reported results from the largest observational dataset of somatropin-treated patients, the KIMS database.[34;36;37] These three studies account for the vast majority of patients in published studies.

KIMS is the Pharmacia and Upjohn International Metabolic Database and pharmacoepidemiological survey of adult GH-deficient patients receiving somatropin replacement therapy. Enrolment of patients began in 1994, and has included over 4000 treated and untreated patients from 25 countries. The patients are seen at clinic at least once per year. Data are collected according to KIMS guidelines,[37] ensuring comparability of data collection and recording. The database is maintained by an independent group of international academics and clinicians.

The other observational studies come from a mix of sources: Ahmad *et al.*[26] compared two brands of somatropin (Genotropin and Humatrope), a study by Bulow and Erfurth[27] was an uncontrolled prospective comparison of different dosing methods in patients with childhood-onset GH deficiency, Davies *et al*.[28] examined three-month changes following somatropin replacement therapy in patients with adult-onset GH deficiency, while Drake *et al.*[29] studied different titration dosage methods, and Murray *et al.*[31] studied influences on quality of life score changes.

Out of the three studies based on the KIMS database, Bengtsson *et al.* showed mean changes in QoL-AGHDA of 7.4 to 4.6 at 12 months in men and 9.8 to 5.0 in women. These were significant for median changes (P<0.001), though the paper does not report the significance of changes in the mean scores. Hernberg-Stahl *et al.* showed significant improvements in the QoL-AGHDA scores from 9.0 to 6.2 at 12 months (P<0.001). Monson *et al.* reported an overall improvement of 3.1 with significant improvements in younger men and women (P<0.001) and older men (P<0.05) but not in older women. The authors suggested that the failure of the score-improvement in the last group to achieve statistical significance might have been due to the small number of patients in this group.

Ahmad *et al* reported a significant improvement in the QoL-AGHDA score 13.3 to 10 (P<0.001). Significant gains were also found by Bulow and Erfurth of 6 to 2 P<0.001), Davies of 10 to 7 (P<0.001), Drake of 14.2 to 7.0 and Murray of 15.3 to 10.4 (P<0.01).

Since these results are not from RCTs, it was decided not to undertake a formal meta-analysis. Nonetheless, it is useful to have some estimate of the size of change being observed in these studies (regardless at this stage of attributability). The average change in QoL-AHDGA scores across the seven studies, weighted by number of patients, is 3.7. This change is noticeably larger for the non-KIMS studies, which had an overall change of 4.9.

3.4.3 Other measures

The Psychological General Well-Being Schedule (PGWB) was used alongside the QoL-AGHDA in the study by Murray *et al.*[31] and alongside the NHP in Wiren *et al.*[33] In the study by Murray *et al.*, there were significant improvements across all the dimensions ($P<0.05$). Wiren *et al.* showed significant changes in well-being, vitality and overall scores versus baseline ($P<0.01$). Riva *et al.*[32] report a reduction in transient anxiety levels, though not in anxiety as a personality trait. It has not been possible to establish the significance of this finding.

3.5 Discussion of quality of life evidence

The search found no RCTs of somatropin replacement therapy reporting quality of life outcomes which had not been included in the Southampton review. However, the current review has been able to use additional data from these studies. The NHP RCT data have been re-analysed along with newly available QoL-AGHDA RCT data.

The impact on the most commonly used quality of life measure in these trials, the NHP, has been ambiguous. The re-run of the meta-analysis has resulted in fewer differences in favour of the control, though the overall picture is still not favourable to somatropin replacement. Only the social isolation dimension is significantly better in the intervention group, and the difference is small. Whilst there are also small differences in favour of the intervention for energy score and emotional reaction, the pain, sleep and physical mobility scores were in favour of placebo. The overall score is better for the intervention, and approaches significance, but an overall score was only calculated in the two trials that favoured the intervention arm in other dimensions. The meta-analysis of the HAD results in the Southampton report also found a non-significant difference in favour of the treatment. The analysis of the two trials using the condition-specific QoL-AGHDA found rather more support for a beneficial impact of somatropin on QoL.

The results of the McKenna (1) and (2) trials are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

On the basis of this review, it is not clear overall whether somatropin therapy confers benefits to patients in terms of their quality of life. However, these trials have a number of shortcomings. They varied considerably in terms of quality, with just two out of the 17 trials achieving Jadad scores of 4 or 5; two scored 3 and the remainder scored 2 out of 5. The number of participants was often low, ranging from 6 to 164. This is very worrying since the baseline

scores were often very different between the arms of the trials. The length of follow-up was also limited. Most trials were six months in duration or less, and only one continued for 18 months. Furthermore, the RCTs used a dosage regimen determined by the patient's weight rather than one based on a titration technique. This may have important implications for adverse events and willingness to continue therapy.

Whilst a large range of quality of life measures were used, only two were developed for this condition: the QoL-AGHDA, which was used in two of the 17 trials, and a questionnaire called the GHDQ, which was used in the largest RCT[17], but about which few details are available. Most of the measures were probably inappropriate for measuring response to therapy in this condition. The NHP, for example, is well known for having a 'ceiling' effect, whereby a large number of people have the highest score of zero and hence cannot get better.

The most important potential limitation of these trials for estimating the effect of somatropin replacement is the selection of patients. The baseline NHP scores reported in the trials were often very low, indicating a comparatively healthy population. These patients may not be typical of the general population of patients with hypopituitarism. NHP scores for a general population sample aged 18-74 have been found to be as follows: energy (16.8), emotional reaction (12.2), pain (6.4), physical mobility (4.1), social isolation (7.0) and sleep (12.9).[47] Two of the five trials contributing to the meta-analysis, Cuneo and Wallymahmed, had scores across the dimensions of below 3.0 (on the zero to 100 scale of the NHP, where zero is the best score). The patients in the study by Baum *et al.* had higher scores than this, but they were similar to the general population scores. Burman had the highest scores at baseline, particularly for energy (37.1) and emotions (23.1), which are thought to be two of the most relevant dimensions for this group. This study generally favoured the intervention, but improvements were also seen in the placebo group, and the difference between the groups was not statistically significant.

The Spanish trial using the QoL-AGHDA has been selected to provide an estimate of the potential benefit of treatment. The reason for this is that the QoL-AGHDA is a measure of QoL which may be more sensitive to response to treatment than the generic measures used in the other trials.

The full results of the McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

The changes reported before and after treatment in the observational and prospective uncontrolled studies were less ambiguous. All studies found significant improvements in quality of life scores. The most commonly used measure has been the QoL-AGHDA with mean changes of 3.7 between the baseline prior to starting treatment and 12 months after. The precise magnitude of the gain varied between studies, ranging from 2.8 to 7.2. There is a suggestion from the studies that the size of the change is associated with the baseline scores. Those patients with worse scores were found to

experience the largest improvement.[30;31] It is also noticeable from the mean scores reported in the observational studies that the two studies with the largest changes are those with the worst baseline scores.[29;31]

However, the main weakness with these studies is the problem of attributing the change to somatropin replacement. It is not possible to infer what the quality of life of these patients would have been without treatment with somatropin. There are well known potential sources of bias in observational evidence, including the placebo effect, regression to the mean and patient selection bias.

The risk of a placebo effect is less likely given the comparatively long follow-up in these studies of up to several years. However, some of the RCTs found significant improvements in the placebo group. It is also possible that the change is the result of a regression to the mean whereby patients are recruited at a low point in their condition and therefore would be expected to improve. Of most concern is the problem of patient selection. Patients diagnosed with GH deficit are initially placed on a six-month clinical evaluation course of somatropin replacement in order to see whether they respond to the treatment in terms of clinical parameters and quality of life. A significant number decide not to pursue somatropin replacement. After the clinical evaluation, a number would drop out each year. As a result, the patients in the observational study are a select group who have apparently responded to therapy. It is possible that the improvements, particularly in quality of life, might have occurred without treatment.

The availability of control evidence is very limited. The study by Gibney *et al.* used the NHP to compare treated patients with a control group in a 10-year follow-up study, and found significant improvements in the treated group in the energy and emotional reactions dimensions of the NHP which were not seen in the untreated group. Patients in the control group had stopped taking somatropin replacement therapy for various reasons and would be a highly select group. Studies by Badia *et al.* and Sanmarti *et al.* found little change in mean QoL-AGHDA scores between baseline and 12 months in an untreated group (9.04 vs. 9.74 and 9.4 and 10.0 respectively), which suggests that patients presenting in clinic have a comparatively stable condition. This may suggest that a regression to the mean does not occur, but it cannot rule out the problem of patient selection. Wiren *et al.* began following up patients 6 months prior to the initiation of treatment. The authors argued that, since they found no significant changes between the first assessment 6 months prior to treatment and the second at baseline, the improvements observed during the administration of somatropin were due to therapy. However, again there is a risk of patient selection bias, and placebo effect.

Due these concerns about bias, it was decided to use in the main ScHARR model a treatment effect size estimated from the Spanish RCT.

3.6 From quality of life to utilities

The review of quality of life has been based on instruments that were not designed for use in economic evaluation. Specifically, they do not generate a preference-based single index that can be used to estimate QALYs. The best method of deriving utility values for the QoL changes would be to directly elicit preferences for these measures, but this is expensive and time-consuming.[54] An alternative approach is to undertake an empirical mapping of the quality of life scales onto an existing preference-based measure. This can be done where the quality of life measures, such as the NHP or QoL-AGHDA, have been used alongside a preference-based measure. Regression can be then used for estimating preference weights from one to the other. Such mapping has been undertaken as part of the Pharmacia & Upjohn submission, and therefore it is possible to translate the gains in QoL-AGHDA reported in the observational studies into a preference-weighted index for use in economic evaluation. The mapping undertaken by Pharmacia & Upjohn is critically reviewed in section 4.1.

3.7 Conclusion

Most RCTs do not provide a good indication of the effectiveness of somatropin because of the poor quality of the studies, the comparatively normal levels of quality of life of patients at baseline in many of the trials which gives little scope for improvement, and the old methods for determining dose levels. The observational evidence, along with the anecdotal evidence and the high adherence rates to somatropin, suggest that this intervention improves the quality of life of some patients. However, the observational data are also poorly reported: they almost certainly over-estimate the extent of any gain, and there is no means of quantifying the size of any bias. It was therefore decided to use in the ScHARR model a treatment effect estimated from an RCT which used the QoL-AGHDA.

4. ECONOMIC EVIDENCE: REVIEW OF SPONSOR SUBMISSIONS

The Southampton Assessment report did not address the economics of somatropin therapy. A few studies of costs are available in the published literature, but the earlier of the two reports by the Wessex Development and Evaluation Committee[55] is the only published work that attempts a cost effectiveness analysis. In the NICE appraisal, three industry models were submitted: Pharmacia & Upjohn (PU),[56] Eli Lilly[57] and Novo Nordisk.[58] Of these, the PU submission was the most sophisticated and caused the greatest problems in interpretation for the NICE appraisal committee. The submissions are reviewed in this chapter, and the analysis which ScHARR has produced is then detailed in chapter 5.

In the economic reports, a number of types of analysis have been attempted. These include cost per % patients 'normalised', and cost per 'normalised' life years gained. Only the cost per QALY estimates have been reviewed.

4.1 Review of Pharmacia & Upjohn submission

4.1.1 Overview

Background

The Pharmacia & Upjohn (PU) model has been developed in three components, with the results from each 'feeding' the next. The first component of the model simulates the events and deaths over 20 years of 1000 patients on treatment or not on treatment. The events are associated with CHD, CVD, MI, stroke, fracture risks in combination with normal mortality risks. The second component of the model uses the results of the simulation to calculate the total costs and quality-adjusted life-years (QALYs) for treatment and non-treatment groups. These are used to find an incremental cost per QALY. The third component of the analysis uses the costs to estimate the budget impact to the NHS.

In this review of the model, the first two components are scrutinised for their validity.

Patient groups

Within this analysis, separate scenarios have been created for different patient groups. At baseline, age and quality of life score were used to determine the grouping. Patients were split into 4 groups according to age, 18-30, 31-54, 55-64 and over 65 years. Quality of life was based on the QoL-AGHDA score of the patient at baseline. The QoL-AGHDA is a score scaling from 0-25 with 0 indicating the best quality of life attainable within the instrument. Patients were split into 4 groups according to QoL-AGHDA score at baseline: 0-5, 6-10, 11-15 and 16+. The inputs and corresponding results create a 4x4 matrix. Thus throughout the PU analysis the results are represented in this way. In the PU analysis it was discovered that patients in the best quality of life group (0-5) showed no increase in terms of their quality

of life, partly since they were at near perfect health at baseline, and consequently had not been included in the analysis.

KIMS database

KIMS is PU's international metabolic survey. It was set up in 1994 and has enrolled more than 5000 patients from over 26 countries. The PU model has used data from the KIMS database to populate the simulation and cost effectiveness model. The advantages with the KIMS database are that patients are followed-up long term, and that it includes a wide variety of patients including those in poor health who may not have qualified for inclusion in the pivotal trials. The primary drawback with this database is the lack of an adequate control arm. A small untreated group of patients has been recruited to KIMS but the number is small and it is not known for what reasons they are untreated. From information available, however, (personal communication: J Monson 15 February 2002) this would appear to be a highly selected patient group, including, amongst others, people who have failed a trial of therapy and patients who have been evaluated by a clinician but not referred for therapy. For the modelling, the patients' characteristics at baseline are assumed to be those of the control group for the duration of the analysis. In the absence of a proper control arm, this is a feasible assumption.

4.1.2 Component 1: Simulation Model

The PU simulation model comprises a patient level simulation. The inputs and outputs to and from the model use Excel spreadsheets whilst the model itself is programmed within Visual Basic Applications (VBA). The code has the following structure:

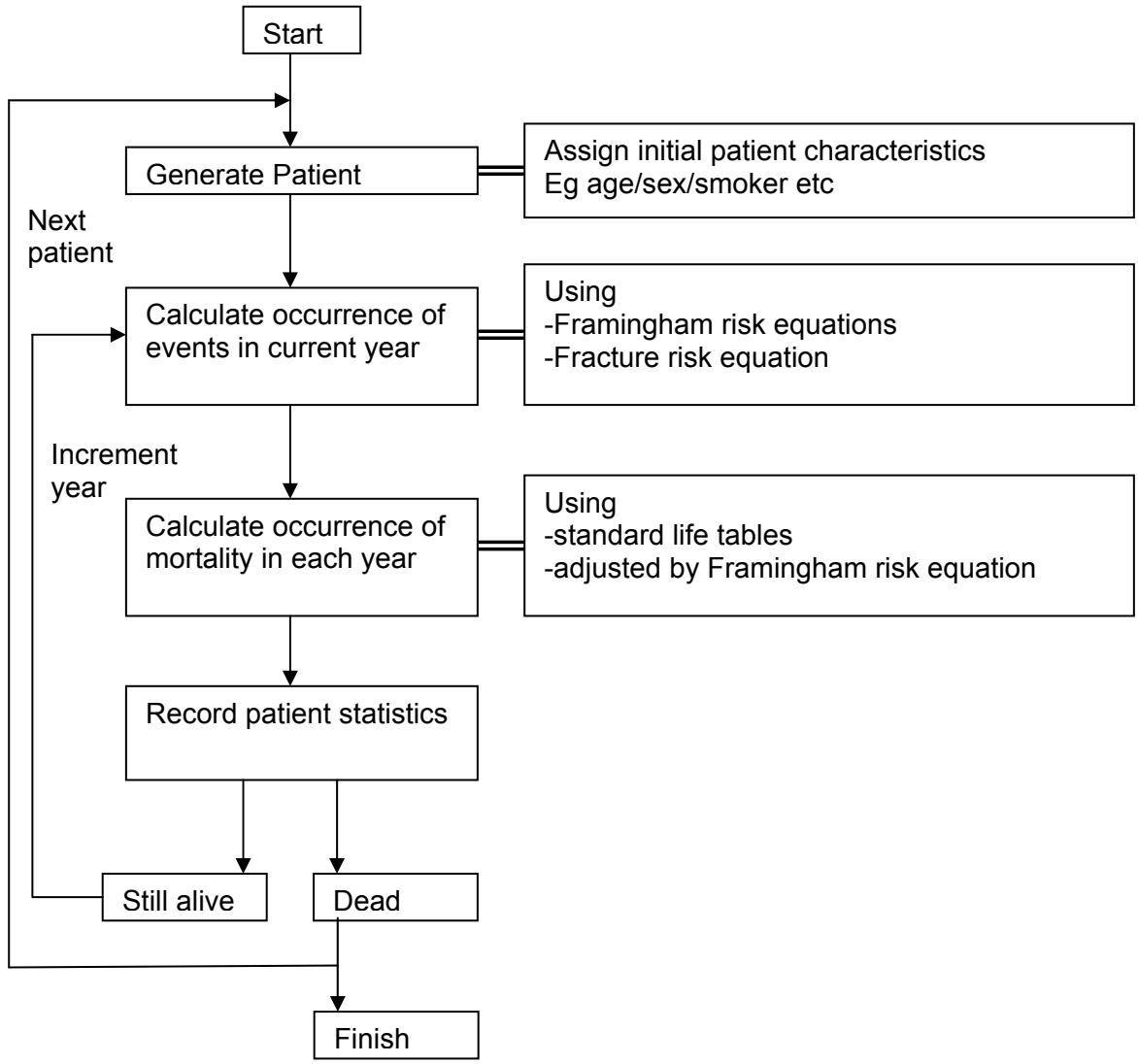


Figure 9: PU simulation model structure

Vascular and Heart Disease predictions

The probability of a CVD, CHD, MI or stroke event or a CVD, CHD death is calculated via the Framingham equations. These equations were developed from the Framingham Heart Study, a 40-year study of 5573 people aged 30 to 74. The Framingham equations are described in Anderson *et al*.[59] Equations were developed for the following outcomes: myocardial infarction (MI), death from coronary heart disease (CHD), a CHD event, a stroke, a cardiovascular disease (CVD) event and a CVD death. There are a number of inputs to the equations: age, sex, whether the patient is diabetic or a smoker, systolic blood pressure and total and high density lipoprotein cholesterol. The equations have not been validated for growth hormone deficient adults but in a number of other disease areas the relationship has been validated, specifically the lipid link has been validated in the treatment of statins.[60] It is a reasonable assumption to include the Framingham equations for human growth hormones, if the evidence for the inputs was validated.

The key parameters for this model, in the sense that human growth hormone is alleged to have a positive impact on them, are systolic blood pressure and total/high density lipoprotein cholesterol. Data for these parameters are derived from the KIMS database.

Table 6: Parameters of mean systolic blood pressure (SBP) from the KIMS database

	No Treatment	Treatment Year 1	Treatment Years 2-5
18-30	117.60	117.40	117.40
31-54	124.70	123.80	123.80
55-64	136.50	136.20	136.20
65+	141.80	141.20	141.20

A range of 85-140 is assumed to be a healthy range with an increasing SBP increasing the risk. Thus it would seem that the patients have a healthy SBP and does not change after somatropin therapy.

Table 7: Parameters of total cholesterol and high density lipoprotein (HDL) cholesterol (KIMS)

	No Treatment	Treatment Year 1	Treatment Years 2-5
18-30	5.20	4.93	4.93
31-54	5.20	4.82	4.82
55-64	5.10	4.76	4.76
65+	5.27	4.69	4.69

The total and HDL cholesterols have not been identified separately. A range of 5.0 to 6.1 of total/HDL cholesterol is identified as a high-risk category. A range of 4.5-4.9 is described a moderate risk. Thus the patients seem to be in a relatively high-risk group. Therefore treatment appears to shift the mean cholesterol levels in patients from the bottom of high risk to the top of the moderate risk categories.

Table 8: Meta-analyses of lipids results over 12 months (taken from table 2.6.4 from Eli Lilly report)

Analysis	No of Studies	HGH n	Mean change (mmol/l) (SD)	Placebo n	Mean change (mmol/l) (SD)	Mean weighted change (mmol/l) (95% CIs)	*p* value
Total cholesterol:	4	58	-0.75 (1.21)	50	0.05 (1.13)	-0.81 (-1.26, -0.35)	0.0005
LDL cholesterol:	4	58	-0.63 (1.37)	50	0.09 (1.20)	-0.72 (-1.21, -0.23)	0.004
Triglycerides:	4	58	-0.19 (1.06)	50	-0.04 (0.91)	-0.16 (-0.54, 0.23)	0.4 (NS)
HDL cholesterol:	4	58	0.06 (0.22)	50	0.01 (0.42)	0.05 (-0.08, 0.18)	0.5 (NS)

NS = Not significant

It is noted that, in the meta-analyses of RCTs, change in HDL cholesterol is not significant. The means and confidence intervals were used to find whether the change in total/HDL cholesterol shown in the KIMS data is a reasonable estimate. KIMS shows a change ranging from 0.27 to 0.58 whilst for reasonable ranges of total cholesterol the meta-analysis shows a range of 0.04 to 0.13.

Therefore the parameters in the PU model will show upper limits of the effects of somatropin on total and HDL cholesterol.

Fracture Risk

A study by De Laet *et al.*[61] has related bone mass density to a risk of a bone fracture. It produces an equation for people aged over 55 relating bone mineral density to risk of hip fracture. This equation is used for all age groups.

Table 9: Bone mineral density (KIMS)

BMD	Treatment Year 1	Treatment Years 2-5	No Treatment
<65	0.89	0.89	0.88
65+	0.82	0.82	0.80

There has been recent criticism of the link between BMD and hip fractures but the assumptions used in the model are still reasonable.[62]

The changes are small and in the over 65 group – the most prevalent group – the risk of a fracture changes from at a maximum 0.12% to 0.10%. To quantify this, for 1000 patients over 20 years there will be at a maximum 4.9 more hip fractures in the untreated group. The KIMS data are thought to be more useful than trials since patients of poorer quality of life were included. The company submission from Eli Lilly worked a meta-analysis of the trials. None of the meta-analyses were shown to be significant (see table 10). The mean change in total BMD was –0.009, less than the minimum from the KIMS study. The minimum change from the meta-analysis was even positive and the maximum was –0.044.

Table 10: Meta-analysis of BMD

Outcome	Duration	Weighted mean average	p value
Total body BMD	6 months	-0.009 g/cm2	NS
Lumbar spine BMD	6 months	-0.005 g/cm2	NS
Femoral neck BMD	6 months	+ 0.005 g/cm2	NS
Total BMD	12 months	-0.006 g/cm2	NS
Femoral neck BMD	12 months	0.00 g/cm2	NS

Withdrawals

The rates of withdrawal have been obtained from the KIMS observational database. Only 3 years of data have been collected for all but the 0-5 quality of life group where 4 years data have been collected. This is not grouped according to age. Withdrawals after 3 (or 4 years) are assumed to be zero. The withdrawals parameter is a weakness in the simulation model, but the results of the simulation are not directly related to the parameter and results can be adjusted without rerunning the simulation.

Summary critique of the simulation model

The simulation model has a sensible and valid mode of action. The code has not been written to optimise the speed of the simulation and thus can take many hours to produce a result, this makes validation difficult. The detailed calculations of the simulation have not been fully validated, though the implementation of the key long term modelling, including the Framingham equations, has been checked. Ideally the model should be replicated to test its robustness but it is apparent that the clinical parameters have so little benefit to the cost effectiveness so this has not been done. The results of the clinical benefits are included in the ScHARR analysis, with the option of taking them out.

4.1.3 Component 2: Cost-Effectiveness model

Costs of treatment

The simulation model has been run once for each age/quality of life group and for treatment and non-treatment groups. The results have then been 'pasted' into a new Excel workbook where individual sheets (n=16) for each of the age and quality of life groups derive the cost effectiveness.

A comprehensive list of the unit costs multiply the numbers of events occurred to find the costs associated with CHD treatment separately, and MI, stroke and hip fracture treatment combined.

Correction 1 – A correction was necessary to the PU model where formulas used for costing the CHD events were incorrect. The formula divided percentages by 100 [row 55 and 64 of all 16 sheets] in a sense multiplying costs by 1/10000. It is worth noting that the correction increases the costs avoided 100 fold, however the impact on the overall cost effectiveness is still small.

Correction 2 – In each cost effectiveness sheet, the cost of treatment and physician visits was based on the number of patients on treatment at the end of the period[rows 53,54 and 62,64]. These should be calculated as the number of patients at the start of the period, and allow for withdrawals by making a half year correction.

The number of physician visits made by patients in the treatment and non-treatment groups is also estimated. These numbers are different in each

age group. Physician visits are not broken down by GP or specialist visits and therefore one is not able to distinguish whether the treatment group is seeing the GP more than a specialist. The number of visits for the treatment group seems to be low, in some cases less than 1. It is thought at a minimum, a patient would have an annual check with an endocrinologist. In an elderly population it is likely that a patient would see a GP four times a year.

The drug dose is also estimated in milligrams, and varies from the first year to successive years of treatment. The dose is different for each age group. Since the same quality of life utilities are used for the under-65, the difference in dose in the key driver to the different cost effectiveness results between the three under-65 age groups.

There is no initial cost for diagnosis or defining the required dose. Also PU offer the first 3 months of treatment at no cost. This discount should not be made in the central analysis.

Mapping

A key component of the Pharmacia & Upjohn submission has been a mapping of the NHP onto a preference-weighted SF-36 index and between a preference-weighted NHP onto the QoL-AGHDA. The models developed for these mappings have been provided to the review team.

The models fitting the NHP onto SF-36 preference-based index would seem to be credible in terms of the regression coefficients and a comparatively good fit (around 55% of variation explained by the model). The model linking NHP to QoL-AGHDA has been produced on the KIMS dataset and also achieved credible coefficients and a reasonable fit (65% of variation explained by the model). Such models are designed to predict SF-6D preference-based index values from NHP and QoL-AGHDA data.

The mapping models are limited by the data from which they are estimated and the degree of overlap of the descriptive systems. The mapping function has been estimated from two different datasets, one from a general population and the other GH-deficient patients. Models estimated from one dataset should not be applied to a substantially different dataset, such as in terms of the severity of health problems. For GH deficiency and the general population sample used in these models, it seems likely that there will be a reasonable degree of overlap. Differences in the descriptive system are likely to be a more important cause of bias since one or other of the descriptive systems may miss the key elements of health. Although these instruments seem reasonably well correlated, there are likely to be gaps in both. The consequence of these differences in the descriptive systems is not known. Despite these concerns, the models presented by Pharmacia and Upjohn compare well with similar mapping exercise recently undertaken in ScHARR[63].

Utility scenarios

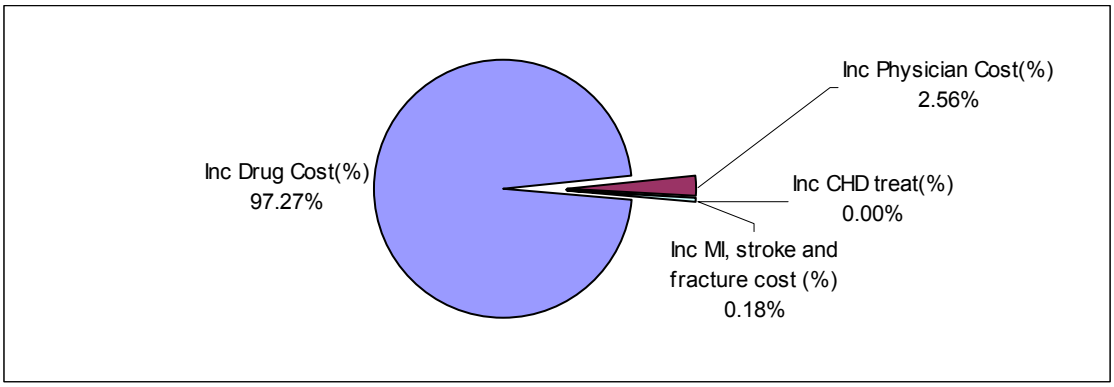
A key assumption in the model concerns the utility profiles of the treatment and non-treatment groups in the model. PU have made two sets of assumptions. The first is called a 'conservative' scenario where the utility improvement observed in the treated patients in the KIMS dataset is assumed to remain for twenty years, along with an unchanged utility score in the untreated group. The KIMS data support this assumption for four years; beyond that time there are no data. The submission refers to the study by Gibney *et al.*[35] which shows an improvement lasting 10 years, but as noted above this study would have been subject to selection bias. There are no control group data to support the assumed pattern for the untreated group. The second set of assumptions is called the 'intuitive' scenario and assumes deterioration in the untreated group for the first five years of 0.01 in the utility score followed by stability. The treated group is assumed to remain unchanged for 10 years, and then 5 years of deterioration at 0.01 followed by stabilisation. Given benefits are discounted this results in a larger overall gain from treatment. The authors of the submission cite Sanmarti and colleagues[64] who found a reduction in QoL-AGHDA from 9.6 (95% CI: 8.4-10.4) to 10.0 (95% CI: 8.8-11.0) over one year. The authors do not present the statistical significance of this change nor do they demonstrate that such deterioration will continue for 5 years. There would seem to be little support for the 'intuitive' scenario.

4.1.4 Results

Impact of clinical events on the economics of therapy

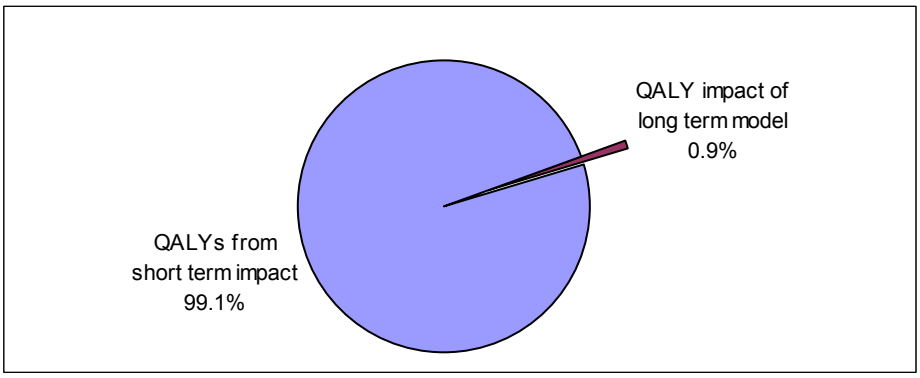
The impact on lipids and bone mineral density was analysed to find its importance in the analysis. Figure 10 indicates that, in terms of the incremental costs, the savings associated with reductions in long-term clinical events, for example fractures and CHD events, are overwhelmingly dominated by the drug cost. The total physician costs are in the order of 10 times greater to the costs of all the clinical events.

Figure 10: Breakdown of the incremental costs (modulus) for the most severe and eldest patient group in the PU analysis



Similarly an analysis on the incremental QALYs is presented in Figure 11. It can be seen that the long-term model results in less than 1% of the total QALY gain from therapy.

Figure 11: Breakdown of the incremental QALYs for the most severe and eldest patient group in the PU analysis



The breakdown of incremental costs and QALYs presented in Figures 10 and 11 are for the most elderly and severe (QoL-AGHDA 16+) patient group. That is the group with most favourable economic characteristics; in the other age/severity groups the relative long-term impact on costs and QALYs is smaller again.

The impact on the cost effectiveness ratios can be seen in tables 12 and 13 in the next section. The impact on the ICER is in the region of £1000.

Therefore the impact of the long term modelling on the economics of therapy is shown to be exceedingly small in the PU model. This issue is addressed again in the ScHARR model to ensure that this still holds true for the independent assessment. The implication of this is that it is inappropriate to use a long time horizon in the analysis.

Economics of therapy in the PU age/severity classes

The 'conservative' results of the PU analysis are shown below. In the PU report, an 'average' figure is presented which simply takes the mean of the cost effectiveness results for each of the age/AGHDA subgroups. The calculation of this figure is seriously flawed and it should be ignored completely. It is later seen in the ScHARR analysis that an overall figure should be derived but this should be calculated using the incidence breakdown for each age/AGHDA groups.

Table 11 PU Conservative Results

QoI-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	£85,425	£76,434	£56,036	£24,456
11-15	£38,005	£34,593	£26,848	£18,327
16+	£28,425	£25,753	£20,256	£16,927

When the corrections 1 and 2 are corrected the estimates are slightly increased.

Table 12 Corrected PU Conservative Results

QoI-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	£86,905	£78,116	£58,483	£26,258
11-15	£38,772	£35,442	£28,054	£19,746
16+	£28,837	£26,231	£21,002	£18,090

When the clinical effects (lipids, fracture risk) are taken out it changes to

Table 13 Corrected PU results without clinical effects

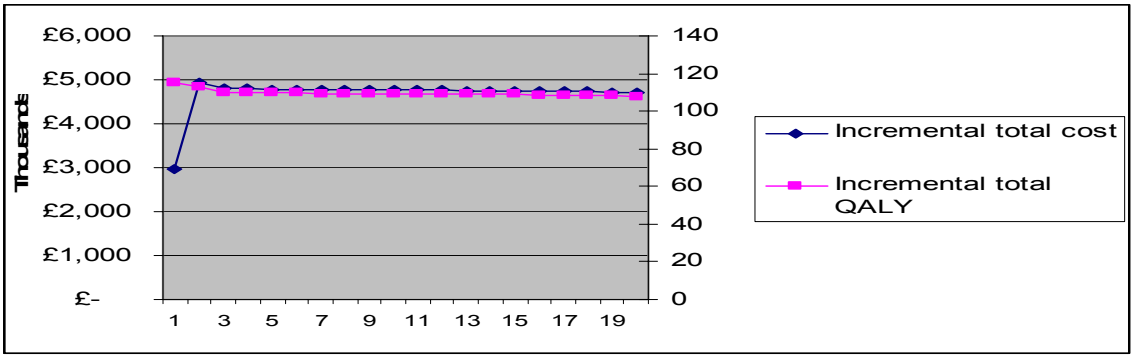
QoI-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	£87,096	£79,626	£65,791	£29,401
11-15	£38,793	£35,472	£29,316	£21,426
16+	£28,869	£26,383	£21,765	£19,430

When discounting is removed (0%, 0%) the results increase considerably.

Table 14 Corrected PU Conservative Results without discounting

QoI-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	£124,581	£111,081	£80,204	£34,665
11-15	£55,481	£50,357	£38,599	£26,070
16+	£41,390	£37,385	£29,043	£23,989

Figure 12 shows that the costs and effects remain constant after the first year of treatment. This indicates that the investment benefits from giving somatropin are small and the 20 year analysis is in some ways misleading since the decrease in cost effectiveness is due primarily to the differential in



discounting rates.

Figure 12 Incremental costs and QALYs over time

The 5 year results are shown below

Table 15 Corrected PU Conservative results with discounting at 5 years

QoI-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	£114,704	£104,598	£84,686	£34,374
11-15	£51,404	£46,901	£38,244	£25,334
16+	£37,798	£34,484	£27,889	£22,876

Table 16 Corrected PU Conservative results with discounting at 4 years without first 3 months treatment free

QoI-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	£122,584	£112,148	£91,532	£36,951
11-15	£54,987	£50,340	£41,386	£27,263
16+	£40,335	£36,921	£30,100	£24,559

The ScHARR analysis further analyses different scenarios using the PU data.

4.2 Review of Eli Lilly economic analysis

4.2.1 Overview

The Eli Lilly economic analysis[57] used simple calculations of Cost and QALY. No long term 'modelling' is undertaken to find an outcome. The time horizon for the analysis is 12 months and therefore discounting is unnecessary.

4.2.2 Costs

Included in the analysis are costs of the drug, diagnosis and monitoring, hospitalisations, visits to health professionals and lost working time. The major difference between this analysis and the PU work is that costs of diagnosis and monitoring are included at the start of treatment. This has also been included for the untreated group. Also, visits to health professionals have been calculated from data from the Lilly observational database. This has been done separately for GPs and specialist endocrinologists. The total annual cost for patients on treatment is £6,244 when lost work time is not included and £1,350 for the non-treatment group. The incremental cost of somatropin treatment is therefore £4,894 per year, but falls to £4,113 when lost work time is included.

4.2.3 QALYs

The change in QALY has been estimated from calculations made in the earlier of the two Wessex DEC reports.[55] However, the methodology used in this report is at best crude, and was subsequently abandoned by the Wessex DEC when they revisited the topic two years later. The baseline utility is

estimated by choosing what the answers to questions on the index of Health Related Quality of Life [65] could be for patients with GH deficiency. The impact of treatment on the different domains is then estimated. The change is then subtracted to find an incremental QALY. The analysis uses 0.16 for 12 months.

4.2.4 Results

The incremental costs and QALYs are then used to calculated the cost effectiveness. When lost work time is not included the incremental cost per QALY is £30,586. When lost work time is included this falls to £25,705.

4.2.5 Critique

The analysis by Eli Lilly is simple but too general for the defined patient group. This is not a reason for the results to be discarded. The 12 month time horizon for the analysis is justified. The costs have been well calculated and offer good evidence for further economic analyses. The let down to the analysis is the quality of life assumptions. The results are in fact only relevant for patients who have a baseline quality of life less than 0.84, else patients could move to a quality of life greater than 1, which excludes a large proportion of the patient group. Starting quality of life is indeed very important.

It is unclear whether a cost of diagnosis and monitoring should be included in the untreated group.

The results are in the order shown by PU for the older patient group.

4.3 Novo Nordisk Review

4.3.1 Overview

The economic analyses presented by Novo Nordisk[58] has not been replicated into the ScHARR modelling. The utility methods used in the Novo analysis have poor validity and the resulting cost effectiveness ratios are consequently flawed. However, the work should not be completely ignored since some interesting and important assumptions are made which are relevant to the final analysis.

4.3.2 Quality of life

The impact on quality of life is estimated using the method described in the Wessex DEC report,[55] previously used in the Eli Lilly submission.[57] In contrast to the Lilly submission, where the actual utilities were taken directly from the DEC report, Novo have redefined the states that patients with growth hormone deficiency would start at. The DEC report considered patients with severe, moderate and mild symptoms. The Novo work considers patients described as either having severe symptoms or hypopituitary patients. The patients with severe symptoms, who respond to treatment, are assumed to improve their quality of life utility by 0.376, based on assumptions that patients

will move from slight physical disability to slight social disability, and from extremely depressed to no distress. This figure is of the order of 3 to 6 times more than predicted in the PU mapping work. Once again, this crude method of predicting QoL gain must be questioned and the figure it produces seems over estimated. The gain for the hypopituitary patients is estimated to be 0.126, a more reasonable estimate and similar to the Lilly analysis.

4.3.3 Clinical events

A crude estimate has been made to estimate the impact of somatropin replacement on lipid profile and BMD. As table 17 shows, and in accordance with the PU analysis, the estimates have no significant impact on the ICER. The difference is in a similar region to that estimated by PU. An important conclusion is 'the link between somatropin replacement and decrease in fracture risk remains to be proven' (pg 32 Novo report).

4.3.4 Other issues

Another interesting assumption is that only 60% of patients are assumed to respond to treatment in the 6-month trial period. This is less than has previously been estimated.

The time horizon is 10, 20 and 40 years. As has been previously discussed this is unreasonable and should be much shorter.

4.3.5 Results

The severe-symptoms patients analysis is not considered. The hypopituitary patients' 10 year results are 16,550 not discounted and 13,068 discounted. Per year this calculates to an incremental cost of £1,655 (not discounted) and a QALY of 0.062 (not discounted).

Table 17: Impact of clinical effects (lipids/fractures) on ICERs (discounted hypopituitary 20 years)

Time horizon	No clinical effects included	Clinical effects included
20 years	£18,288	£17,424

4.3.6 Conclusion

The conclusions of the analysis are similar to the Eli Lilly submission.

4.4 Conclusion

The most important conclusion from the analysis of the industry submissions is that the impact of therapy on short-term quality of life is the most important parameter in the cost effectiveness model. By contrast the benefits on lipid lowering and BMD impact, have little to no economic impact. The evidence for any positive impact of somatropin on fracture risk is not conclusive. The observational evidence on total and HDL cholesterol shows some significant,

but very small, beneficial changes in these surrogate end points. However, the cost savings due to the potential for fewer adverse events are completely outweighed by the drug costs. When it is seen that it is the improvement in QoL and not the costs that are the most important and sensitive variable, the effects on fracture-risk and cardiovascular-risk can reasonably be disregarded.

The QALY estimates made by Eli Lilly and Novo Nordisk have low validity. Neither estimate is based upon patient-administered questionnaires, and both have used very crude assumptions. The gains should in no way be interpreted as central estimates.

The costs used are useful and well presented. However, due to the weaknesses in the QALY estimates, the estimated ICERs presented in the industry submissions should be viewed as extremely conservative.

5. ECONOMIC EVIDENCE: ScHARR ASSESSMENT

It was deemed necessary to develop a new analysis to determine the cost effectiveness of human growth hormones in adults.

The ScHARR model has been created in excel with a small amount of VBA. The most useful evidence is that presented by PU. As described earlier, the KIMS database provides some useful data for the treatment group not available from the clinical trials.

As identified in the previous section, the long-term impact of therapy on mortality and costs has very little impact on the economics of treatment. Whilst the validity of the long term modelling may remain in question, due to its small impact this is not a key economic issue. The PU long-term results have, therefore, been included within the ScHARR analysis. Costs and resource usage from Eli Lilly, Novo Nordisk and the Southampton assessment have been used in the ScHARR analysis. The model has been made to be flexible so the results of the PU and Lilly analyses can be replicated. The model can then easily be recalculated to allow a number of the important variables that make up the QALYs and costs to be changed. Thus the impacts of each variable can easily be explored and the differences between the company and ScHARR estimates can transparently be seen. A number of different pieces of evidence have been used to populate the parameters which develop the costs and the QALYs.

5.1 Parameters

Utilities

- The PU 'conservative' scenario
- The PU 'intuitive' scenario
- The Eli Lilly utility gain taken from the Wessex DEC report
- The ScHARR 'optimistic' scenario = to the PU 'conservative scenario
- The SCHARR 'McKenna 1' scenario where the result of the Spanish RCT gain of x is used.[10] The full results of McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

Discounting

- NICE guidance of 6% cost, 1.5% effect
- 3%,1.5%
- 1.5%,1.5%
- 0%,0%

Time Horizon

- 1 year
- 4 year
- 10 year
- 20 year

Calculation Point

- The end of year calculation as PU
- The mid year adjustment, so costs and QALYs are estimated at the midpoint of each year.

Withdrawals

- PU assumption. Withdrawal rates from KIMS for first 3 years and then no withdrawals there after.
- Eli Lilly assumption. No withdrawals over the year analysis
- ScHARR progression assumptions. Withdrawals are assumed to carry on past 3 years. The first 3-year withdrawal rates are calculated and then assumed constant past 3 years.
- Wessex assumption. 20% patients fail to respond after 6 months and then follow assumption below.[66] ScHARR assumption above past 1 year.
- Novo assumption – 40% withdrawal after 6 months, then ScHARR assumption past 1 year.

Clinical effects

The effects on total and HDL cholesterol and bone mineral density can be included or excluded. Excluding these effects removes the difference in life years lived between treatment and non treatment groups and excludes all costs associated with cardiovascular adverse events and fractures.

Drug Costs

The individual vial prices for each of the 5 drugs is included in the analysis. Genotropin (Pharm&Upjohn), Humatrope (Lilly), Norditropin (NovoNordisk), Saizen (Serono), Zomacton (Ferring).[55]

Drug Dose

Dosing levels from KIMS are assumed to be the most accurate available. The data from the Eli Lilly submission were used to calculate the average dose of patients in Lilly's observational database. This was much higher than KIMS.

Initial Costs

There were no initial diagnosis and dose scheduling costs in the PU model. These costs were taken from the Wessex report.[55] Costs from Eli Lilly analysis are also included.

Ongoing costs

An ongoing monitoring cost has been developed from the Wessex report.[55]

Physician costs

The physician costs from the PU and Lilly analyses have been incorporated into the model. Costs have been separated for GP and specialist visits.

Cost of clinical events

Some of the cost vectors from the (personal communication: Dr Matt Stevenson 15 February 2002) HTA Osteoporosis report have been included into the model.

5.2 Results

The results of the ScHARR analysis are presented below. With the uncertainty of the quality of life benefits, the purpose of the model is to identify a likely range for the incremental cost effectiveness and explore its sensitivity to the different parameters. A central estimate is not presented. Instead, a central estimate for the costs has been selected from the list of possible parameters, and then the most optimistic QoL gain is presented. Then the QoL gain is changed to show the importance of the parameter.

The model has been developed so that different inputs and scenarios can be easily interchanged. Thus one can move, by changing the inputs, from the PU estimates, to the Lilly estimates, to ScHARR's (the Novo analysis has not been included). The model is easily adaptable and entirely transparent, so if new or better parameters are found they can be included.

The model presents ICERs separately for each age and QoL-AGHDA group in the PU analysis. However, we have a number of concerns about this method of presentation. Firstly, the focus on age is rather misleading since the large differences in ICER between age groups results from differences in dosage. A younger patient seems to need a greater amount of dose titration to achieve a response and thus incurs a greater treatment cost. Secondly, we are not sure whether QoL-AGHDA scores could be used to control access to this intervention at the patient level. We are aware of the use of QoL-AGHDA as part of the overall assessment and clinical evaluation, but using subjective assessments to determine who receives treatment raises major practical problems.

Thus the correct way to present the cost effectiveness is an overall figure for the whole adult growth hormone deficient population. The overall cost effectiveness is calculated using the population breakdown of the age/QoL-AGHDA groups to reconstruct a total population average incremental cost and QALY. The data from KIMS have been used to estimate this population breakdown (summarised in Table 18).

Only results were presented for the over-65 and under-65 population by QoL-AGHDA group. Thus the under-65 population has been broken down evenly by the number of years in each group.

Table 18 An estimate of the distribution of patients on treatment

Qol-AGHDA group	Age Group			
	18-30	31-55	56-64	65+
0-5	*	*	*	*
6-10	8%	16%	6%	3%
11-15	8%	16%	6%	3%
16+	9%	17%	6%	2%

5.2.1 Scenario 1

A 4 year analysis has been chosen in the first scenario. The reason for this is as follows:

- Long-term benefits associated with somatropin therapy are economically insignificant
  - either clinical effects (total and HDL cholesterol or BMD)
  - The impact of somatropin has an onset within a few months and effect is sustained only with continued treatment
  - i.e. there is no long term investment benefit in using somatropin therapy
- The misleading benefit from differential discounting is reduced
- KIMS provides data only for 4 years.

Dosing is different in the first year to subsequent.

The utilities are varied in the three sub scenarios. A mixture of costs from the PU, Eli Lilly and Wessex Assessment cost model have been used. Each parameter can be easily identified for its value and reference in the model. Major differences from the PU estimates are:

- initial costs from the Wessex report is included for the treatment arm whereas no initial cost is assumed for the non-treatment arm
- physician costs which are broken down by specialist and GP visits are taken from the Eli Lilly submission
- no annual treatment costs have been included.

An overall ICER can be derived. For scenario 1 they are

Scenario	Utility Description	ICER
1a	PU conservative estimate/ScHARR optimistic	£51,457
1b	McKenna 1 trial	*
1c	No quality of life benefit	£21,090,631

* The full conclusions on the health economics are not reported here because they were based upon the results of the McKenna (1) trial which was submitted to NICE in confidence by Pharmacia & Upjohn.

A comprehensive sensitivity analysis has been performed and is presented in full in Appendix 5. A summary of the sensitivity analysis results is included in Table 19. It should be noted that the low and high cost effectiveness figures for the age/QoL-AGHDA groups referred to do not constitute an uncertainty interval for the population cost effectiveness.

Table 19 Summary results of sensitivity analysis

	Overall	Age 65+ AGHDA 16+	Age 18-30 AGHDA 6-10
[1a] ScHARR 'Optimistic'	£51,457	£25,286	£124,941
[1b] ScHARR 'McKenna 1 trial'	*	*	*
[1c] ScHARR 'No QAL benefit'	£20,090,631	£1,788,694	£254,012,046
Sensitivity Analyses			

[2a] 1 Year	£62,011	£32,251	£141,303
[2b] 4 Year no discounting	£53,968	£26,448	£131,272
[2c] 10 Years	£44,207	£21,284	£110,385
[2d] 20 Years	£37,496	£18,481	£96,087
[2e] 20 Years no discounting	£49,862	£23,411	£127,629
[3] No clinical impact	£51,617	£25,538	£125,006
[4a] High Costs	£55,172	£27,600	£133,358
[4b] Low Costs	£45,337	£21,828	£110,554
[5] 40% Initially fail (Novo)	£57,714	£28,528	£139,500
[6] Productivity included	£42,003	£18,899	£104,294
[7c] PU correct	£36,084	£18,090	£86,905

The most optimistic assumption for the utility gain provides an incremental cost per QALY in the region of £50,000. The actual utility gain is disputed and the cost effectiveness estimate is very sensitive this figure. It is clear that £50,000 per QALY is at the floor of the cost effectiveness.

*The full conclusions on the health economics are not reported here because they were based upon the results of the McKenna (1) trial which was submitted to NICE in confidence by Pharmacia & Upjohn.

5.3 Conclusion

The full conclusions on the health economics are not reported here because they were based upon the results of the McKenna (1) trial which was submitted to NICE in confidence by Pharmacia & Upjohn.

In contrast the economic submissions by the industry estimated the short-term quality of life impact of somatropin therapy as:

- Eli Lilly 0.16
- Novo Nordisk 0.126 - 0.376
- PU 0.02 – 0.12

The validity of the assumptions underlying the industry submissions has been discussed in this document but all show significant weaknesses and are consistently overestimates.

The economic submissions by the industry estimated the ICER for somatropin as:

- Eli Lilly £25,700 - £30,600
- Novo Nordisk £13,600 - £22,400
- PU £27,500 – £37,600

The most optimistic assumption for the utility gain within the ScHARR analysis gives an incremental cost per QALY in the region of £52,000. The actual utility gain is disputed but it is clear that the cost effectiveness is very sensitive to this figure. The estimate of £52,000 per QALY is at the floor of the estimates.

The full conclusions on the health economics are not reported here because they were based upon the results of the McKenna (1) trial which was submitted to NICE in confidence by Pharmacia & Upjohn.

The benefits of somatropin, in terms of improved total and HDL cholesterol levels and BMD, have very little long-term economic impact. The effect on the ICER of removing the long-term clinical effects is to change the ICER from £51,457 to £51,617.

The time horizon of the analysis has some effect on the ICER. This is primarily due not to the inclusion of long-term clinical benefits but to differential discounting of the short-term clinical benefits and costs. A 20 year analysis gives an ICER of £38,000 per QALY.

All other uncertain parameters within the model have only marginal impacts upon the cost effectiveness. The treatment costs impact the ICER by +/- £8000 per QALY. Including lost productivity decreases the ICER by around £10,000 per QALY.

6. DISCUSSION

Contribution of this review

This review has sought to make more of the existing evidence in order to address the concerns of the NICE Appraisal Committee. The Southampton review was severely limited by the poor nature of the RCTs conducted in this area. By restricting the review to RCTs, they concluded that somatropin has little or no benefit for QoL and is unlikely to be cost-effective.

Extending the scope of the review to examine the longer-term benefits of somatropin does not alter this conclusion since these benefits are comparatively small. The additional trial data using the NHP have not altered the conclusions of the meta-analysis that the overall treatment effect is small and contradictory.

The full results of the McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

The review of observational studies has provided some additional evidence, but it must be interpreted with caution. The overall conclusion from these additional reviews has been to be rather more positive about the likelihood of a beneficial impact of somatropin for many individuals. However, it is difficult to be able to offer a precise estimate of the size of the benefit. The evidence overall probably suggests that, at best, the gain is small to moderate for those who stick with therapy, and the economic implications of such a gain have been discussed in the preceding chapters.

Using QoL-AGHDA to allocate treatment

An important claim in the Pharmacia & Upjohn submission is that the cost per QALY of somatropin therapy is related to baseline QoL-AGHDA score and age. There would seem to be a certain amount of support from the observational literature that, the higher the initial QoL-AGDHA score, the larger the gain.

The full results of McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

The implication of the QoL-AGDHA breakdown of the cost per QALY results would seem to be that limiting treatment to those with higher QoL-AGDHA scores would improve the economics of treatment.

This raises the practical question of whether it would be possible to use an individual's baseline QoL-AGDHA score to determine whether or not to offer somatropin therapy. The questionnaire is easy to use and, given the simple summation scoring method, it would be straightforward to tally the score and decide whether or not the patient met the criteria for treatment. Currently, the QoL-AGDHA is routinely used on many GH-deficient patients, and it can be used in assessment prior to treatment. However, it is not currently used to

preclude treatment. The problem with using this method to ration care is that it is subjective and open to manipulation by the patient to determine whether or not they receive treatment.

Weaknesses

The primary weakness continues to be the poor quality of evidence for the impact of somatropin on QoL. However, given that the observational studies probably provide an upper estimate of the likely benefit, better evidence is unlikely to improve the cost per QALY position of this intervention over and above that suggested by the observational studies.

Strengths

The estimates for the other parameters in the model assessment, such as long-term effects and costs of the intervention have little impact on the result. In this respect the conclusions drawn from the economic analyses are robust.

Reference List

1. Bryant, J., Loveman, E., Chase, D., Mihaylova, B., Cave, C., Gerard, K., and Milne, R. Clinical and cost-effectiveness of growth hormone in adults: quality of life. 2001.
   Ref Type: Report

2. Nicolson A, Toogood AA, Rahim A, Shalet SM. The prevalence of severe growth hormone deficiency in adults who received growth hormone replacement in childhood. *Clinical Endocrinology* 1996;**44**:311-6.

3. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavanagh DJ *et al*. Assessing the quality of reporting of randomised clinical trials: is blinding necessary? *Controlled Clinical Trials* 1996;**17**:1-12.

4. Bengtsson BA, Eden S, Lonn L, Kvist H, Stokland A, Lindstedt G *et al*. Treatment of adults with growth hormone (GH) deficiency with recombinant human GH. *Journal of Clinical Endocrinology & Metabolism* 1993;**72**:309-17.

5. Bronstein MD, Musolino NRC, Cunha-Neto MB, Soares CN, Caires MAF, Rosenthal MC *et al*. Impact of treatment with growth hormone on the psychological and somatic profile of adults with growth hormone deficiency. *Endocrinology and Metabolism* 1997;**4**:167.

6. Dean HJ, McTaggart TL, Fish DG, Friesen HG. The educational, vocational, and marital status of growth hormone-deficient adults treated with growth hormone during childhood. *AJDC* 1985;**139**:1105-10.

7. Degerblad M, Almqvist O, Grunditz R, Hall K, Kaijser L, Knutsson E *et al*. Physical and psychological capabilities during substitution therapy with recombinant growth hormone in adults with growth hormone deficiency. *Acta Endocrinologica* 1990;**123**:185-93.

8. Mardh G, Lundin K, Borg G, Jonsson B, Lindeberg A. Growth hormone replacement therapy in adult hypopituitary patients with growth hormone deficiency: combined data from 12 European placebo-controlled clinical trials. *Endocrinology and Metabolism* 1994;**1**:43-9.

9. McKenna SP, Koppeschaar HPF, Zelissen PMJ, Snel YEM, Whalley D, Doward LC. Impact of growth hormone replacement on quality of life: results of a clinical trial in the Netherlands. *Endocrinology and Metabolism* 1997;**4**:167.

10. McKenna SP, Vazquez JA, Pico A, Escobar F, Torres E, Ulied A *et al*. Quality of life in adult patients with growth hormone deficiency in Spain and the effects of replacement therapy with growth hormone. *Endocrinology and Metabolism* 1997; **4**:168.

11. Salomon F, Cuneo RC, Hesp R, Sonksen PH. The effects of treatment with recombinant human growth hormone on body composition and

metabolism in adults with growth hormone deficiency. *New England Journal of Medicine* 1989;**321**:1797-803.

12. Whitehead HM, Boreham C, McIlrath EM, Sheridan B, Kennedy L, Atkinson AB *et al*. Growth hormone treatment of adults with growth hormone deficiency: results of a 13-month placebo controlled cross-over study. *Clinical Endocrinology* 1992;**36**:45-52.

13. Whitehead HM, Aiken B, Lewis S, Sheridan B, Hadden DR. Physiological growth hormone secretion in adult growth hormone deficiency: comparison with normal controls. *Clinical Endocrinology* 1991;**34**:371-6.

14. Attanasio AF, Lamberts SW, Matranga AM, Birkett MA, Bates PC, Valk NK *et al*. Adult growth hormone (GH)-deficient patients demonstrate heterogeneity between childhood onset and adult onset before and during human GH treatment. Adult Growth Hormone Deficiency Study Group. *Journal of Clinical Endocrinology & Metabolism* 1997;**82**:82-8.

15. Baum HBA, Katznelson L, Sherman JC, Biller BM, Hayden DL, Schoenfeld DA *et al*. Effects of physiological growth hormone (GH) therapy on cognition and quality of life in patients with adult-onset GH deficiency. *Journal of Clinical Endocrinology & Metabolism* 1998;**83**:3184-9.

16. Beshyah SA, Freemantle C, Shahi M, Anyaoku V, Merson S, Lynch S *et al*. Replacement treatment with biosynthetic human growth hormone in growth hormone-deficient hypopituitary adults. *Clinical Endocrinology* 1995;**42**:73-84.

17. Cuneo RC, Judd S, Wallace JD, Perry-Keene D, Burger H, Lim-Tio S *et al*. The Australian Multicenter Trial of Growth Hormone (GH) Treatment in GH-Deficient Adults. *Journal of Clinical Endocrinology & Metabolism* 1998;**83**:107-16.

18. Deijen JB, de Boer H, van der Veen EA. Cognitive changes during growth hormone replacement in adult men. *Psychoneuroendocrinology* 1998;**23**:45-55.

19. Giusti M, Meineri I, Malagamba D, Cuttica CM, Fattacciu G, Menichini U *et al*. Impact of recombinant human growth hormone treatment on psychological profiles in hypopituitary patients with adult-onset growth hormone deficiency. *European Journal of Clinical Investigation* 1998;**28**:13-9.

20. McGauley GA. Quality of life assessment before and after growth hormone treatment in adults with growth hormone deficiency. *Acta Paediatrica Scandinavica - Supplement* 1989;**356**:70-2.

21. Soares CN, Musolino NR, Cunha NM, Caires MA, Rosenthal MC, Camargo CP *et al*. Impact of recombinant human growth hormone (RH-GH) treatment on psychiatric, neuropsychological and clinical profiles of

GH deficient adults. A placebo-controlled trial. *Arquivos de Neuro-Psiquiatria* 1999;**57**:182-9.

22. Verhelst J, Abs R, Vandeweghe M, Mockel J, Legros JJ, Copinschi G *et al*. Two years of replacement therapy in adults with growth hormone deficiency. *Clinical Endocrinology* 1997;**47**:485-94.

23. Wallymahmed ME, Foy P, Shaw D, Hutcheon R, Edwards RH, MacFarlane IA. Quality of life, body composition and muscle strength in adult growth hormone deficiency: the influence of growth hormone replacement therapy for up to 3 years. *Clinical Endocrinology* 1997;**47**:439-46.

24. Burman P, Broman JE, Hetta J, Wiklund I, Erfurth EM, Hagg E *et al*. Quality of life in adults with growth hormone (GH) deficiency: response to treatment with recombinant human GH in a placebo-controlled 21-month trial. *Journal of Clinical Endocrinology & Metabolism* 1995;**80**:3585-90.

25. Florkowski CM, Stevens I, Joyce P, Espiner EA, Donald RA. Growth hormone replacement does not improve psychological well-being in adult hypopituitarism: a randomized crossover trial. *Psychoneuroendocrinology* 1998;**23**:57-63.

26. Ahmad AM, Hopkins MT, Thomas J, Ibrahim H, Fraser WD, Vora JP. Body composition and quality of life in adults with growth hormone deficiency; effects of low-dose growth hormone replacement. *Clinical Endocrinology* 2001;**54**:709-17.

27. Bulow B,.Erfurth EM. A low individualized GH dose in young patients with childhood onset GH deficiency normalized serum IGF-I without significant deterioration in glucose tolerance. *Clinical Endocrinology* 1999;**50**:45-55.

28. Davies JS, Obuobie K, Smith J, Rees DA, Furlong A, Davies N *et al*. A therapeutic trial of growth hormone in hypopituitary adults and its influence upon continued prescription by general practitioners. *Clinical Endocrinology* 2000;**52**:295-303.

29. Drake WM, Coyte D, Camacho-Hubner C, Jivanji NM, Kaltsas G, Wood DF *et al*. Optimizing growth hormone replacement therapy by dose titration in hypopituitary adults. *Journal of Clinical Endocrinology & Metabolism* 1998;**83**:3913-9.

30. Hayes FJ, Fiad TM, McKenna TJ. Gender difference in the response of growth hormone (GH)-deficient adults to GH therapy. *Metabolism: Clinical & Experimental* 1999;**48**:308-13.

31. Murray RD, Skillicorn CJ, Howell SJ, Lissett CA, Rahim A, Shalet SM. Dose titration and patient selection increases the efficacy of GH replacement in severely GH deficient adults. *Clinical Endocrinology* 1999;**50**:749-57.

32. Riva G, Molinari E, Sartorio A, Morabito F. Psychosocial characteristics before and after growth hormone treatment in adults with GH deficit. *Acta Medica Auxologica* 1993;**25**:169-75.

33. Wiren L, Bengtsson BA, Johannsson G. Beneficial effects of long-term GH replacement therapy on quality of life in adults with GH deficiency. *Clinical Endocrinology* 1998;**48**:613-20.

34. Bengtsson BA, Abs R, Bennmarker H, Monson JP, Feldt-Rasmussen U, Hernberg-Stahl E *et al*. The effects of treatment and the individual responsiveness to growth hormone (GH) replacement therapy in 665 GH-deficient adults. KIMS Study Group and the KIMS International Board. *Journal of Clinical Endocrinology & Metabolism* 1999;**84**:3929-35.

35. Gibney J, Wallace JD, Spinks T, Schnorr L, Ranicar A, Cuneo RC *et al*. The effects of 10 years of recombinant human growth hormone (GH) in adult GH-deficient patients. *Journal of Clinical Endocrinology & Metabolism* 1999;**84**:2596-602.

36. Hernberg-Stahl E, Luger A, Abs R, Bengtsson BA, Feldt-Rasmussen U, Wilton P *et al*. Healthcare consumption decreases in parallel with improvements in quality of life during GH replacement in hypopituitary adults with GH deficiency. *Journal of Clinical Endocrinology & Metabolism* 2001;**86**:5277-81.

37. Monson JP, Abs R, Bengtsson BA, Bennmarker H, Feldt-Rasmussen U, Hernberg-Stahl E *et al*. Growth hormone deficiency and replacement in elderly hypopituitary adults. KIMS Study Group and the KIMS International Board. Pharmacia and Upjohn International Metabolic Database. [see comments]. *Clinical Endocrinology* 2000;**53**:281-9.

38. Badia X, Lucas A, Sanmarti A, Roset M, Ulied A. One-year follow-up of quality of life in adults with untreated growth hormone deficiency. *Clinical Endocrinology* 1998;**49**:765-71.

39. Carroll PV, Littlewood R, Weissberger AJ, Bogalho P, McGauley G, Sonksen PH *et al*. The effects of two doses of replacement growth hormone on the biochemical, body composition and psychological profiles of growth hormone-deficient adults. *European Journal of Endocrinology* 1997;**137**:146-53.

40. Holmes SJ,.Shalet SM. Factors influencing the desire for long-term growth hormone replacement in adults. *Clinical Endocrinology* 1995;**43**:151-7.

41. Bowling, A. Measuring health: a review of quality of life and measurement scales. 1997. Milton Keynes, Open University Press. Ref Type: Serial (Book,Monograph)

42. Sartorio A, Molinari E, Riva G, Conti A, Morabito F, Faglia G. Growth hormone treatment in adults with childhood onset growth hormone

deficiency: effects on psychological capabilities. *Hormone Research* 1995;**44**:6-11.

43. Hunt SM, McEwen J, McKenna SP. Measuring health status. London: Croom Helm, 1986.

44. McGauley GA, Cuneo RC, Salomon F, Sonksen PH. Psychological well-being before and after growth hormone treatment in adults with growth hormone deficiency. *Hormone Research* 1990;**33 Suppl 4**:52-4.

45. Chong JSWL, Benbow S, Foy P, Wallymahmed ME, Wile D, MacFarlane IA. Elderly people with hypothalamic-pituitary disease and growth hormone deficiency: lipid profiles, body composition and quality of life compared with control subjects. *Clinical Endocrinology* 2000;**53**:551-9.

46. Rosen T, Wiren L, Wilhelmsen L, Wiklund I, Bengtsson BA. Decreased psychological well-being in adult patients with growth hormone deficiency. *Clinical Endocrinology* 1994;**40**:111-6.

47. Brazier J, Harper R, Jones NMB. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;**305**:160-4.

48. Hunt SM, McKenna SP, Doward LC. Preliminary report on the development of a disease-specific instrument for assessing quality of life of adults with growth hormone deficiency. *Acta Endocrinologica* 1993;**128 Suppl 2**:37-40.

49. Hunt SM. Developing a measure of quality of life for adults with growth hormone deficiency. *Drug Information Journal* 1994;**28**:1-11.

50. McKenna SP,.Doward LC. Quality-of-life assessment of adults with growth hormone deficiency. Implications for drug therapy. [Review] [44 refs]. *Pharmacoeconomics* 1994;**6**:434-41.

51. Holmes SJ, McKenna SP, Doward LC, Hunt SM, Shalet SM. Development of a questionnaire to assess the quality of life of adults with growth hormone deficiency. *Endocrinology and Metabolism* 1995;**2**:1-69.

52. Wiren L, Whalley D, McKenna S, Wilhelmsen L. Application of a disease-specific, quality-of-life measure (QoL-AGHDA) in growth hormone-deficient adults and a random population sample in Sweden: validation of the measure by rasch analysis. [letter; comment]. *Clinical Endocrinology* 2000;**52**:143-52.

53. Barkan AL. The "quality of life-assessment of growth hormone deficiency in adults" questionnaire: can it be used to assess quality of life in hypopituitarism? *Journal of Clinical Endocrinology & Metabolism* 2001;**86**:1905-7.

54. Brazier J, Deverill M, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technology Assessment* 1999;**3**:1-164.

55. Anthony, D. Growth hormone (somatropin) for growth hormone deficient adults. 45. 1995. Bristol, Wessex Institute for Health Research and Development. Wessex DEC reports.
Ref Type: Report

56. Pharmacia and Upjohn. Submission to NICE. 2001. Pharmacia and Upjohn.
Ref Type: Generic

57. Eli Lilly. Submission to NICE. 2001. Eli Lilly.
Ref Type: Generic

58. Novo Nordisk. Submission to NICE. 2001. Novo Nordisk.
Ref Type: Generic

59. Anderson, K. M., Odell, P. M., Wilson, P. W. F., and Kannel, W. B. Cardiovascular disease risk profiles. American Heart Journal 121(1(pt2)), 293-298. 1991.
Ref Type: Generic

60. Morris, S. A comparison of economic modelling and clinical trials in the economic evaluation of cholesterol-modifying pharmacotherapy. Health Economics 6, 589-601. 1997.
Ref Type: Generic

61. De Laet, C. E. D. H, van Hout, B. A., Burger, H., Weel, A. E. A. M., Hofman, A, and Pols, H. A. P. Incremental cost of medical care after hip fracture and first vertebral fracture : the Rotterdam Study. Osteoporosis Int. 10, 66-72. 1999.
Ref Type: Generic

62. Marshall, D. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ 312, 1254-1259. 1996.
Ref Type: Generic

63. Tsuchiya, A., Brazier, J., and Roberts, R. Estimating preference-weights for the AQLQ from the EQ-5D. Euroqol Scientific Meeting Proceedings . 2001. Copenhagen, Denmark.
Ref Type: Report

64. Sanmarti A, Lucas A, Hawkins F, Webb SM, Ulied A. Observational study in adult hypopituitary patients with untreated growth hormone deficiency (ODA study). Socio-economic impact and health status. Collaborative ODA (Observational GH Deficiency in Adults) Group. [erratum appears in Eur J Endocrinol 2000 Feb;142(2):209]. *European Journal of Endocrinology* 1999;**141**:481-9.

65. Rosser, R, Cottee, M, and Rabin, R. Index of Health Related Quality of Life : measures of the quality of life. 1992. London, Royal College of Physicians.
Ref Type: Generic

66. Monson, J. and Nussey, S. Adult growth hormone replacement. Endocrinology . 2002.
Ref Type: Electronic Citation

67. DeWitt SJ, Sparks JW, Swank PB, Smith K, Denson SE, Landry SH. Physical growth of low birthweight infants in the first year of life: impact of maternal behaviors. *Early Human Development* 1997;**47**:19-34.

68. Murray RD, Skillicorn CJ, Howell SJ, Lissett CA, Rahim A, Smethurst LE *et al*. Influences on quality of life in GH deficient adults and their effect on response to treatment. *Clinical Endocrinology* 1999;**51**:565-73.

APPENDIX 1

Search strategy

Table 1 - Sources searched

EMBASE
Medline
NHS CRD DARE
NHS CRD EED
NHS CRD HTA
OHE HEED
PsycINFO
Quality of Life in Medicine
Web of Science

EMBASE
1980 to 2002
SilverPlatter WebSpirs 4.0 Version
Search undertaken January 2002

#86 #81 or #82 or #83 or #85
#85 growth hormone deficiency questionnaire*
#84 ghdq*
#83 QoL-AGHDA*
#82 #7 and #80
#81 #7 and #37
#80 #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57 or #58 or #59 or #60 or #61 or #62 or #63 or #64 or #65 or #66 or #67 or #68 or #69 or #70 or #71 or #72 or #73 or #74 or #75 or #76 or #77 or #78 or #79
#79 state trait anxiety
#78 stai
#77 sjoberg* mood*
#76 smq*
#75 self esteem scale*
#74 ses
#73 symptom* checklist* ninety
#72 symptom* checklist* 90
#71 scl 90
#70 social adjustment scale*
#69 sas
#68 schedule for affective disorder* near2 schizophrenia
#67 sads
#66 profile of mood state*
#65 poms
#64 nottingham health profile

#63 nhp
#62 minnesota multiphas*
#61 mmpi*
#60 mental fatigue syndrome*
#59 mfs
#58 life fulfilment scale*
#57 lfs*
#56 symptom rating test*
#55 srt*
#54 kellner* symptom*
#53 ksq*
#52 impact scale*
#51 hopkin* symptom*
#50 hscl*
#49 hamilton depression
#48 hds
#47 hospital anxiety near2 depression
#46 hads
#45 general health questionnaire*
#44 ghq
#43 disease specific questionnaire*
#42 dsq
#41 comprehensive psychopatholog*
#40 cprs
#39 beck* depression
#38 bdi
#37 #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36
#36 rosser
#35 qwb*
#34 quality of well being
#33 quality of wellbeing
#32 hui*
#31 health utilit*
#30 pgwb*
#29 psychological general wellbeing
#28 psychological general well being
#27 health* year* equivalen*
#26 hye*
#25 quality adjusted life year*
#24 qaly*
#23 eq5d
#22 eq 5d
#21 euroqol
#20 ql
#19 qol
#18 mos
#17 medical outcomes survey
#16 shortform

#15 short form
#14 sf*
#13 hql
#12 hrqol
#11 hrql
#10 life quality
#9 quality of life
#8 explode 'quality-of-life' / all subheadings
#7 #1 or #2 or #3 or #4 or #5 or #6
#6 'body-height' / all subheadings
#5 ghd
#4 gh deficien*
#3 growth hormone deficien*
#2 'growth-retardation' / all subheadings
#1 explode 'growth-disorder' / all subheadings

MEDLINE
1966 to 2002
Ovid Biomed version
Search undertaken January 2002

1 exp Growth Disorders/
2 growth hormone deficien$.tw.
3 growth deficien$.tw.
4 gh deficien$.tw.
5 ghd.tw.
6 or/1-5
7 exp Quality of Life/
8 quality of life.tw.
9 life quality.tw.
10 hrql.tw.
11 hrqol.tw.
12 hql.tw.
13 sf$.tw.
14 short form.tw.
15 shortform.tw.
16 medical outcomes survey.tw.
17 mos.tw.
18 qol.tw.
19 ql.tw.
20 euroqol.tw.
21 eq 5d.tw.
22 eq5d.tw.
23 qaly$.tw.
24 quality adjusted life year$.tw.
25 hye$.tw.
26 health$ year$ equivalen$.tw.
27 psychological general well being.tw.
28 psychological general wellbeing.tw.
29 pgwb$.tw.

30 health utilit$.tw.
31 hui$.tw.
32 quality of wellbeing.tw.
33 quality of well being.tw.
34 qwb$.tw.
35 rosser.tw.
36 or/7-35
37 bdi.tw.
38 beck$ depression.tw.
39 cprs.tw.
40 comprehensive psychopatholog$.tw.
41 dsq.tw.
42 disease specific questionnaire$.tw.
43 ghq.tw.
44 general health questionnaire$.tw.
45 hads.tw.
46 (hospital anxiety adj2 depression).tw.
47 hds.tw.
48 hamilton depression.tw.
49 hscl$.tw.
50 hopkin$ symptom$.tw.
51 impact scale$.tw.
52 ksq$.tw.
53 kellner$ symptom$.tw.
54 srt.tw.
55 symptom rating test$.tw.
56 lfs$.tw.
57 life fulfilment scale$.tw.
58 mfs.tw.
59 mental fatigue syndrome$.tw.
60 mmpi$.tw.
61 minnesota multiphase.tw.
62 nhp.tw.
63 nottingham health profile.tw.
64 poms.tw.
65 profile of mood state$.tw.
66 sads.tw.
67 (schedule for affective disorder$ adj2 schizophrenia).tw.
68 sas.tw.
69 social adjustment scale$.tw.
70 scl 90.tw.
71 symptom$ checklist$ 90.tw.
72 symptom$ checklist$ ninety.tw.
73 ses.tw.
74 self esteem scale$.tw.
75 smq$.tw.
76 sjoberg$ mood$.tw.
77 stai.tw.
78 state trait anxiety.tw.
79 or/37-78

80 6 and 36
81 6 and 79
82 80 or 81
83 QoL-AGHDA$.tw.
84 assessment of growth hormone deficiency in adults.tw.
85 ghdq$.tw.
86 growth hormone deficiency questionnaire$.tw.
87 or/83-86
88 82 or 87

NHS CRD DARE, EED and HTA
All years
Internet version
Search undertaken January 2002

All fields searched using the following terms:

Growth deficien*
Growth disorder
Gh deficienc*
Ghd

OHE HEED
CD ROM version
Search undertaken January 2002

Fields searched:

Abstract
All data
Article title
Book title
Keywords

Terms searched:

Growth deficien*
Growth disorder
Gh deficienc*
Ghd

PsycINFO
1887 - 2002
SilverPlatter WebSpirs 4.0 Version
Search undertaken January 2002

#6 #1 or #2 or #3 or #4 or #5
#5 ghd
#4 gh deficien*

#3 growth deficien*
#2 growth hormone deficien*
#1 growth disorder*

Quality of Life in Medicine
CD ROM version
Search undertaken January 2002

All fields searched using the following terms:

Growth deficien*
Growth disorder
Gh deficienc*
Ghd

Web of Science
Internet version
Searches undertaken January 2002

Topic=(growth disorder or growth retardation or growth hormone deficienc* or gh deficienc* or ghd) and (quality of life or life quality or hrql or hrqol or hql or sf 36 or sf36 or sf12 or sf 12 or sf60 or sf 60 or short form or shortform or medical outcomes survey or mos or qol or ql or euroqol or eq 5d or eq5d or qaly* or quality adjusted or hye* or healthy year* equivalent* or psychological general wellbeing or psychological general well being or pqwb* or health utilit* or hui* or quality of wellbeing or quality of well being or qwb* or rosser); DocType=All document types; Language=All languages; Databases=SCI-EXPANDED, SSCI, A&HCI; Timespan=All Years;

Topic=(growth disorder or growth retardation or growth hormone deficienc* or gh deficienc* or ghd) and (bdi or beck* depression or cprs or comprehensive psychopatholog* or dsq or disease specific questionnaire or ghq or general health questionnaire or hads or hospital anxiety or hds or hamiltondepression or hscl* or hopkin* symptom* or impact scale* or ksq or kellner* smyptom* or srt or symptom rating test or lfs or life fulfilment scale or mfs or mental fatigue syndrome or mmpi or minnesota multiphas* or nhp or nottingham health profile or poms or profile of mood state or sads or schedule for affective disorder or sas or social adjustment scale or scl 90 or symptom checklist 90 or ses or self esteem sclae or smq or sjoberg* mood or stai or state trait anxiety); DocType=All document types; Language=All languages; Databases=SCI-EXPANDED, SSCI, A&HCI; Timespan=All Years;

Topic=QoL-AGHDA or assessment of growth hormone deficiency in adults or ghdq or growth hormone deficiency questionnaire; DocType=All document types; Language=All languages; Databases=SCI-EXPANDED, SSCI, A&HCI; Timespan=All Years;

APPENDIX 2 - INCLUDED STUDIES

Reference and design	Intervention	Participants	Outcome measures
Author: Ahmad et al 2001[26] Country: UK Type of study: prospective uncontrolled Length of treatment: 3 months Loss to follow-up: not stated Jadad score: not applicable	Name of GH: subjects were randomly assigned either Genotropin or Humatrope Dose: treatment was initiated at 0.4 (Genotropin) or 0.5 (Humatrope) IU/day and titrated to achieve and maintain IGF-1 standard deviation score between the median and upper end of the age-related reference range (mean dose at 3 months 0.80±0.12 IU/day – range 0.5-1.2 IU) Did any patients receive somatropin before trial: no Other hormone replacements: additional pituitary hormone – 43 patients	Total number: 46 Isolated or multiple deficiencies: multiple 43 Comorbidities: not stated Adult or childhood onset: adult Causes of GH deficiency: nonfunctioning pituitary adenoma - 2 prolactinoma - 12 Cushing's disease – 4* glioma - 2 craniopharyngioma - 2 dermoid cyst - 1 Rathke's cystic tumour - 1 epidermoid cyst - 1 acromegaly - 1 * in remission for a mean of 10.2 years (range 6-14) Definition of GH deficiency: peak GH response <9 mU/l to insulin-induced hypoglycaemia (blood glucose <2.2 mmol/l) or glucagon stimulation test Peak GH concentrations: <0.5 mU/l – 37 patients 0.5-5.0 mU/l – 7 patients Mean duration of GH deficiency: 10.6 years (range 0.75-22 years) Sex: 22 men, 24 women Mean age at diagnosis: 50.4	Quality of life scales used: Qol-AGHDA

		(range 26-72) Mean time from diagnosis to recruitment: 10.6 years (range 0.75-22)	
Results **Baseline** **QoL-AGHDA** 13.3±6.4	**QoL-AGHDA** * P<0.01 ** P<0.001 † P<0.001 vs 1 month	**1 month** 11.5±6.6*	**3 months** 10.0±6.6** †
Methodological comments Randomisation method: not applicable Patients blinded to treatment: no Outcome assessors blinded to treatment: no Baseline characteristics: Dropouts and withdrawals: no data Compliance: no data **General comments** Conflict of interests: supported by Eli Lilly & Co and Pharmacia & Upjohn; partly funded by the Royal Liverpool & Broadgreen Hospital Trust Other:			

Adverse events leading to withdrawal	No data

Number of specific adverse events	No data

Reference and design	Intervention	Participants	Outcome measures
Author: Attanasio et al, 1997[14] Country: Multinational Type of study: RCT Length of treatment: 18 months Loss to follow-up: Not reported clearly, 7 withdrew due to side effects Jadad score: 3/5	Brand name of somatropin: Humatrope, Lilly Research Centre Dose: 6.25 μg/kg for the first 4 weeks, then 12.5 μg/kg per day maximum for 6 months. All patients then on open-label somatropin for an additional 12 months or placebo. Did any patients receive somatropin before trial: no somatropin treatment in previous 2 years Other hormone replacements: Replacement therapy with cortisol, thyroxine, sex steroids and vasopressin stable for at least 6 months before study	Total number: 173 Isolated or multiple deficiencies: Not reported Comorbidities: Not reported Adult or childhood onset: 99 AO, 74 CO Cause of GH deficiency: *Adult onset* Functional adenoma 30; Nonfunctional adenoma 25; Craniopharyngioma 19; Dysgerminoma, pinealoma, epidermoid cyst 6; Posttubercular condition, histiocytosis 2; Trauma, Sheehan syndrome, empty sella 9; Idiopathic, hypothalamic origin 7 *Childhood onset* Idiopathic: Isolated GH deficiency 19; GH plus TSH deficiency 7: GH plus LH/FSH deficiency 4; Multiple deficiency 30; Trauma, empty sella, posttubercular condition 4; Craniopharyngioma, dysgerminoma 3 Definition of GH deficiency: peak serum GH level <5 μ g/L Mean GH concentrations: Mean not reported	Quality of life scales used: NHP

		Sex: 116 male, 57 female Mean age: Childhood onset 28.8 ± 8; adult onset 43.5 ± 10 (p<0.05)	

Mean NHP scores (CI) and mean reference level of age and sex matched controls in CO and AO patients at baseline.

Baseline Scores

	Mean (CI)	**Reference level**
Childhood onset (n=61)		
Social isolation	5.9 (3.2-8.6)	4.6
Physical mobility	8.8 (4.7-12.9)	1.4
Emotional reaction	14.0 (9.4-18.6)	8.5
Energy level	14.8 (8.6-20.9)	6.4
Sleep	14.8 (9.4-20.1)	8.5
Pain	8.2 (4.5-11.9)	2.8
Adult onset (n=87)		
Social isolation	7.4 (4.7-10.1)	5.1
Physical mobility	17.2* (12.8-21.7)	3.4
Emotional reaction	14.7 (11.3-18.1)	9.6
Energy level	28.4* (22.0-34.7)	12.0
Sleep	20.7 (15.8-25.6)	12.1
Pain	9.5 (6.3-12.7)	5.0

p<0.01 vs. Childhood onset

End of Trial scores
No values reported. Significant improvement for placebo and somatropin treated patients in both AO and CO groups during double blind therapy phase. There was a significant treatment effect compared to placebo for social isolation and physical mobility in AO but not CO patients (p<0.01). In AO patients these improvements persisted with somatropin therapy and at 18 months physical mobility and energy level were significantly (p<0.01 for each) improved from baseline. In the CO group there were no significant effects of somatropin therapy for any NHP scores at 12 and 18 months.

Methodological comments
Randomisation method: Not reported
Patients blinded to treatment: Yes
Outcome assessors blinded to treatment: Not reported
Baseline characteristics: Baseline characteristics were compared between those with childhood onset and those with adult onset only.
Dropouts and withdrawals: 7 patients withdrew due to side effects
Compliance:

General comments
Conflict of interests: This work was supported by Eli Lilly Industries
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal	

AO group: 1 recurrence of craniopharyngioma	CO group: 1 hepatitis
AO group: 1 because of hypertension and arthralgia	CO group: 1 increased liver enzyme levels
AO group: 1 abnormal glucose tolerance	CO group: 1 joint disorder.
AO group: 1 viral illness	

Number of specific adverse events 6 month endpoint	AO somatropin/ somatropin	AO placebo/ somatropin	CO somatropin/ somatropin	CO placebo/ somatropin
Oedema, peripheral oedema	15 (28.8%)	2 (4.3%)	2 (6.3%)	0
Arthralgia, myalgia, joint disorder	12 (23.1%)	3 (6.5%)	2 (6.3%)	0
Parasthesia, hypesthesia	3 (5.8%)	2 (4.3%)	2 (6.3%)	0
Hypertension	0	1 (2.2%)	0	0

Number of specific adverse events 18 month endpoint*	AO somatropin/ somatropin	AO placebo/ somatropin	CO somatropin/ somatropin	CO placebo/ somatropin
Oedema, peripheral oedema	17 (32.7%)	13 (28.3%)	3 (9.4%)	3 (10%)
Arthralgia, myalgia, joint disorder	19 (36.5%)	11 (23.9%)	4 (12.5%)	2 (6.7%)
Parasthesia, hypesthesia	10 (19.2%)	6 (13%)	3 (9.4%)	0
Hypertension	4 (7.7%)	2 (4.3%)	0	0

*GH treatment only; throughout 8 months of study for somatropin/somatropin and from 6 month baseline to 18 month endpoint for placebo/somatropin.

Reference and design	Intervention	Participants	Outcome measures
Author: Baum et al., 1998[15] Country: USA Type of study: RCT Length of treatment: 18 months Loss to follow-up: all patients accounted for. Jadad score: 5/5	Name of somatropin: Nutropin, Genetech, South San Francisco, CA Dose: Initial starting dose of 10 μg/kg per day Did any patients receive somatropin before trial: Not reported Other hormone replacements: Standard thyroid or adrenal hormone replacement therapy was in place for at least 6 months before the trial in those patients that needed it.	Total number: 40 Isolated or multiple deficiencies: isolated 2, multiple 38 Comorbidities: none reported Adult or childhood onset: Adult Definition of GHD: < 5 μg/L in standard stimulation test Cause of GHD: 19 with clinically non-functioning pituitary adenoma, 11 with prolactinoma, 5 with craniopharyngioma, 2 with Cushing's, 2 with idiopathic hypopituitarism and 1 with apoplexy. Mean GH concentrations: 4±2 μg/kg per day Sex: all male Mean age: median 51 (range: 24-64)	Quality of life scales used: NHP, PGWB, GHQ, MMPI-2

Mean ± SEM

NHP

	Somatropin	Placebo	GH (Mean change)	Placebo (Mean Change)
Emotional Reactions	7.8 ± 3.1	12.2 ± 4.9	10.7 ± 4.9 (2.2 ± 3.6)	3.0 ± 1.7(-6.7 ± 4.4)
Energy	18.3 ± 6.2	19.3 ± 8.2	15.6 ± 9.1 (-4.4 ± 9.1)	8.9 ± 5.1(-4.8 ± 11.0)
Pain	3.1 ± 2.5*	12.5 ± 4.8	4.2 ± 2.9† (4.2 ± 2.9)	2.5 ± 1.8 (-6.7 ± 4.2)
Sleep	15.0 ± 5.6	14.0 ± 5.1	8.0 ± 3.3 (-1.3 ± 5.3)	10.7 ± 3.3 (1.3 ± 3.6)
Social Isolation	3.2 ± 1.7	3.0 ± 2.2	1.3 ± 1.3 (-1.4 ± 1.4)	0.0 ± 0.0 (-1.3 ± 1.3)
Physical Mobility	5.3 ± 2.8	10.5 ± 4.0	3.3 ± 1.9 (0.9 ± 1.6)	5.8 ± 3.0 (-4.5 ± 2.1)

PSGW (maximum 110)

	83 ± 13	85 ± 16	84 ± 18	86 ± 8

GHQ (maximum 180)

	37 ± 17	36 ± 19	36 ± 19	31 ± 8

MMPI-2 (T scores: mean=50, SD=10)

Hypochondriasis				
	52 ± 10	55 ± 11	57 ± 9	53 ± 11
Depression	55 ± 11	55 ± 10	54 ± 6	55 ± 11
Hysteria	52 ± 10	55 ± 9	57 ± 12‡	53 ± 10

*p<0.05 compared with baseline value for placebo group
†p<0.05 for change from baseline to 18 months compared with placebo
‡ p<0.03 for change from baseline to 18 months compared with placebo

Methodological comments
Randomisation method: computerised randomisation
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: No significant differences between the two groups
Dropouts and withdrawals: 5 patients withdrew from the GHD group and one from the placebo group
Compliance: Measured by vial count.

General comments
Conflict of interests: This study was supported by a research grant from Genentech.
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1+1
Was the study described as double blind?	1+1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal
1 in placebo group with pneumonia
1 in somatropin group with seizure
1 in somatropin group because of tachycardia
1 in somatropin group due to cerebrovascular accident

Number of specific adverse events
2 in somatropin group with oedema
1 in somatropin group with myalgias

Reference and design	Intervention	Participants	Outcome measures
Author: Bengtsson et al 1999[34] Country: European multinational Type of study: observational (data from KIMS dataset) Length of treatment: 12 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Genotropin Dose: initiated at a maximum of 0.125 IU/kg/week, subsequently increased to a maximum of 0.25 IU/kg/week according to individual patient requirements. The guidelines for therapy did not preclude the use of dose titration independent of body weight, based on clinical response & serum IGF-1 measurements Mean maintenance doses after 6 and 12 months were 0.43 and 0.53 mg/day for men and women respectively Did any patients receive somatropin before trial: some did, but it had been discontinued at least 6 months before enrolment Other hormone replacements: no data	Total number: 665 Isolated or multiple deficiencies: isolated 64, multiple 601 Comorbidities: 3% on hypolipidaemic drugs Adult or childhood onset: adult 493, childhood 172 Causes of GHD: stated only for isolated GHD (pituitary tumour 41%, craniopharyngioma 4.7%, idiopathic 28%) Definition of GHD: not given. Peak GH response to stimulation of <3 µg/l in 616 patients (97%), 3-4.9 µg/l in 17 patients and >5 µg/l in 1 patient Mean duration of GHD: not stated Sex: 333 men, 332 women Mean age: 44	Quality of life scales used: QoL-AGHDA

Results

QoL-AGHDA

	Baseline	6 months		12 months	
		Mean change	Median change	Mean change	Median change
Men	7.4	-2.2	-1*	-2.8	-1**
Women	9.8	-2.8	-3*	-4.8	-4*

* P<0.0001
** P=0.0004

Patients receiving the highest doses of somatropin were said to demonstrate the greatest improvement in quality of life (r=0.28, P<0.01). A significant change in score was seen after 6 months in patients with adult-onset GHD, but not in those with childhood-onset GHD; however, the number of patients included in this analysis was small (36 childhood-onset and 99 adult-onset patients), and the reason for this was not given.

Methodological comments

Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics:
Dropouts and withdrawals: not stated
Compliance: no information provided

General comments
Conflict of interests: study supported by Pharmacia and Upjohn
Other:

Number of specific adverse events	No data provided

Reference and design	Intervention	Participants	Outcome measures
Author: Bengsson et al, 1993[4] Country: Sweden Type of study: Randomised, cross-over double blind trial Length of treatment: 6 months then cross over to other treatment Loss to follow-up: one patient withdrew Jadad score: 5/5	Name of somatropin: Humatrope, Eli Lilly Co. Dose: maximum 0.5 U/kg/week (0.026 mg/kg/day) Did any patients receive somatropin before trial: no Other hormone replacements: glucocorticoids, thyroid hormones, bromocriptine and sex hormones	Total number: 10 Isolated or multiple deficiencies: multiple Comorbidities: not reported Adult or childhood onset: all adult onset Definition of GHD: < 5 μg/L in response to two stimuli Cause of GHD: chromophobe adenoma 4; prolactinoma 5, meningeoma 1 Mean GH concentrations: <1mU/L Sex: 9 male, 1 female Mean age: range 34-58 years	Quality of life scales used: Comprehensive Psychological Rating Scale (CPRS) and the Symptom Check List–90 (SCL-90)

Results
No baseline scores were reported.

CPRS
No values reported. After 26 weeks of somatropin treatment, seven patients had a decreased score, one had an unchanged score and one had an increased score ($p<0.05$).

SCL-90
No significant changes in SCL-90 results were noted.

There were four patients in the somatropin/placebo group. All 4 patients had higher scores on the CPRS scale 26 weeks after withdrawal of somatropin. On the SCL-90 scale, 2 patients had an increased score and 2 had unchanged scores. One patient suffered from withdrawal to the extent that she was evaluated for mild depression.

Methodological comments
Randomisation method: randomisation codes
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: cross over trial
Dropouts and withdrawals: one patient withdrew
Compliance:

General comments
Conflict of interests: The work was supported in part by a grant from Eli Lilly Co.
Other:

Quality assessment for RCTs (Jadad Score)	

Question	Score
Was the study described as randomised?	1+1
Was the study described as double blind?	1+1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal
1 patient withdrew due to atrial fibrillation

Number of specific adverse events
Oedema n=2
Transient oedema n=1
Swollen fingers n=1
Tinnitus n=1
Carpal tunnel syndrome n=1
Arthralgia n=1

Reference and design	Intervention	Participants	Outcome measures
Author: Beshyah et al 1995[16] Country: UK Type of study: RCT Length of treatment: 6 months as an RCT, followed by 12 months when all subjects, including those initially receiving placebo, were given somatropin therapy Loss to follow-up: 2 withdrew from the somatropin group during the 6-month double-blind phase (1 due to fluid retention attributed to therapy, 1 as a result of an accident). 11 withdrew from the open phase because of lack of perceived benefit or adverse events Jadad score: 4/5	Name of somatropin: Norditropin Dose: 0.02- 0.05 IU/kg daily (starting dose 0.05 IU/kg (maximum 4 IU/day), later adjusted according to patients' tolerance by 25 or 50% reductions) or placebo Did any patients receive somatropin before trial: not stated Other hormone replacements: thyroxine, cortisol, fluorocortisone, desmopressin, sex steroids	Total number: 40 Isolated or multiple deficiencies: the majority had Comorbidities: priority multiple/deficiencies not stated Adult or childhood onset: somatropin group: 16 adult, 4 childhood Placebo group: 16 adult, 4 childhood Causes of GHD: idiopathic – somatropin group 3, placebo group 1 prolactinoma – somatropin group 7, placebo group 3 non-functioning adenoma – somatropin group 2, placebo group 7 craniopharyngioma – somatropin group 3, placebo group 6 Cushing's disease – somatropin group 2, placebo group 2 other pathology - somatropin group 3, placebo group 1 Definition of GHD: serum GH response of <6 mU/l to insulin-induced hypoglycaemia (blood glucose <2.0 mmol/l) or to oral clonitidine (50 µl) Mean GH concentrations: not stated Mean duration of GHD: somatropin group 8 years (range 1-20) Placebo group 9	Quality of life scales used: General Health Questionnaire (GHQ-60), Comprehensive Psychopathological Rating Scale (CPRS)

		years (range 1-23) Sex: somatropin group: 9 men, 11 women Placebo group: 10 men, 10 women Mean age: somatropin group: 46 (range 19-67) Placebo group: 42 (range 26-59)	

Results at baseline

somatropin placebo
GHQ – median (range)
3 (0-47) 12 (0-37)*
P<0.004 vs somatropin group

CPRS – median (range)
8 (4-34) 20 (3-31)**
P<0.06 vs somatropin group

RCT: Results at 6 months

somatropin placebo
GHQ – median (range)
(0-55) 4 (0-47)*

CPRS – median (range)
7 (1-23) *15 (3-23)*

During the RCT phase, 11 patients on somatropin and 4 on placebo reported improvement in general well-being (increased energy and stamina, more alertness, a more positive attitude to life, increased capacity to work longer hours, and the ability to perform more physical activities (P<0.01).

Open phase

Baseline	6 months (n=34)	12 months (n=27)	18 months (n=11)
GHQ – median (range)			
3 (0-47)	4 (0-55)	1 (0-44)	1 (0-47)
CPRS – median (range)			
12 (3-34)	9 (1-23)*	5 (2-18)**	6 (0-19)

* P<0.05 vs baseline
** P<0.001 presumably vs baseline

Methodological comments

Randomisation method: not given
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not stated
Baseline characteristics: significant differences between the two groups in GHQ and CPRS scores
Dropouts and withdrawals: 2 from the controlled phase
Compliance: no data

General comments

Conflict of interests: study supported by Novo Nordisk Pharmaceuticals
Other: only 11 patients (ie 11 of the 20 in the original treatment group) still chose to be on treatment at 18 months. 7 patients in the treatment group needed dose reductions because of side effects in the first 6 months. 2 of the placebo group did not accept treatment in the open phase. 11 withdrew from the open phase because of adverse events (n=3) or perceived lack of benefit (n=8)

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	2
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal (RCT phase)	Somatropin	Placebo
Generalised oedema, tiredness	1	0
Parachuting accident	1	0

Adverse events leading to withdrawal (open phase)	
Fluid retention	1
Fluid retention + bilateral carpal tunnel syndrome	1
Emotional lability and depression	1
Expansion of a prolactinoma	1
Diabetic glucose tolerance test	2
Possible pituitary apoplexy	1

Number of specific adverse events (excluding withdrawals) – RCT phase	Somatropin	Placebo
Disturbed sleep pattern	2	0
Oedema	7	4
Carpal tunnel syndrome	2	0
Dizziness	1	1
Tiredness	1	2
Arthralgia	2	0
Limb or finger swelling	3	0
General or limb ache	0	2
Vaginal bleeding	0	1
Puffy face	0	1
Bloating	0	1
Total events	18	12
Total patients with events	11	7

Number of specific non-accidental adverse events (including withdrawals) – all patients who received somatropin	Somatropin
Disturbed sleep pattern	2
Oedema	20
Carpal tunnel syndrome	5
Dizziness	1
Tiredness	2
Arthralgia	4
Limb or finger swelling	6
General or limb ache	2
Vaginal bleeding	1
Puffy face	3
Bloating	7
Weight gain	4
Stiff fingers	2
Polyuria	1
Acne	1
Muscular discomfort	2
Headache	1
Transient BP elevation	1
Skin rash, itching	1
Elbow pain	1
Expansion of prolactinoma	1

Emotionally unstable	2
Depression	1
Low backache	1
Dyspnoia	1
Chest pain	1
Total events	74
Total patients with events	34

Reference and design	Intervention	Participants	Outcome measures
Author: Bulow & Erfurth 1999[27] Country: Sweden Type of study: prospective uncontrolled Length of treatment: 9 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Genotropin Dose: intended to achieve a serum IGF-1 level in the middle of the normal range (median dose at end of study 0.14 IU/kg/week) Did any patients receive somatropin before trial: all, but none more recently than 5 years previously Other hormone replacements: glucocorticoids, thyroid hormone, ADH-analogue and gonadal steroids as required	Total number: 10 Isolated or multiple deficiencies: 3 isolated, 7 multiple Comorbidities: none stated Adult or childhood onset: childhood Causes of GHD: craniopharyngoma - 3 suprasellar cyst - 1 prolactinoma - 1 optic glioma - 1 idiopathic hypopituitarism - 2 idiopathic GHD - 2 Definition of GHD: serum GH response of ≤1.6 mIU/l to insulin-induced hypoglycaemia (blood glucose <2.2 mmol/l) Mean GH concentrations: no data Mean duration of GHD: not stated Sex: 8 men, 2 women Median age: 27 (range 21-28)	Quality of life scales used: QoL-AGHDA
Results **Baseline** *QoL-AGHDA (median + range)* 6 (1-23)		**9 months** *QoL-AGHDA (median + range)* 2 (0-18) (P=0.008 vs baseline)	

Methodological comments

Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics:
Dropouts and withdrawals: no data
Compliance: no data

General comments

Conflict of interests: funded by the University of Lund and Pharmacia & Upjohn
Other:

Adverse events leading to withdrawal	No data

Number of specific adverse events	
Oedema	2
Arthralgia	1
Total events	3
Total patients with events	2

Reference and design	Intervention	Participants	Outcome measures
Author: Burman et al, 1995[24] Country: Sweden Type of study: Cross-over, double blind trial Length of treatment: 9 months for each treatment period separated by 3 month washout period Loss to follow-up: one patient withdrew but was replaced Jadad score: 2/5	Name of somatropin: Norditropin, Novo Nordisk Pharma Dose: 2 U/m^2 maximum Did any patients receive somatropin before trial: no Other hormone replacements: levothyroxine, adrenal steroids, sex steroids and desmopressin	Total number: 36 Isolated or multiple deficiencies: not reported Comorbidities: 15 (42%) had less than full time employment Adult or childhood onset: all adult onset Definition of GHD: 3 μg/L or less during insulin-induced hypoglycaemia Cause of GHD: 1 with craniopharyngioma and 1 with a hypothalamic disorder (onset during adolescence); the others had acquired pituitary insufficiency as a result of surgery for and/or irradiation of a pituitary tumour: non-functioning adenoma n=18; craniopharyngioma n=4, prolactinoma n=4, ACTHoma n=3 OR as a consequence of empty sella n=3, injury n=1 or pituitary apoplexy n=1. Mean GH concentrations: 2.4 U (1.25 U/m^2) Sex: 21 male, 15 female Mean age: 46 (range: 28-57 years)	Quality of life scales used: Hopkins Symptom Checklist (HSCL-56); Nottingham Health Profile (NHP), Psychological General Well-Being Index (PGWB) and 12 item partner questionnaire

Results

HSCL (mean ± sd)			*p values*		
Baseline (A)	Post (B) somatropin	Post placebo (C)	A vs. B	A vs. C	B vs. C
Total score					
89 ± 18.9	80.2 ± 18.5	84.0 ± 21.3	0.0003	0.06	0.06
Anxiety					
10.8 ± 2.9	9.1 ± 2.0	9.6 ± 2.6	0.0001	0.002	0.06
Cognition					
13.2 ± 3.6	11.7 ± 3.6	12.3 ±3.9	0.0008	0.07	0.18
Depression					
20.5 ± 6.6	18.5 ± 6.2	19.8 ± 6.8	0.01	0.39	0.14
Interpersonal Sensitivity					
10.9 ± 2.8	10.0 ± 3.0	10.6 ± 3.1	0.05	0.49	0.11
Somatization					
20.3 ± 4.9	19.0 ± 5.0	19.0 ± 5.0	0.06	0.04	0.95
Fearfulness					
16.9 ± 4.2	14.8 ± 3.3	15.3 ± 3.8	0.0001	0.003	0.20
Inferiority					
5.9 ± 1.9	5.5 ± 2.0	5.9 ± 2.1	0.15	1.0	0.07
Tension					
3.3 ± 0.9	2.8 ± 0.8	2.9 ± 1.0	0.005	0.07	0.23
NHP					
Total score					
16.7 ± 15.7	10.4 ± 14.2	14.0 ± 17.9	0.01	0.21	0.08
Emotions					
23.1 ± 25.3	12.1 ± 20.9	16.5 ± 24.1	0.003	0.08	0.14
Sleep					
13.4 ± 19.1	12.7 ± 21.9	15.3 ± 21.6	0.80	0.39	0.27
Energy					
37.1 ± 39.6	16.4 ± 24.2	25.1 ± 38.6	0.003	0.04	0.16
Pain 8.7 ± 18.8	8.7 ± 16.9	8.8 ± 21.7	0.99	0.96	0.97
Social Isolation					
9.9 ± 21.9	4.5 ± 14.6	8.5 ± 19.6	0.09	0.65	0.08
Physical Activity					
7.8 ± 11.2	7.7 ± 12.6	9.7 ± 14.4	0.97	0.23	0.21
PGWB					
Total score					
92.0 ± 15.5	97.4 ± 15.4	93.9 ± 16.6	0.04	0.43	0.16
Anxiety					
21.6 ± 4.8	23.9 ± 3.9	22.9 ± 3.9	0.009	0.07	0.13
Depression					
14.7 ± 3.0	15.4 ± 2.8	15.0 ± 3.2	0.14	0.63	0.44
Well being					
14.3 ± 3.5	15.3 ± 3.7	14.5 ± 3.9	0.11	0.76	0.14
Self-control					
14.1 ± 3.0	14.9 ± 2.7	14.4 ± 2.7	0.13	0.55	0.19
Health					
14.1 ± 2.8	14.2 ± 2.7	13.9 ± 3.0	0.80	0.59	0.52
Vitality					
14.2 ± 3.9	15.8 ± 3.5	14.7 ± 4.4	0.03	0.34	0.10

Partner Questionnaire

	Placebo (%)	somatropin (%)	p
More alert	0.0	69.0	<0.0001
More active	3.7	1.8	<0.001
Higher endurance	3.6	60.7	<0.0001
Less easily annoyed	7.1	28.6	<0.10
Less worried	6.9	37.9	<0.05
More extrovert	3.4	37.9	<0.01
More industrious	3.3	46.7	<0.001
More happy	11.1	48.1	<0.01
Better looks	10.3	51.7	<0.01
More satisfied with his/her occupation	7.7	34.6	<0.05
Fewer family conflicts	3.4	24.1	<0.10
Better personal relationships	3.4	34.5	<0.01

Methodological comments
Randomisation method: none described
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: yes
Baseline characteristics: crossover trial
Dropouts and withdrawals: one withdrawal who was replaced.
Compliance:

General comments
Conflict of interests: The work was supported by Novo Nordisk Pharma AB
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	0
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal	None reported

Number of specific adverse events	None reported

Reference and design	Intervention	Participants	Outcome measures
Author: Cuneo et al 1998[17] Country: Australia Type of study: RCT Length of treatment: 6 months RCT, followed by 6 months when all subjects, including those receiving placebo initially, were given somatropin therapy Loss to follow-up: 3 subjects were said to have withdrawn consent before initiation of treatment. 19 patients withdrew from the somatropin group and 11 from the somatropin/placebo group; 13 of these withdrew because of side effects. However, results were only available for 115 at 12 months Jadad score: 4/5	Name of somatropin: Genotropin Dose: 0.125 U/kg/week (maximum 2 IU/day) for first month, and 0.25 U/kg/week for following 5 months, in 7 daily doses. Dose reduced if side effects developed. Maximum daily dose (irrespective of body weight) = 4 IU/day. If subjects suffered oedema or arthralgia, the dose was reduced to 0.125 U/kg/week, and only increased with resolution of the symptoms. If side effects were persistent, therapy was stopped, and resumed with the resolution of symptoms, with the dose remaining at 0.125 U/kg/week. Average maintenance dose at 9 months 2.6±0.8 IU/day Did any patients receive somatropin before trial: not stated Other hormone replacements: an unspecified number of patients received pituitary hormone replacement	Total number: 166 Isolated or multiple deficiencies: not stated Comorbidities: There were no differences in medications between the groups. Overall, 10% received bromocryptine; 9% received antidepressants, anxiolytics or hypnotic medication; 6% anticonvulsants; 7% asthma treatment; 4% treatment for hypercholesterolaemia Adult or childhood onset: approximately a third of subjects were said to have received somatropin treatment for short stature in childhood Causes of GHD: pituitary tumour – somatropin group 48, placebo group 39 irradiation – somatropin group 13, placebo group 19 craniopharyngioma – somatropin group 6, placebo group 15 idiopathic – somatropin group 8, placebo group 13 Cushing's disease – somatropin group 6, placebo group 3 trauma – somatropin group 5, placebo group 1 septooptic dysplasia – somatropin group 1, placebo group 0 other – somatropin group 12, placebo group 11	Quality of life scales used: Nottingham Health Profile (NHP) part I, GH Deficiency Questionnaire (GHDQ), social history

		Definition of GHD: peak GH level <5 mU/l after insulin-induced hypoglycaemia Mean GH concentrations: not stated Estimated mean duration of GHD: somatropin group: 9.3±0.8 Placebo group: 12.1±1.0 P<0.05 Sex: somatropin group: 50 men, 33 women Placebo group: 41 men, 39 women Mean age: somatropin group: 41.2±1.5 Placebo group: 39.8±1.5	

Results at baseline

NHP (mean ± SEM)

Energy	
somatropin (n=73)	1.03±0.06
placebo (n=76)	1.17±0.05
Emotional reaction	
somatropin (n=72)	1.38±0.12
placebo (n=78)	0.70±0.07
Social isolation	
somatropin (n=73)	0.48±0.06
placebo (n=79)	0.24±0.03
Sleep	
somatropin (n=73)	1.14±0.07
placebo (n=79)	0.99±0.06
Pain	
somatropin (n=73)	0.77±0.10
placebo (n=79)	0.23±0.06
Mobility	
somatropin (n=73)	0.54±0.07
placebo (n=78)	0.38±0.08
SUM score	
No data	

For each item on the NHP, between 37 and 77% of subjects scored 0 at baseline, thus making it impossible to measure any improvement in their quality of life

GHDQ

No differences between groups at baseline

Social history

Satisfaction with lives	
somatropin group	63%
Placebo group	73%

Although the placebo-controlled phase ends at 6 months, 12 month scores are also given for both groups, as the study also functions as a before-and-after study for both groups

NHP (mean (SEM)

	6 months	*12 months*
Energy		
somatropin	(n=70) 0.55±0.05	(n=54) 0.23±0.04*
placebo	(n=76) 0.56±0.04**	(n=61) 0.36±0.04
Emotional reaction		
somatropin	(n=70) 0.65±0.07	(n=54) 0.31±0.06
placebo	(n=76) 0.58±0.05	(n=61) 0.55±0.06†
Social isolation		
somatropin	(n=70) 0.27±0.04	(n=54) 0.18±0.04
placebo	(n=76) 0.31±0.04	(n=61) 0.29±0.04
Sleep		
somatropin	(n=70) 0.85±0.07	(n=54) 0.70±0.08††
placebo	(n=76) 0.55±0.05	(n=61) 0.78±0.07
Pain		
somatropin	(n=70) 0.34±0.04	(n=54) 0.45±0.09‡
placebo	(n=75) 0.28±0.05	(n=60) 0.49±0.09
Mobility		
somatropin	(n=70) 0.61±0.06	(n=54) 0.58±0.07
placebo	(n=76) 0.37±0.05	(n=61) 0.54±0.06
SUM score		
somatropin	(n=69) 4.67±0.11	(n=53) 5.09±0.12#
placebo	(n=71) 4.43±0.12	(n=57) 4.88±0.13

NB on the SUM score a positive score is good

* P<0.001 compared with baseline

** P=0.016 vs somatropin group

† P<0.001 vs placebo group

†† P=0.011 compared with baseline

‡ P=0.047 compared with baseline

P=0.037 compared with baseline

GHDQ

No significant treatment effects seen between the groups in the mood, energy or sleep scales, but at 12 months sleep had improved in the somatropin group compared with baseline (P=0.011)

Social history

Satisfaction with lives at 12 months	
somatropin group	85%
Placebo group	83%
Mean days off sick to 6 months	
somatropin group	0.45±0.11
Placebo group	0.41±0.09

Methodological comments

Randomisation method: computer-generated list
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not stated
Baseline characteristics: the groups were comparable at baseline except in relation to mean estimated duration of GHD, which was longer in the placebo group
Dropouts and withdrawals: 19 patients withdrew from the original somatropin group and 11 from the original placebo group. 40% of these withdrawals were related primarily to a treatment-related adverse event. No explanation is provided in relation to the nature of the adverse events which caused these withdrawals (other than that 13 of those in the somatropin group were due to oedema or arthralgia), or to the causes of the other withdrawals.
Compliance: high rates of compliance (>90%) were seen in 70% and 73% of patients who completed 6 and 12 months of treatment respectively

General comments
Conflict of interests: assistance provided by Pharmacia
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	2
Was the study described as double blind?	2
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal (RCT phase)	Somatropin (n=83)	Placebo (n=80)
Oedema or arthralgia	13	Seemingly none
Other withdrawals (not necessarily due to adverse events)	6	11
Total withdrawals	19	11

Number of specific adverse events excluding withdrawals	RCT phase		Open phase
	Somatropin (n=83)	Placebo (n=80)	N=130
Oedema (including generalised, peripheral or facial oedema, carpal tunnel symptoms, and peripheral swelling or tightness)	48%	30%	43%
Myalgia/arthralgia (including arthritis, arthrosis, myalgia, muscle stiffness, tendonitis and muscle weakness)	30%	13%	25%
Paraesthesia/anaesthesia	12%	4%	15%
Increased sweating	3.6%	0%	No data
Aggressive reactions	0%	3.8%	No data
Moniliasis	0%	3.8%	No data
Adrenal insufficiency	5 patients	0	No data
Operation for pituitary tumour	1 patient	1 patient	No data
Collapse	1 patient, 2 events	0	No data
Amaurosis fugax and chest pain	1 patient	0	No data
Total events	290*	219*	411*
Total patients with events	70*	60*	99*

* reported numbers, impossible to relate to data given

Reference and design	Intervention	Participants	Outcome measures
Author: Carroll et al, 1997[39] Country: UK Type of study: retrospective analysis of two RCTs Length of treatment: 1 year Loss to follow-up: data not available for 4 patients from study 1 Jadad score: 2/5	Name of somatropin: Genotropin Dose: Study 1 (n=24) 0.024 mg/kg/day vs placebo; study 2 (n=18) 0.012 mg/kg/day vs placebo Did any patients receive somatropin before trial: not reported Other hormone replacements: stable replacement of all other pituitary deficiencies (including sex steroids)	Total number: 42 Isolated or multiple deficiencies: not reported Comorbidities: not reported Adult or childhood onset: men 90.7% adult onset; women 96.1% adult onset Definition of GHD: peak GH response of < 3 μg/1mU/l Cause of GHD: not reported Mean GH concentrations: not reported Sex: 21 male, 17 female Mean age: (± SE) 42.9 ± 1.9	Quality of life scales used: NHP, PGWB

*Results **(Mean ± SEM)***

Overall score	*Baseline*	*NHP after 6 months' somatropin therapy*
NHP		
somatropin group	14.9±2.3	4.9±1.1*‡
placebo group	12.4±2.6	8.9 ±1.9
PGWB		
somatropin group	74.9 ± 2.6	85.8±2.2**‡
placebo group	75.6 ± 4.1	86.1±2.6†

* P<0.01 vs baseline
** P<0.001 vs baseline
† P<0.05 vs baseline

‡ results at end of initial 6 months of treatment, whether this was during the randomised or the open label part of the study. Thus in this group comparisons with baseline should be with the whole group, not just the group originally randomised to somatropin

The psychological benefits were stated to be similar with the higher and lower doses of somatropin

Methodological comments
Randomisation method: not reported
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: not reported
Dropouts and withdrawals: not reported
Compliance:

General comments
Conflict of interests:
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal
No adverse events reported

Reference and design	Intervention	Participants	Outcome measures
Author: Davies et al, 2000[28] Country: UK Type of study: open therapeutic trial Length of treatment: 3 months Loss to follow-up: 2 patients withdrew due to adverse effects Jadad score: N/A	Name of somatropin: Humatrope, Eli Lilly Dose: 0.01 iU/kg/day increased to 0.015 iU/kg in males and 0.02 iU/kg/day in females Did any patients receive somatropin before trial: Not reported Other hormone replacements: thyroxine, hydrocortisone, sex steroids and desmopressin	Total number: 39 (an additional 24 refused treatment) Isolated or multiple deficiencies: multiple Comorbidities: one patient had overt evidence of cardiovascular disease, 7 had cured Cushing's disease and 3 had cured acromegaly. Adult or childhood onset: ? all adult Definition of GHD: peak response <10mU/l to standard provocative testing Cause of GHD: Craniopharyngioma 3; apoplexy 1; non-functional pituitary tumour 13; Cushing's 6; acromegaly 3; prolactinoma 4; Sheehan's 3; dysgerminoma 1; idiopathic 1; medulloblastoma 1; dysgerminoma 1; meningioma 1; acute lymphoblastic leukaemia Mean GH concentrations: Not reported Sex: 20 male; 19 female Mean age: 46.4 ± 14.4	Quality of life scales used: Assessment of Growth Hormone Deficient Adults (QoL-AGHDA)

Results

QoL-AGHDA

Baseline: 10.0 ± 4.0
Post treatment: 7.0 ± 4.1 p<0.001 (from table)
In text: QoL-AGHDA fell significantly after treatment (7 vs. 4, p<0.001). Five patients derived no improvement or a slight increase in QoL-AGHDA score after somatropin therapy. There was a significant correlation between the percent change in body composition and change in QoL score (r=0.34, p<0.05) though no significant correlation with percent change in cholesterol/HDL ratio (r=0.19, p>0.05).

Methodological comments
Randomisation method: not randomised
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics: Patients who did and did not somatropin therapy were compared. There were significant differences with regard to age (group receiving somatropin therapy were younger) and peak somatropin responses (somatropin group had significantly lower responses).
Dropouts and withdrawals: 2 patients withdrew due to adverse effects
Compliance:

General comments
Conflict of interests:
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	n/a
Was the study described as double blind?	n/a
Was there a description of withdrawals and dropouts?	n/a

Adverse events leading to withdrawal
One patient withdrew due to headaches
One patient withdrew due to fatigue

Number of specific adverse events
Mild swelling of fingers and toes n=2

Reference and design	Intervention	Participants	Outcome measures
Author: Degerblad et al 1990[7] Country: Sweden Type of study: placebo-controlled cross-over trial Length of treatment: 12 weeks each of somatropin and placebo, with washout period of at least 12 weeks Loss to follow-up: none Jadad score: 3/5	Name of somatropin: Somatonorm Dose: 4 IU for 6 or 7 days/week, depending on body weight, to correspond to a dose of 0.5-0.6 IU/kg^{-1} /week^{-1} Did any patients receive somatropin before trial: the 5 with childhood-onset GHD had at some time, but had ceased treatment at least 5 years prior to study entry Other hormone replacements: the 5 patients with multiple pituitary hormone deficiency received cortisone acetate, thyroxine, and testosterone or oestrogen and progesterone; one of them also received desmopressin for diabetes insipidus	Total number: 6 Isolated or multiple deficiencies: isolated 1, multiple 5 Comorbidities: diabetes insipidus - 1 Adult or childhood onset: adult 1, childhood 5 Causes of GHD: idiopathic - 3 craniopharyngioma - 1 perinatal asphyxia, congenital toxoplasmosis - 1 Cushing's disease, pituitary irradiation - 1 Definition of GHD: GH <3.4 following arginine-insulin tests Mean GH concentrations: 0.67 µg/l Mean duration of GHD: not stated Sex: 3 men, 3 women Mean age: 29 (20-38)	Quality of life scales used: Profile of Mood Scales (POMS), Sjoberg mood questionnaire

Results at baseline			Results at 12 weeks		
	somatropin	placebo		somatropin	placebo
POMS *(Mean ± SEM)*			**POMS *(Mean ± SEM)***		
Tension	2.67±0.26	2.60±0.21	Tension	2.65±0.25	2.77±0.29
Depression	2.17±0.20	2.47±0.42	Depression	1.93±0.24	2.55±0.39
Anger	1.97±0.23	2.13±0.39	Anger	2.10±0.27	2.50±0.30
Fatigue	2.77±0.24	2.97±0.34	Fatigue	2.50±0.44	2.93±0.23
Confusion	2.20±0.28	2.73±0.31	Confusion	2.40±0.35	2.58±0.30
Sjoberg *(Mean ± SEM)*			**Sjoberg *(Mean ± SEM)***		
Activity	2.78±0.23	2.70±0.20	Activity	2.97±0.28	2.60±0.18
Social orientation	3.12±0.13	2.50±0.30	Social orientation	2.92±0.18	2.95±0.19
Control	2.77±0.19	2.70±0.16	Control	2.97±0.17	2.72±0.19
Extraversion	2.70±0.12	2.67 ±0.08	Extraversion	2.82±0.11	2.57±0.07
Calmness	2.53±0.26	2.60±0.13	Calmness	2.70±0.14	2.42±0.17
Pleasantness	2.93 ±0.16	2.55±0.19	Pleasantness	2.76±0.28	2.58±0.19

Methodological comments
Randomisation method: not given
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not stated
Baseline characteristics: as this was a cross-over trial, both groups were identical
Dropouts and withdrawals: none
Compliance: no data

General comments
Conflict of interests: supported by grants from the Karolinska Institute, the Swedish Medical Research Council, Svenska Sallskapet for Medicinsk Forskning, Nordisk Insulin Foundation, Magnus Bergvall Foundation and Clas Groschinskys minnesfond
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal	None reported

Number of specific adverse events	Somatropin	Placebo
Fluid retention	1	0
Arthralgia	1	0
Total events	2	0
Total patients with events	1	0

Reference and design	Intervention	Participants	Outcome measures
Author: Deijen et al 1998[18] Country: the Netherlands Type of study: placebo-controlled dose-finding RCT Length of treatment: subjects starting on somatropin remained on that dose for 24 months; those on placebo switched to somatropin (2 IU/m^2) after 6 months and continued active treatment for 24 months 2 years Loss to follow-up: excluded for poor compliance – 2 withdrew during first 6 months – 2 incomplete patient data – 1 withdrew during 2nd year - 2 Jadad score: 1/5	Name of somatropin: Dose: 1, 2 or 3 IU/m^2. Subjects started on one third of target somatropin dose, and the dose was then increased by one-third over the next 2 months until the target dose was reached Did any patients receive somatropin before trial: not for at least a year (mean 7.5±4.5 years) Other hormone replacements: testosterone undecanoate (n=29), levo-thyroxine (n=27), hydrocortisone (n=20), vasopressin (n=6)	Total number: 50 Isolated or multiple deficiencies: isolated 17, multiple 31 Comorbidities: not stated Adult or childhood onset: childhood Causes of GHD: congenital GHD - 41 Craniopharyngioma - 7 Definition of GHD: peak GH response to 100 µg GH-releasing hormone or insulin-induced hypoglycaemia of <7 µg/l Mean GH concentrations: not stated Mean duration of GHD: not stated Sex: 50 men, no women Mean age: 26.7 (19-37)	Quality of life scales used: Hopkins Symptom Checklist (HSCL), shortened Dutch version (32 items) of Profile of Mood States (POMS), State-Trait Anxiety Inventory (STAI)

Results

Full data were not published. No treatment effects were seen at 6 months. The state anxiety score was significantly lower at month 24 in the pooled group of treated patients compared with baseline (baseline 34.7±10.4, month 24 31.9±10.8, P=0.04), but no other significant changes were seen over 24 months

Methodological comments

Randomisation method: not given
Patients blinded to treatment: not specified
Outcome assessors blinded to treatment: not specified
Baseline characteristics: no data were given to enable comparison of the groups at baseline
Dropouts and withdrawals: no data. 14 patients were said to require dose reductions because of side effects: these were dose-related.
Compliance:

General comments

Conflict of interests: none stated
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**

Was the study described as randomised?	1
Was the study described as double blind?	0
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal	None reported

Number of specific adverse events	Not stated

Reference and design	Intervention	Participants	Outcome measures
Author: Drake et al 1998[29] Country: UK Type of study: prospective uncontrolled Length of treatment: 6 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Genotropin Dose: starting dose 0.8 IU/day (0.4 IU/day in 2 patients with essential hypertension or impaired glucose tolerance). Doses were adjusted, if necessary, at 4, 8 and 12 weeks to maintain serum IGF-1 concentrations between the median and the upper end of the age-related reference range. Median dose 0.8 IU/day (range 0.4-1.6) in men, 1.2 (range 0.8-2.0) in women Did any patients receive somatropin before trial: no Other hormone replacements: no data	Total number: Isolated or multiple deficiencies: isolated 10, multiple 40 Comorbidities: cranial diabetes insipidus - 11 Adult or childhood onset: adult Causes of GHD: clinically non-functioning pituitary adenoma - 15 corticotropinoma - 12 prolactinoma - 12 Sheehan's syndrome - 1 post-cerebral irradiation for acute lymphoblastic leukaemia - 2 craniopharyngioma - 3 somatotropinoma - 1 Idiopathic - 3 tuberculous meningitis - 1 Definition of GHD: peak GH level <9 mU/l after insulin-induced hypoglycaemia or a glucagon stimulation test Mean GH concentrations: not stated Mean duration of GHD: not stated Sex: 17 men, 33 women Mean age: 45 (range 18-69)	Quality of life scales used: QoL-AGHDA

Results	
Mean QoL-AGHDA score (SD) **Baseline** 14.2 (5.9)	*Mean QoL-AGHDA score (SD)* **3 months** **6 months** 7.4 (4.5)* 7.0 (5.5) * P<0.001 compared with baseline
Methodological comments Randomisation method: not applicable Patients blinded to treatment: no Outcome assessors blinded to treatment: no Baseline characteristics: Dropouts and withdrawals: no data Compliance: no data **General comments** Conflict of interests: supported by Pharmacia & Upjohn Other: the main purpose of the study was to compare the results in patients treated do novo with a dose-titration regimen with those in patients previously treated initially in an RCT using a weight-based regimen which was then titrated during routine follow-up; quality of life scores were said to be not significantly different in the two groups, but actual data were not presented for the ex-RCT group. The time taken to reach a maintenance dose was significantly shorter in men (4 weeks, range 2-12) than in women (11 weeks, range 2-26, P<0.0001). Arthralgia resolved following a reduction in dose of 0.4 IU/day for 2 weeks, and did not recur when the original dose was restored.	

Adverse events leading to withdrawal	No data

Number of specific adverse events	
Arthralgia	8
Total events	8
Total patients with events	8

Reference and design	Intervention	Participants	Outcome measures
Author: Florkowski et al, 1998[25] Country: New Zealand Type of study: crossover RCT Length of treatment: 3 months Loss to follow-up: all patients completed the study Jadad score: 3/5	Name of somatropin: Genotropin 16 pen (Pharmacia and Upjohn) Dose: 0.125 U/kg/week for 1 month up to 0.25 U/kg/week Did any patients receive somatropin before trial: 2 had received somatropin during childhood but not in past 4 years Other hormone replacements: all except 3 were on long-term replacement with two or more pituitary hormones	Total number: 20 Isolated or multiple deficiencies: multiple Comorbidities: not reported Adult or childhood onset: 18 adult onset, 2 childhood onset Definition of GHD: GH <3 μg/l Cause of GHD: Adenoma 9; empty pit fossa 1; Cushing's 2; prolactinoma 2; head injury 1; craniopharyngioma 2; idiopathic 2; post Tb hypopituitarism 1 Mean GH concentrations: not reported Sex: 17 male, 3 female Mean age: 47 (range: 20-69 years)	Quality of life scales used: Disease Specific Questionnaire (DSQ); Symptom Checklist-90 (SCL-90); Social Adjustment Scale (SAS)

Results

SAS

Baseline: 1.8 for whole group; Baseline DSQ and SCL-90 total score values not reported.

SCL depression subscale-baseline

0.81 ± 0.18 (for those receiving somatropin first)
0.55 ± 0.12 (for those receiving somatropin second)

There was a significant decline with respect to time for both SAS and SCL-90 (p=0.03 and 0.013 respectively). This was for the whole group and not a function of active treatment. The DSQ showed a trend to decline (p=0.06) but no effect for active treatment.

SCL-depression subscale

0.52 ± 0.19 (for those receiving somatropin first)
0.28 ± 0.09 (for those receiving somatropin second)
No significant time-group interaction on ANOVA.
None of the subscale scores of the SCL-90 showed any significant change on active treatment.

Methodological comments
Randomisation method: not reported
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: n/a crossover trial
Dropouts and withdrawals: none
Compliance:

General comments
Conflict of interests:
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal
No withdrawals

Number of specific adverse events
None reported

Reference and design	Intervention	Participants	Outcome measures
Author: Gibney et al 1999[35] Country: UK & Australia Type of study: Observational (10-year follow-up of patients who took part in an RCT[11]) Length of follow-up: 10 years Loss to follow-up: 2 Jadad score: not appicable	Name of somatropin: Dose: mean dose 0.025 IU/kg/day Did any patients receive somatropin before trial: the study compares patients who had taken somatropin for at least 9 out of 10 years, and continuously for 1 year before the study with those who had taken it for less than 1 of the 10 years, and had not taken it for the year before the study Other hormone replacements: corticosteroids – 9 in each group T_4 – 9 in each group gonadal steroids – 9 in each group desmopressin – 1 in treated group, 2 in untreated group fluorocortisone –2 in treated group	Total number: Isolated or multiple deficiencies: most, if not all multiple Comorbidities: not specified Adult or childhood onset: not stated Causes of GHD: Cushing's disease – 4 in treated group, 2 in untreated group prolactinoma or chromophobe adenoma – 4 in treated group, 8 in untreated group, idiopathic – 2 in treated group Definition of GHD: peak GH level <3 mU/l after insulin tolerance test Mean GH concentrations: not stated Mean duration of GHD: not stated Sex: treated group 7 men 3 women untreated group 8 men 3 women Mean age: treated group 48 (range 31-58) untreated group 49 (range 31-61)	Quality of life scales used: Nottingham Health Profile

Results

NHP (mean ± SEM)
Baseline

	somatropin group (n=10)	Untreated group (n=11)
Energy	43.3±12.6	36.4±12.6
Emotional reaction	19.8±6.3	11.9±3.9
Social isolation	9.1±4.9	17.9±8.9
Sleep	5.2±2.7	15.0±6.6
Pain	0.0±0.0	3.5±2.4
Mobility	4.9±1.9	3.3±1.6
Overall	18.8±6.1	14.6±3.7

NHP (mean ± SEM)
10 years

	somatropin group (n=10)	Untreated group (n=11)
Energy	5.3±3.5*	36.3±8.3
Emotional reaction	3.7±2.6*	14.6±8.4
Social isolation	4.0±2.7	13.1±8.6
Sleep	14.5±6.1	29.9±11.7
Pain	4.8±4.8	7.8±3.0
Mobility	5.9±3.7	11.2±4.6
Overall	7.5±2.5*	18.8±4.5

*P<0.02 vs baseline, P<0.02 change from baseline vs change from baseline in untreated group

Methodological comments
Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics: subjects who stopped taking somatropin did not differ either at baseline or in response to somatropin in the original trial from those who continued treatment.
Reasons for discontinuing somatropin treatment included inability to get it prescribed (n=6), perceived side effects (n=2) and lack of interest (n=3)
Dropouts and withdrawals: 1 subject of the original trial was excluded as not fitting either category of this study, and 2 were lost to follow-up
Compliance: not stated

General comments
Conflict of interests: none stated
Other: GHD was measured at the outset of the original study, in 1987

Adverse events leading to withdrawal	None during this phase of the study

Number of specific adverse events	None during this phase of the study

Reference and design	Intervention	Participants	Outcome measures
Author: Giusti et al, 1998[19] Country: Italy Type of study: RCT Length of treatment: 6 months RCT, followed by 6 months when those initially receiving placebo were given somatropin therapy Loss to follow-up: 1 patient Jadad score: 1/5	Name of somatropin: Genotropin, Pharmacia Dose: 0.5-1.0 UI daily Did any patients receive somatropin before trial: not reported Other hormone replacements: adrenal replacement, desmopressin, gonadal replacement and thyroid replacement	Total number: 26 Isolated or multiple deficiencies: not reported Comorbidities: not reported Adult or childhood onset: all adult onset Definition of GHD: GH < 3.5 μg/l Cause of GHD: Adenoma 9; prolactinoma 6; Willis aneurysm 1; craniopharyngioma 4; ACTH-oma 1; empty sella 2; meningioma 3. Mean GH concentrations: 4.6 μg/kg Sex: 12 male, 14 female Mean age: 51.0 (range: 21-74 years)	Quality of life scales used: Kellner Symptom questionnaire (KSQ) Italian version; Hamilton Depression Scale (HDS)

Results

KSQ scores

Somatropin Group			Placebo Group		
Baseline	3 months	6 months	Baseline	3 months	6 months
Total score					
23.8 ± 3.5	22.5 ±1.0	19.0 ± 4.0	24.4 ± 3.3	23.6 ± 3.0	19.6 ± 3.5
Anxiety					
6.9 ± 1.2	5.9 ± 1.2	5.6 ± 1.0	5.1 ± 0.8	4.3 ± 0.8	4.5 ± 0.9
Depression					
6.0 ± 1.3	5.3 ± 1.2	4.0 ± 1.2	6.3 ± 1.0	5.3 ± 0.9	4.5 ± 1.3
Somatisation					
6.6 ± 1.2	6.1 ± 1.2	5.4 ± 1.3*	9.9 ± 1.9	10.3 ± 2.0	9.3 ± 2.3
Hostility					
4.9 ± 0.9†	5.1 ± 1.3	4.4 ± 1.2*	2.3 ± 0.6	3.6 ± 0.9	2.4 ± 0.6

† p=0.04 vs. corresponding experimental time in the placebo-treated group.
*p=0.2 vs. corresponding experimental time in the placebo-treated group.
There was no difference in overall scores on the KSQ between somatropin and placebo groups on entry but subsection analysis of items showed significantly higher scores for hostility in the somatropin group compared to the placebo group on entry.

HDS

No differences in HDS score at baseline between groups. After 6 months, there was a significant decrease in HDS score (most values not reported) in the somatropin group but not in the placebo group compared to baseline. Significances were p=0.008 at 3 months for somatropin group vs. baseline; p=0.02 [HDS score: 28 ± 1 to 25 ± 1] at 6 months somatropin group vs. baseline; p=0.09 for somatropin group vs. placebo at 3 months and p=0.2 for somatropin group vs. placebo at 6 months. Values reported in figures.

Methodological comments
Randomisation method: not reported
Patients blinded to treatment: not reported
Outcome assessors blinded to treatment: not reported
Baseline characteristics: no significant differences between two groups
Dropouts and withdrawals: 1 drop out
Compliance:

General comments
Conflict of interests: none reported
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	0
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal
None reported

Number of specific adverse events
Severe swelling n=1
Headache, swelling n=1 (not included in data analysis)

Reference and design	Intervention	Participants	Outcome measures
Author: Hayes et al 1999[30] Country: Ireland Type of study: prospective uncontrolled Length of treatment: 12 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Genotropin Dose: 0.125 IU/kg/week in daily doses for 4 weeks, 0.25 IU/kg/week thereafter Did any patients receive somatropin before trial: not stated Other hormone replacements: appropriate pituitary hormone replacement – 11 testosterone – 5 oestrogen - 5	Total number: 12 Isolated or multiple deficiencies: isolated 1, multiple 11 Comorbidities: none Adult or childhood onset: 9 adult, 3 childhood Causes of GHD: craniopharyngioma - 3 pituitary adenoma/ apoplexy - 4 idiopathic -1 other - 4 Definition of GHD: peak GH response of <10 mIU/l following insulin-induced hypoglycaemia (n=10) or L-Dopa (n=2) Mean GH concentrations: not stated Mean duration of GHD: 11.0±2.5 years Sex: 7 men, 5 women Mean age: 35.4±2.5	Quality of life scales used: Nottingham Health Profile

Results

Baseline

NHP (Mean ± SEM)

All:	119.7±32.6
Men:	92.4±4.4
Women:	160.9±45.8

12 months

NHP (Mean ± SEM)		*P value*
All:	48.9±23.9	<0.05
Men:	23.1±11.5	P<0.07
Women:	89.3±55.8	P=0.3

2 subjects with a baseline score of 0, leaving no room for improvement, were excluded from the analysis. The improvement in scores was highest in those with the highest baseline score and therefore the greatest opportunity for improvement (r=-0.75, P<0.02)

Additional unpublished data (pers comm. Prof TJ McKenna)

Baseline

NHP (Mean ± SEM)

N = 10

Emotional reaction	24.73±7.32
Sleep	20.21±9.23
Energy	44.80±13.46
Pain	5.12±4.03
Mobility	10.96±4.61
Social life	12.81±8.43
Total score	118.6±31.9

Additional unpublished data (pers comm. Prof TJ McKenna)

12 months

NHP (Mean ± SEM)

N = 10

Emotional reaction	7.23±3.26
Sleep	14.13±7.53
Energy	13.68±9.19
Pain	5.39±5.39
Mobility	8.49±2.40
Social life	3.95±3.95
Total score	50.9±23.7

Methodological comments

Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics:
Dropouts and withdrawals: no data
Compliance: no data

General comments

Conflict of interests: supported by Pharmacia
Other: the purpose of this study was to seek to determine whether baseline clinical characteristics can identify those individuals likely to derive the most benefit from somatropin replacement.

Adverse events leading to withdrawal	No data

Number of specific adverse events	
Oedema & arthralgia	3 patients

Reference and design	Intervention	Participants	Outcome measures
Author: Hernberg-Stahl et al 2001[36] Country: multinational Type of study: observational (data from KIMS database) Length of treatment: 12 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Genotropin Dose: maximum starting dose 0.125 IU/kg/week with a subsequent increment to a maximum of 0.25 IU/kg/week based on individual requirement and responsiveness; the guidelines for therapy did not preclude the use of dose titration independent of body weight, based on clinical response & serum IGF-1 measurements Did any patients receive somatropin before trial: not during adulthood Other hormone replacements: gonadal steroids – 93.9% of men, 39.3% of women	Total number: 304 Isolated or multiple deficiencies: 4.7% of men and 9.7% of women had isolated GHD Comorbidities: not specified Adult or childhood onset: 284 adult, 20 child Causes of GHD: not stated Definition of GHD: not stated Mean GH concentrations: not stated Mean duration of GHD: men 8.2±8.3 years women 10.1±9.3 years Sex: 150 men, 154 women Age: men 50.8±18.1 women: 48.6±12.6	Quality of life scales used: QoL-AGHDA, KIMS Patient Life Situation Form

Results

Baseline
Mean + SD

Men	Women	All subjects
QoL-AGHDA		
7.7±6.0	10.4±6.5*	9.0±6.4

* P<0.001 vs men

PLFS

No of days sick leave in previous 6 months		
12.0±36.5	7.2±23.6	9.5±30.4
No of hospital days in previous 6 months		
2.1±12.8	1.3±4.1	1.7±9.4
No of doctor visits in previous 6 months		
2.1±3.5	2.0±4.3	2.1±3.9
Leisure-time physical activity		
42.1±26.8	39.7±29.2	40.8±28.0
Satisfaction with physical activity		
44.7±27.2	38.7±30.1	41.6±28.8
Need for assistance with daily activities (%)		
11*	31	21

* P<0.001 vs women

12 months
Mean + SD

Men	Women	All subjects
QoL-AGHDA at 12 months		
5.0±5.6*	7.2±6.0	6.2±5.9**

* P<0.01 vs women
** P<0.001 vs baseline

PLFS

No of days sick leave in previous 6 months		
3.2±9.9	4.3±15.6	3.8±13.2*
No of hospital days in previous 6 months		
0.8±5.6	0.3±1.7	0.6±4.1*
No of doctor visits in previous 6 months		
1.2±2.5	1.4±2.0	1.4±2.3**
Leisure-time physical activity		
54.9±25.7‡	47.5±29.6	51.1±27.9†
Satisfaction with physical activity		
52.6±27.7‡	45.1±30.7	48.8±29.5†
Need for assistance with daily activities (%)		
6**	26#	16*

* P<0.05 vs baseline
** P<0.01 vs baseline
‡ P <0.05 vs women
† P<0.001 vs baseline
# P<0.001 vs men

Methodological comments
Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics:
Dropouts and withdrawals: not stated
Compliance: not stated

General comments
Conflict of interests: study supported by Pharmacia and Upjohn
Other: women had significantly lower quality of life scores than men both at baseline and after treatment, as did patients treated with radiotherapy. In patients previously treated for Cushing's disease, quality of life was lower at baseline but not after 12 months' treatment. There was no significant correlation between QoL-AGHDA scores and age at onset of pituitary disease (including childhood onset).

Adverse events leading to withdrawal	None

Number of specific adverse events	No serious adverse events

Reference and design	Intervention	Participants	Outcome measures
Author: Holmes & Shalet, 1995[40] Country: UK Type of study: analysis of data relating to patients who did and did not choose to continue somatropin therapy after end of study Length of treatment: 6 months RCT followed by 6 months when all patients received somatropin therapy Loss to follow-up: 2 patients withdrew from placebo arm Jadad score: 2/5	Name of somatropin: Genotropin Dose: 0.125 IU/kg/week for first month, increasing to 0.25g for the following 5 months (maximum dose 4 IU) Did any patients receive somatropin before trial: no somatropin in year before study Other hormone replacements: stable for at least 6 months prior to study	Total number: 65 Isolated or multiple deficiencies: not reported Comorbidities: continuing on somatropin: IGHD-11 (37%) Not continuing on somatropin: IGHD –9 (27%) p=0.60 Adult or childhood onset: Continuing on somatropin: Adult 16, child 14; not continuing on somatropin: adult 20, child 13 p=0.74 Definition of GHD: peak GH response of <10 μg/1mU/l Cause of GHD: Continuing Not on somatropin continuing on somatropin Non-functioning pituitary adenoma 8 Prolactinoma 12 Cushing's disease 3 FHS-secreting pituitary adenoma 3 Acromegaly 1 Craniopharyngioma 13 Cerebral tumour 12 Idiopathic GHD 7 Mean GH concentrations: not reported Sex: continuing on somatropin: 12 male, 18 female; not continuing on somatropin 15 male, 18 female (p=0.86) Mean age: continuing on somatropin: 28.3	Quality of life scales used: NHP; PGWB

		(20.1-55.6); not continuing on somatropin: 34.7 (21.3-59.5) (p=0.17)	

Results

NHP (median (range))

	Baseline	p value	6 months	p value
Energy				
Continuing (n=30)	43.60 (1-100)		-24.00 (-100 to 36.80) (n=30)	
Not continuing (n=33)	24.00 (0-100)	0.06	0 (-100 to 24.00) (n=26)	0.06
Emotional reaction				
Continuing	9.76 (0-69.80)		0 (-41.40 to 25.52)	
Not continuing	9.76 (0-90.69)	0.80	0 (-35.99 to 12.01)	0.44
Social isolation				
Continuing	0 (0-100)		0 (-61.50 to 20.13)	
Not continuing	0 (0-100)	0.84	0 (-41.89 to 0)	0.51
Sleep				
Continuing	0 (0-77.63)		0 (-77.63 to 21.70)	
Not continuing	0 (0-77.63)	0.49	0 (-65.06 to 77.63)	0.62
Pain				
Continuing	0 (0-63.50)		0 (-19.74 to 1.69)	
Not continuing	0 (0-28.70)	0.34	0 (-12.91 to 11.22)	0.67
Mobility				
Continuing	0 (0-66.09)		0 (-32.56 to 21.77)	
Not continuing	0 (0-31.07)	0.26	0 (-20.09 to 19.87)	0.60
Total				
Continuing	68.08 (0-439.87)		-38.17 (-208.48 to 70.4)	
Not continuing	44.33 (0-282.12)	0.26	-18.99 (-172.89 to 53.63)	0.09

PWBS (median (range))

Baseline		
Anxiety		P
Continuing with somatropin (n=17)	20 (8-25)	
Not continuing with somatropin (n=16)	22 (12-25)	0.15
Depressed mood		
Continuing with somatropin	12 (8-15)	
Not continuing with somatropin	14 (6-15)	0.08
Positive well-being		
Continuing with somatropin	11 (5-18)	
Not continuing with somatropin	15 (5-20)	0.10
Self-control		
Continuing with somatropin	13 (5-15)	
Not continuing with somatropin	14 (7-15)	0.35
General health		
Continuing with somatropin	12 (5-15)	
Not continuing with somatropin	12 (5-15)	0.77
Vitality		
Continuing with somatropin	8 (2-16)	
Not continuing with somatropin	12 (3-17)	0.06
Total score		
Continuing with somatropin	72 (43-102)	
Not continuing with somatropin	89 (46-104)	0.09

There was no significant difference in change in score on any subsection or in total score (data not given)

Methodological comments
Randomisation method: not reported
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: not reported
Dropouts and withdrawals: 2 withdrawals, reason not reported
Compliance: not reported

General comments
Conflict of interests: The study was supported by Pharmacia
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal
No adverse effects reported

Reference and design	Intervention	Participants	Outcome measures
Author: McGauley 1989[20], McGauley et al 1990[67], Salomon et al 1989[11] Country: UK Type of study: RCT Length of treatment: 6 months Loss to follow-up: 1 Jadad score: 5/5	Name of somatropin: Genotropin Dose: 0.07 iu/kg/day, reduced by 50% in patients who became hypertensive or who developed marked discomfort from fluid retention Did any patients receive somatropin before trial: 2 had been treated in childhood, but treatment had been discontinued 4 and 5 years before the study Other hormone replacements: 10 patients in the somatropin and 9 in the placebo group were receiving corticosteroids, 11 in the somatropin group and 10 in the placebo group were receiving thyroxine, 10 in the somatropin group and 8 in the placebo group were receiving gonadal steroids, 3 in the somatropin group and 2 in the placebo group were receiving desmopressin, and 1 in the somatropin group and 2 in the placebo group were receiving fluorocortisone	Total number: 24 Isolated or multiple deficiencies: not stated Comorbidities: somatropin group: hyperprolactinaemia - 2 stable ulcerative colitis - 1 Placebo group: parkinsonism + mild hypertension - 1 hyperprolactinaemia - 1 Adult or childhood onset: adult 22, childhood 2 Causes of GHD: Cushing's disease - somatropin group 6, placebo group 3 prolactinoma, chromophobe adenoma, craniopharyngioma – 6 in each group idiopathic hypopituitarism – placebo group 3 Definition of GHD: GH concentration <3mU/l during an insulin tolerance test, with a venous plasma glucose concentration of ≤2.0mmol/l and/or associated with symptoms of hypoglycaemia Mean GH concentrations: not stated Mean known duration of GHD: somatropin group: 6 years (range 1-25) Placebo group: 10	Quality of life scales used: Nottingham Health Profile (NHP), Psychological General Well-being Schedule (PGWS), GHQ

		years (range 1-21) Sex: somatropin group: 8 men, 4 women Placebo group: 8 men, 4 women Mean age: somatropin group: 39 (range 21-51) Placebo group: 38 (range 21-51)	

Results – baseline

NHP (mean ± SEM)	**All subjects**
Overall score	17.9±2.5
Energy	44.5±8.9
Emotional reaction	20.9±4.4
Social isolation	16.8±4.1

PGWS (mean ± SEM)	
Overall score	69.6±3.5
General health	10.0±0.7
Self-control	11.5±0.6
Vitality	10.2±0.9
Anxiety	15.7±1.0
Mood	11.6±0.5
Well-being	10.6±0.8

The authors stated that there was no significant difference in NHP or PGWS overall scores or in GHQ scores between the treatment and placebo groups at either baseline or 1 month

Results at 6 months

NHP (mean ± SEM)	**somatropin group** (n=11)	**placebo group** (n=12)
Overall score	2.5±1.2*	8.2±1.5
Energy	2.18±2.2**	21.8±6.7

* P<0.01 vs placebo group
** P=0.015 vs placebo group

PGWS at 6 months (mean ± SEM)
Overall score
No significant difference between groups

Mood	Baseline	6 months
somatropin group (n=11)	14.4±0.4	12.3±0.5†

† P=0.015 vs baseline

The authors attribute the lack of significant differences between the 2 groups in overall scores at 1 month and the PGWS overall score at 6 months to the placebo effect, and to dose-related adverse effects in the treatment group (clinical oedema in 6 subjects, and transient arthralgia in 5)

At 6 months, the somatropin group showed a reduction in psychological distress as measured by the GHQ.

Methodological comments
Randomisation method: randomisation code provided by Kabi-Vitrum
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not stated
Baseline characteristics: no statistically significant difference between the groups
Dropouts and withdrawals: 1 patient withdrew from the somatropin group 3 days after beginning treatment
Compliance: assessed by counting returned empty vials and expressing that number as a percentage of the total number of vials needed for the treatment period. Compliance was higher in the somatropin group (96±2%) than in the placebo group (78±6%, P<0.05).

General comments
Conflict of interests: supported in part by a grant from KabiVitrum Peptide Hormones AB and the Swiss National Foundation for Scientific Research
Other: most side effects disappeared spontaneously during the first 2 or 3 months of treatment, but dose reductions were required in 3, and one needed additional treatment with a diuretic agent

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	2
Was the study described as double blind?	2
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal	Somatropin group	Placebo group
Generalised misery & depression	1	0
Total withdrawals	1	0

Number of specific adverse events	Somatropin group	Placebo group
Fluid retention (including increase in body weight, swollen ankles, carpal tunnel compression and sensation of tightness in the hands)	6 patients	0
Hypertension	1 patient	0
Arthralgia	5 patients	0
Mild discomfort in large proximal muscle groups	3 patients	0
Encephalocoele at site of previous transphenoidal surgery in association with peripheral oedema	1 patient	0

Reference and design	Intervention	Participants	Outcome measures
Author: McKenna et al 1997(2) [9] Country: Netherlands Type of study: RCT Length of treatment: 6 months RCT, followed by 6 months with all patients receiving somatropin Loss to follow-up: not reported Jadad score: 2/5	Name of somatropin: Genotropin, Pharmacia & Upjohn Dose: 0.10 IU/kg/week for 4 weeks followed by 0.20 IU/kg/week for remainder of 6 months; maximum daily dose 3 IU Did any patients receive somatropin before trial: not reported Other hormone replacements: not reported	Total number: 30 Isolated or multiple deficiencies: not reported Comorbidities: not reported Adult or childhood onset: not reported Definition of GHD: not reported Cause of GHD: not reported Mean GH concentrations: not reported Sex: 15 men/15 women Mean age: 49	Quality of life scales used: QoL-AGHDA, NHP

Results

The full results of the McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

QoL-AGHDA

Values not reported. Said to be a statistically significant improvement from baseline to the end of trial when all patients were receiving somatropin ($P<0.01$); no data for somatropin vs. placebo

NHP

Values not reported. Said to be a statistically significant improvement in the energy dimension only ($P<0.05$).

Methodological comments

Randomisation method: not reported
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: not reported
Dropouts and withdrawals: not reported
Compliance: not reported

General comments

Conflict of interests:
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal
No adverse events reported

Reference and design	Intervention	Participants	Outcome measures
Author: McKenna et al.1997(1) [10] Country: Spain Type of study: RCT Length of treatment: 6 months RCT, plus 6 months with all patients on somatropin Loss to follow-up: not reported Jadad score: 2/5	Name of somatropin: Genotropin, Pharmacia & Upjohn Dose: 0.125 IU/kg/week for 4 weeks then 0.250 IU/kg/week for rest of 6 month period Did any patients receive somatropin before trial: not reported Other hormone replacements: not reported	Total number: 69 Isolated or multiple deficiencies: not reported Comorbidities: not reported Adult or childhood onset: not reported Definition of GHD: not reported Cause of GHD: not reported Mean GH concentrations: not reported Sex: 42 men/27 women Mean age: 37.7	Quality of life scales used: QoI-AGHDA, NHP

Results

The full results of the McKenna (1) trial are not reported here because they were submitted to NICE in confidence by Pharmacia & Upjohn.

QoL-AGHDA
Baseline: 11.1
End of trial: 6.9 (P<0.0001); no data for somatropin vs. placebo

NHP
Significant improvements for energy level, pain, emotional reactions, social isolation and physical mobility sections of NHP, values not reported

Methodological comments

Randomisation method: not reported
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: not reported
Dropouts and withdrawals: not reported
Compliance: not reported

General comments

Conflict of interests:
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal from RCT phase
No adverse events reported

Reference and design	Intervention	Participants	Outcome measures
Author: Mardh, et al.1994[8] Country: Sweden Type of study: reanalysis of data from 12 European studies of similar design (1988-1993); QoL data for 125 patients from 7 trials. Length of treatment: 6 months RCT, then 6-12 months open study Loss to follow-up: 18 Jadad score: n/a	Name of somatropin: Genotropin Dose: 0.125 IU/kg/week for first month then 0.25 IU/kg/week, maximum 4 IU/day Did any patients receive somatropin before trial: no somatropin therapy in previous 12 months Other hormone replacements: if necessary, stable for previous 6 months	Total number: 233 Isolated or multiple deficiencies: some with multiple, numbers not reported Comorbidities: not reported Adult or childhood onset: not reported Definition of GHD: peak GH response to stimulation of < 2μg/l in 84%, 2-5 μg/l in 14%, and 5-7 μg/l in 2% Cause of GHD: Pituitary adenoma 39% Idiopathic 31% Craniopharyngioma 18% Other 12% Mean GH concentrations: not reported Sex: 153 male, 80 female Mean age: somatropin group: 38 ± 13; placebo group: 35 ± 13	Quality of life scales used: NHP, PGWB
Results *Baseline data not given in full* **n=125** *Results for NHP given only as bar charts - and not for all domains. There was a significant difference between the treatment and placebo groups at 6 months only in relation to energy (p=0.02)* The overall PGWB score increased significantly in the treatment group vs placebo (p=0.03)			

Methodological comments
Randomisation method: n/a
Patients blinded to treatment: yes in individual studies
Outcome assessors blinded to treatment: not reported
Baseline characteristics: not reported
Dropouts and withdrawals: 18; 9 due to adverse events and 0 due to poor compliance or not wishing to continue
Compliance: not reported

General comments
Conflict of interests: study reported by Pharmacia
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	N/a
Was the study described as double blind?	N/a
Was there a description of withdrawals and dropouts?	N/a

Adverse events leading to withdrawal
9 patients withdrew due to adverse events, but specific events not reported

Most common types of adverse events during first 6 months of therapy somatropin-treated n=115, placebo treated n=118

Number of specific adverse effects	Somatropin group (%)	Placebo group (%)
Oedema	37.4	3.4
Arthralgia	19.1	1.7
Muscle pain	15.7	3.4
Upper respiratory tract infection	11.3	8.5
Paraesthesia	7.8	0.8
Headache	2.6	1.7
Nausea	2.6	2.5
Diarrhoea	1.7	3.4
Carpal tunnel syndrome	1.7	0.0
Dizziness	0.9	3.4

Reference and design	Intervention	Participants	Outcome measures
Author: Monson et al 2000[37] Country: multinational Type of study: observational (data from KIMS dataset) Length of treatment: 12 months Loss to follow-up: at the time of analysis, only 64 subjects had been reviewed at 6 months or later. It was not stated whether the remaining 45 subjects had withdrawn, or whether sufficient time had not elapsed for them to have received at least 6 months' therapy Jadad score: not applicable	Name of GH: Genotropin Dose: maximum starting dose 0.125 IU/kg/week with a subsequent increment to a maximum of 0.25 IU/kg/week based on individual requirement and responsiveness; the guidelines for therapy did not preclude the use of dose titration independent of body weight, based on clinical response & serum IGF-1 measurements (mean dose 1.1 U/day) Did any patients receive GH before trial: not for at least 6 months prior to study entry Other hormone replacements: of gonadotrophin-deficient patients aged >65, 22% of females and 90% of males were on sex steroid replacement, compared with 71% of females and 89% of men in the younger group	Total number: 109 aged >65, 64 of whom completed at least 6 months of GH treatment 863 aged <65, 220 of whom completed at least 6 months of GH treatment Isolated or multiple deficiencies: 90.9% of the older group and 97.1% of the younger group had multiple pituitary deficiencies Comorbidities: 8% of the older group and 4% of the younger group had diabetes mellitus (P=0.04); 33% of the older group and 14% of the younger group had treated hypertension (P<0.0001) Adult or childhood onset: adult onset Causes of GHD: pituitary adenoma – 84% of older and 59% of younger group (P<0.001) craniopharyngioma – 4% of older and 12% of younger group surgery – 2% of younger group irradiation – 1% of younger group trauma – 2% of younger group idiopathic – 4% of older and 7% of younger group other – 8% of older and 18% of younger group Definition of GHD: not stated Mean GH	Quality of life scales used: QoL-AGHDA

		concentrations: peak GH serum of <3 mg/l on dynamic testing in 97% of patients, and 3-5 mg/l in the rest Median duration of GHD: Older group 9.0 years (range 0.1-30) Younger group 6.8 years (range 0.2-24) (P=0.003) Sex: Older group 61% men Younger group 51% men Median age: Older group 68 (range 65-82) Younger group 46 (range 18-65)	

Results

QoL-AGHDA
The AGHDA score improved in all groups. The improvement was statistically significant in men and women under 65 (P<0.001) and in men over 65 (P<0.05). This result was only published in bar chart form. However, fuller data (though without P values) were provided by the author. The authors suggest that the failure of the older women to achieve significance may be due to their smaller number

Data from personal communication:

Mean ± SD	Men	Women	All
At baseline			
Age <65	7.91±6.01(n=265)	10.73±6.85 (n=280)	9.36±6.60 (n=544)
Age 65-75	7.20±6.36 (n=44)	9.62±6.81(n=29)	8.16±6.61 (n=73)
Age 75+	8.00±5.20 (n=3)-	-	8.00±5.20 (n=3)
Change in QoL-AGHDA at 6 months (n not given)			
Age <65	-1.92±5.65	-4.38±4.99	-3.20±5.44
Age 65-75	-3.36±5.62	-2.44±6.77	-3.00±5.96
Age 75+	-1.33±3.21	-	-1.33±3.21
Change in QoL-AGHDA at 12 months (n not given)			
Age <65	-2.23±5.81	-3.80±4.33	-2.94±5.24
Age 65-75	-5.29±4.64	-0.75±3.77	-3.64±4.74
Age 75+	-9.50±0.71	-	-9.50±0.71

Methodological comments
Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics: there were some statistically significant differences (other than age) between the two groups at baseline (see above)
Dropouts and withdrawals: it is not clear how many of the older group withdrew from the study. 643 of the younger group appear to have done so
Compliance: no data

General comments
Conflict of interests: involvement of Pharmacia and Upjohn
Other: the purpose of this study was to compare the longitudinal response to GH replacement in patients aged >65 years with adult-onset GHD with that in untreated younger patients, 220 of whom then completed > 6 months of GH therapy. The comparison is presumably with the younger patients who completed at least 6 months of therapy, rather than with the whole of the younger group.

Adverse events leading to withdrawal	No data

Number of specific adverse events	No data

Reference and design	Intervention	Participants	Outcome measures
Author: Murray et al 1999a[31], 1999b[68] Country: UK Type of study: prospective uncontrolled Length of treatment: 8 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Genotropin Dose: starting dose 0.8 IU/day titrated by 0.4 IU increments to normalise the IGF-1 SDS between –2.0 and +2.0 SD of the age-related normal range (mean dose at 8 months: men 1.11 U/day, women 1.35 U/day) Did any patients receive somatropin before trial: not in the 6 months prior to study entry Other hormone replacements: an unspecified number of patients received pituitary hormone replacement	Total number: 65 Isolated or multiple deficiencies: isolated 25, multiple 40 Comorbidities: not stated Adult or childhood onset: adult 45, childhood 20 Causes of GHD: hypothalamo-pituitary pathology, or treatment thereof – 36 (including 5 with acromegaly and 4 with Cushing's disease) cranial irradiation for primary brain tumour or prophylaxis in childhood-onset acute lymphoblastic leukaemia - 29 Definition of GHD: peak GH response of <9 mU/l to provocative testing Mean GH concentrations: not stated Mean duration of GHD: 8.6 years (range 0.5-29) Sex: 25 men, 40 women Mean age: 38.7 (range 17-72)	Quality of life scales used: PGWB, QoL-AGHDA

Results

Baseline PGWB *(Mean ± SD)*		3 months	8 months
Anxiety	15.0±5.6	18.0±3.8‡	17.6±4.9†
Depressed mood	9.6±3.6	11.8±3.0‡	11.5±3.2‡
General health	8.7±2.7	10.3±2.9‡	10.0±2.9†
Positive well-being	9.1±4.0	11.8±3.8‡	11.8±4.0‡
Self control	0.1±3.3	11.5±2.8†	11.4±3.2†
Vitality	7.1±4.5	11.3±3.9‡	11.5±4.1‡
Overall score	9.7±19.9	75.8±15.0*	73.7±19.5**

Subgroup analysis failed to demonstrate any significant effect on baseline PGWB score from gender, number of hormone deficiencies or pathology. However, multiple linear regression analysis found that 9% of the variation in baseline PGWB score could be explained by whether the patient had adult- or childhood-onset GHD (P=0.05), subjective quality of life being more impaired in those with adult-onset GHD.

Improvement in PGWB score relative to baseline

Baseline PGWB <60	27.1±12.8#	25.6±1.48#
Baseline PGWB >60	6.7±12.0	3.3±18.8

Gender, number of hormone deficiencies or pathology did not significantly influence the improvement in PGWB score following somatropin replacement, but the greatest improvements were seen in those patients with the lowest baseline scores and in those with childhood-onset GHD.

QoL-AGHDA (Mean ± SD) 15.3±6.0

QoL-AGHDA *(Mean ± SD)* *10.4±6.2** 9.8±6.5*

However, multiple linear regression analysis found that 20% of the variation in baseline QoL-AGHDA score could be explained by whether the patient had adult- or childhood-onset GHD (P=0.025), subjective quality of life being more impaired in those with adult-onset GHD.

Improvement in QoL-AGHDA score relative to baseline

Baseline QoL-AGHDA <15	-2.5±4.5	-4.0±5.7
Baseline QoL-AGHDA >15	-5.1±6.5	-6.1±6.1

Multiple linear regression analysis found that only baseline QoL-AGHDA score was predictive of improvement of score with therapy (P=0.004).

† P<0.05 vs baseline
‡ P<0.01 vs baseline
* P<0.001 vs baseline
** P=0.001 vs baseline
# P<0.001 vs baseline PGWB >60 group

Methodological comments
Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics: patients were selected for subjectively poor quality of life despite optimal replacement of other pituitary hormones
Dropouts and withdrawals: no data
Compliance: no data

General comments
Conflict of interests: involvement of Pharmacia & Upjohn
Other:

Adverse events leading to withdrawal	None reported

Number of specific adverse events	None reported

Reference and design	Intervention	Participants	Outcome measures
Author: Riva et al 1993[32], Sartorio et al 1995[42] Country: Italy Type of study: prospective uncontrolled Length of treatment: 6 months Loss to follow-up: not stated Jadad score: not applicable	Name of somatropin: Saizen Dose: 0.5 IU/kg/week Did any patients receive somatropin before trial: treatment discontinued at least 7 years prior to current study Other hormone replacements: no data	Total number: 8 Isolated or multiple deficiencies: isolated 3, multiple 5 Comorbidities: not stated Adult or childhood onset: childhood Causes of GHD: not stated Definition of GHD: peak GH response of <5 ng/ml to GH stimulation testing Mean GH concentrations: not stated Mean duration of GHD: not stated Sex: 8 men, no women Mean age: 29.6±1.2 (range 25-34)	Quality of life scales used: the State-Trait Anxiety Inventory (STAI) and Experiential World Inventory (EWI)

Results

STAI
Baseline scores not given

EWI *(Mean ± SD)*

Sensitivity	38.9±4.9
Time	42.9±5.4
Body	41.7±3.2
Self	38.9±5.7
Others	41.0±9.1
Thought	43.9±5.4
Dysphoria	35.5±7.4
Impulsiveness	43.9±7.5
Hyperaesthesia	44.5±6.5
Hypoaesthesia	44.5±6.6
Euphoria	43.5±10.0
Anxiety	44.1±7.1

STAI
Scores not given. Said to be a reduction in transient anxiety levels though not in anxiety as a personality trait

EWI *(Mean ± SD)*	*At 6 months*	*6 months after withdrawal of therapy*
Sensitivity	32.7±2.9*	35.9±6.7
Time	38.0±6.1	40.6±9.2
Body	37.6±4.2	36.9±6.9
Self	36.6±5.1	37.8±4.4
Others	37.6±7.2	40.6±4.3
Thought	39.4±5.9**	39.6±6.7
Dysphoria	33.9±7.6	33.8±6.8
Impulsiveness	39.6±4.3**	40.4±12.7
Hyperaesthesia	37.5±7.1	38.6±8.5
Hypoaesthesia	37.5±5.8	36.6±7.8
Euphoria	37.0±9.1	45.0±6.6
Anxiety	34.6±4.2†	37.2±7.0**

* P<0.03 vs baseline
** P<0.05 vs baseline
† P<0.02 vs baseline

Methodological comments
Randomisation method: not applicable
Patients blinded to treatment: no
Outcome assessors blinded to treatment: no
Baseline characteristics:
Dropouts and withdrawals: no data
Compliance: no data

General comments
Conflict of interests: none reported
Other:

Adverse events leading to withdrawal	None reported

Number of specific adverse events	None reported

Reference and design	Intervention	Participants	Outcome measures
Author: Noaves Soares et al, 1999[21] Country: Brazil Type of study: double blind RCT Length of treatment: 6 months RCT followed by 6 months open trial Loss to follow-up: 1 patient dropped out Jadad score: 3/5	Name of somatropin: Genotropin, Pharmacia Dose: 0.125 IU/kg/week for first month increased to 0.250 IU/kg/week for following 5 months Did any patients receive somatropin before trial: no somatropin treatment in 12 months before trial Other hormone replacements: not reported	Total number: 10 Isolated or multiple deficiencies: multiple Comorbidities: 6 of 9 had had previous depressive episodes Adult or childhood onset: 7 adult; 2 childhood Definition of GHD: 0.2-4.5 ng/ml Cause of GHD: adenoma 4; empty sella 1; prolactinoma 2; idiopathic 1, craniopharyngioma 1 Mean GH concentrations: not reported Sex: 6 male, 3 female Mean age: 39.4 (range: 28-52 years)	Quality of life scales used: Beck Depression Inventory (BDI) and Hamilton Depression Scale (HDS)

Results (Mean ± SD)

HDS

	Baseline	6 months	12 months
Somatropin group (n=5)	7.60 ± 5.81	2.20 ± 1.64*	0.6 ± 0.54
Placebo group (n-4)	4.75 ± 1.26	2.50 ± 2.64	0.5 ± 1.0 (then receiving somatropin)

No significant difference between somatropin and placebo groups at baseline.

BDI

Somatropin group (n=5)	12.60 ± 7.02	4.20 ± 1.92*	1.80 ± 0.83
Placebo group (n=4)	7.0 ± 3.16	4.5 ± 1.29	1.75 ± 1.7 (then receiving somatropin)

No significant difference between groups at baseline; *p=0.043 compared to baseline
The placebo group at 6 months did not improve after somatropin treatment following the initial placebo response.

Methodological comments
Randomisation method: not reported
Patients blinded to treatment: yes
Outcome assessors blinded to treatment: not reported
Baseline characteristics: not reported
Dropouts and withdrawals: one drop out due to non-compliance
Compliance: not reported

General comments
Conflict of interests:
Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal
No adverse events reported

Reference and design	Intervention	Participants	Outcome measures
Author: Verhelst et al 1997[22] Country: Belgium Type of study: RCT Length of treatment: 6 months RCT, followed by 18 months when all subjects, including those receiving placebo initially, were given somatropin therapy Loss to follow-up: 15 subjects withdrew during the first 6 months, 28 in months 7-12 and 14 in months 13-24. 2 stopped at baseline, 20 withdrew because of adverse effects and 35 because of insufficient subjective improvement Jadad score: 2/5	Name of somatropin: Genotropin Dose: 0.125 IU/kg/week for the first month, then 0.25 IU/kg/week (1.5 IU/m^2/day) in daily doses (max 4 IU/day). Mean dose 2.50±0.71 IU/day after 6 months Did any patients receive somatropin before trial: not in the previous 12 months Other hormone replacements: TSH – somatropin group 50, placebo group 58 ACTH – somatropin group 55, placebo group 56 LH/FSH – somatropin group 54, placebo group 60 ADH – somatropin group	Total number: 148 Isolated or multiple deficiencies: isolated 16, multiple 132 Comorbidities: none reported Adult or childhood onset: adult 134, child 14 Causes of GHD: pituitary tumour – somatropin group 47, placebo group 43 craniopharyngioma – somatropin group 6, placebo group 8 idiopathic – somatropin group, placebo group 5 trauma – somatropin group 2, placebo group 3 other – somatropin group 9, placebo group 18 Definition of GHD: peak GH response of <10 mU/l to provocative testing by either insulin-induced hypoglycaemia, glucagons or clonidine Mean GH concentrations: not reported Mean duration of GHD: not stated Sex: Somatropin group: 42 men, 29 women Placebo group: 47 men, 30 women Mean age: Somatropin group: 43.5 Placebo group: 44.1	Quality of life scales used: Nottingham Health Profile (NHP), social self-reporting questionnaire

Results *NHP* Data given in graphic form only *Social self-reporting questionnaire* Days sick leave over 6-month period Somatropin group: 12.17±3.90 Placebo group: 10.1±5.6 Hospitalisation rate Somatropin group: 14.9% Placebo group: 13.0%	*NHP* Data given in graphic form only. At 6 months, both the somatropin and the placebo group had significant improvements in energy, emotions and sleep; although the somatropin group performed slightly better than the placebo group in all these domains, the difference did not reach significance. There was a significant difference between the 2 groups in the pain domain (P=0.02), but the improvement was in the placebo group. *Social self-reporting questionnaire* 6 months 12 months 18 months 24 months Days sick leave over 6-month period somatropin group: 7.15±3.50* 2.93±1.55** 0.39±0.17† 3.3±2.51‡ Placebo group: 11.6±5.57# Hospitalisation rate Somatropin group: 7.0 % 5.8% 3.8% 7.7% Placebo group: 14.1% * P=0.009 vs baseline ** P=0.01 † P<0.001 ‡ P=0.026 # P=0.039 vs somatropin group
Methodological comments Randomisation method: not stated Patients blinded to treatment: yes Outcome assessors blinded to treatment: not stated Baseline characteristics: both groups were comparable at baseline Dropouts and withdrawals: 15 subjects withdrew from the placebo-controlled phase. However, they were not attributed to treatment groups, and the reasons for withdrawal were not specified. Compliance: no data **General comments** Conflict of interests: none stated Other:	

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	0

Adverse events leading to withdrawal from RCT phase	No data

Number of specific adverse events – RCT phase	Somatropin group (n=71)	Placebo group (n=77)
Arthralgia	15.4%	2.4%
Peripheral oedema	12.6%	1.25
Generalised oedema	5.6%	0%
Myalgia	4.2%	0%
Paraesthesia	2.8%	0%
Stiffness in extremities	2.8%	1.2%

Carpal tunnel syndrome	2.8%	0%
Depression	2.8%	1.3%
Dyspepsia	2.8%	0%
Nervousness	2.8%	1.3%
Hyperuricaemia	1.4%	1.3%
Flu	1.4%	1.3%
High blood pressure	1.4%	1.4%
Headaches	1.4%	1.3%
Tendonitis	1.4%	1.2%
Tiredness	0%	1.3%
Insomnia	0%	2.6%
Cutaneous rash	0%	2.8%

Reference and design	Intervention	Participants	Outcome measures
Author: Wallymahmed,et al. 1997[23] Country: UK Type of study: RCT Length of treatment: 6 months RCT, followed by 6 months open trial Loss to follow-up: no losses at 6 months, 2 wtihdrew at 6 months, 19 continued treatment after 12 months Jadad score: 3/5	Name of somatropin: Genotropin, Pharmacia Ltd. Dose: 0.125 u/kg/week for first month, 0.25 u/kg/week for following 5 months, in 7 daily doses. Dose reduced if side effects developed. Maximum daily dose (irrespective of body weight) = 4 u/day Did any patients receive somatropin before trial: not reported Other hormone replacements: not reported	Total number: 32 Isolated or multiple deficiencies: not reported Comorbidities: not reported Adult or childhood onset: 28 adult and 2 child (both had craniopharyngiomas) Definition of GHD: peak GH level < 10mU/l Cause of GHD: n=30 Hypothalamic pituitary disorders: Pituitary adenoma: 15 Other intracranial tumours: 7 Cranial/transsphenoidal surgery & external irradiation: 17 Surgery alone: 9 External irradiation alone: 4 Mean GH concentrations: not reported Sex: GH group 7 male, 10 female. Placebo group 3 male, 10 female Mean age: Somatropin group 37 ± 12.9; placebo group 33 ± 11.2	Quality of life scales used: NHP, HAD, Self Esteem Scale (SE), Mental Fatigue Questionnaire (MFQ), Life Fulfilment Scale (LFS), Impact scale (IS) (adapted for use in this population.

NHP (mean (SD))
Energy

	Baseline	*6 months (n=30)*	*12 months (n=30)*
Somatropin group	1.76 (1.0)	1.05 (0.9)*	0.82 (1.0)*
Placebo group	1.30 (1.1)	0.84 (1.2)	0.92 (1.3)
Emotional reaction			
Somatropin group	2.52 (2.9)	1.82 (2.7)	1.88 (2.1)
Placebo group	1.38 (1.5)	1.53 (2.1)	1.30 (1.9)
Social isolation			
Somatropin group	0.52 (0.9)	0.76 (1.2)	1.0 (1.3)
Placebo group	0.62 (1.0)	0.92 (1.4)	0.54 (0.8)
Sleep			
Somatropin group	1.35 (1.4)	1.41 (1.4)	1.41 (1.5)
Placebo group	1.15 (1.3)	1.0 (1.4)	0.84 (1.3)
Pain			
Somatropin group	1.29 (2.1)	2.0 (2.8)	1.52 (2.6)
Placebo group	0.84 (1.3)	1.15 (1.9)	0.61 (1.3)
Mobility			
Somatropin group	0.88 (1.1)	1.58 (1.5)	1.11 (1.6)
Placebo group	0.61 (1.0)	0.92 (1.2)	0.84 (1.3)

* P<0.01 compared to baseline

HADS (mean (SD)

Anxiety			
Somatropin group	7.8 (3.4)	7.3 (3.4)	7.3 (3.5)
Placebo group	6.6 (2.9)	5.5 (3.3)	5.9 (2.8)
Depression			
Somatropin group	5.5 (2.8)	5.1 (3.2)	5.7 (3.9)
Placebo group	4.6 (3.0)	4.7 (4.5)	3.1 (3.3)

Self esteem (mean (SD))

Somatropin group	27.8 (4.9)	28.5 (5.9)	28.6 (6.5)
Placebo group	28.4 (3.5)	30.9 (4.4)**	32.3 (5.7)**

** P<0.05 compared with baseline

MFQ (mean (SD))

Somatropin group	20.5 (8.9)	18.2 (8.1)	16.5 (7.3)
Placebo group	15.8 (4.3)	15.5 (6.0)	14.6 (4.7)

Life fulfilment scale (mean (SD))

Life fulfilment (personal)			
Somatropin group	34.0 (13)	38.1 (16.6)	31.2 (14.5)#
Placebo group	29.6 (9.5)	39.5 (13.4)†	27.9 (14.6)††
Life fulfilment (material)			
Somatropin group	11.1 (7.4)	17.2 (14.6)‡	12.29 (8.6)
Placebo group	9.6 (5.3)	15.8 (12.7)	9.8 (4.7)

# P<0.05, 6-12 months
† P<0.01compared with baseline
†† P<0.0001 compared with 6 months
‡ P<0.05 compared with baseline

Impact scale (mean (SD))

Somatropin group	22.1 (8.2)	19.1 (7.5)##	19.5 (8.2)
Placebo group	22.1 (5.7)	21.1 (7.3)	18.8 (8.5)

##P<0.05 compared with baseline)
Although the placebo-controlled phase ends at 6 months, 12 month scores are also given for both groups, as the study also functions as a before-and-after study for both groups. NB that at 12 months the placebo group has received 6 months' somatropin

Combined data for the 13 patients still on GH and available for qol assessment at 36 months – the only scores to show significant change

(Mean (SD))

	Baseline	24 months	36 months
NHP Energy	1.61 (1.1)	0.38 (0.5)*	0.92 (1.0)

Methodological comments
Randomisation method: not reported Patients blinded to treatment: yes Outcome assessors blinded to treatment: not reported Baseline characteristics: not reported Dropouts and withdrawals: no losses at 6 months, 2 withdrew at 6 months due to adverse effects, only 19 continued with treatment after 12 months. Compliance: measured by self reported number of injections missed and by vial count.
General comments
Conflict of interests: not reported Other:

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	1
Was the study described as double blind?	1
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal
2 withdrew but adverse effects not described

Number of specific adverse events
Oedema of the hands and feet and/or arthralgia n=11
Mild transient adverse effects not requiring reduction in dose n=7

Reference and design	Intervention	Participants	Outcome measures
Author: Whitehead et al 1992[12 13] Country: UK Type of study: placebo-controlled cross-over trial with 1-month washout period Length of treatment: 6 months Loss to follow-up: 4 (3 due to adverse events, 1 to inconvenience of daily injection) Jadad score: 3/5	Name of somatropin: Genotropin Dose: 0.5 u/kg/week in 7 daily doses Did any patients receive somatropin before trial: not stated Other hormone replacements: hydrocortisone, thyroxine, oestrogen/mixed testosterone esters	Total number: 14 Isolated or multiple deficiencies: isolated 3, multiple 11 Comorbidities: not stated Adult or childhood onset: adult-onset 6, childhood-onset 8 Causes of GHD: craniopharyngioma – 5, idiopathic hypopituitarism – 4, isolated GHD – 3, post-partum Sheehan's syndrome – 1, post-traumatic hypopituitarism - 1 Definition of GHD: peak response of <7 mU/l to insulin-induced hypoglycaemia (blood glucose <2.0 mmol/l) Mean GH concentrations: 1.5 mU/l Mean duration of GHD: not stated Sex: 9 men, 5 women Mean age $\pm$ SEM: 29.4$\pm$2.7 (range 19.5-52.0)	Quality of life scales used: Nottingham Health Profile, PGWB

Results Baseline data were not published, but have been provided by Dr Whitehead: Before somatropin NHP1 (n=10) 44.27 NHP2 (n=12) 1.3 Before placebo NHP1 (n=10) 43.82 NHP2 (n=10) 1.2	*NHP* Results were not published as the authors claimed that the effect of somatropin could not be assessed as several values were equal pre- and post-treatment so that non-parametric statistical values could not be applied. Data provided by Dr Whitehead relating to results after 6 months are as follows: After somatropin NHP1 (n=10) 33.34 NHP2 (n=12) 1.1 After placebo NHP1 (n=10) 54.73 NHP2 (n=10) 1.1 *PGWB* Psychological well-being was said to be unaltered by somatropin therapy, but figures were not published; however, median scores in all subscales did not differ from American population norms or from a small population in England
Methodological comments Randomisation method: the study is not stated to have been randomised, although presumably it was, and no details are given of the method of randomisation Patients blinded to treatment: yes Outcome assessors blinded to treatment: not reported Baseline characteristics: Dropouts and withdrawals: 29% of patients withdrew Compliance: not stated **General comments** Conflict of interests: study sponsored by Kabivitrum Ltd Other:	

Quality assessment for RCTs (Jadad Score)	
Question	**Score**
Was the study described as randomised?	0
Was the study described as double blind?	2
Was there a description of withdrawals and dropouts?	1

Adverse events leading to withdrawal	Somatropin	Placebo
Carpal tunnel compression	1	0
Fluid retention	1	0
Transient amblyopia	0	1
Total withdrawals due to adverse events	2	1

Number of specific adverse events (excluding withdrawals)	Somatropin	Placebo
Fluid retention	1	0
Painful knees & shoulder	1	0
Acute back pain	1	1
Total events	3	1
Total patients with events	2	1

Reference and design	Intervention	Participants	Outcome measures
Author: Wiren et al, 1998 (stage 2) Country: Sweden Type of study:retrospective study Length of treatment: more than 12 months Loss to follow-up: 9 patients did not complete questionnaire Jadad score: n/a	Name of somatropin: not reported Dose: somatropin at 6 µg/kg/day (0.018 IU/kg/day) for first 4 weeks, thereafter 12 µg/kg/day (0.036 IU/kg/day) (n=104), initial dose 0.17 mg/day (n=30), initial dose 0.33 mg/day (n=27) Did any patients receive somatropin before trial: all had received somatropin therapy for more than 12 months (mean duration of treatment 33 months) Other hormone replacements: glucocorticoids, thyroid hormone, gonadal steroids and/or desmopressin given as adequate and stable therapy where necessary	Total number: 161 Isolated or multiple deficiencies: not reported Comorbidities: not reported Adult or childhood onset: adult 138, child 23 Definition of GHD: <3.0 µg/l Cause of GHD: Non-secreting pituitary adenoma 78 GH- or ACTH-secreting pituitary adenoma 16 Prolactinoma 8 Craniopharyngoma 21 Idiopathic 10 Other 28 Mean somatropin concentrations: not reported Sex: 94 male, 67 female Mean age: 50.5 (range: 21-78)	Quality of life scales used: questionnaire developed specifically for this study
Results 141/152 patients had experienced positive effects of somatropin therapy. 123/152 stated that they had more energy, 90/152 that they were happier, 68/152 that they slept less, 84/152 that they were physically fitter and 34/152 that their memory had improved. 85/152 said that family or friends had noticed an improvement in their condition			
Methodological comments Randomisation method: n/a Patients blinded to treatment: n/a Outcome assessors blinded to treatment: n/a Baseline characteristics: n/a Dropouts and withdrawals: Compliance: **General comments** Conflict of interests: Other:			

Quality assessment for RCTs (Jadad Score)	
Question	**Score**

Was the study described as randomised?	N/a
Was the study described as double blind?	N/a
Was there a description of withdrawals and dropouts?	N/a

Adverse events leading to withdrawal
No adverse events reported

Appendix 3: List of excluded studies

Dean HJ, McTaggart TL, Fish DG, Friesen HG. The educational, vocational, and marital status of growth hormone-deficient adults treated with growth hormone during childhood. *AJDC* 1985;**139**:1105-10 (subjects not assessed for GHD at study entry).

Holmes SJ, McKenna SP, Doward LC, Hunt SM, Shalet SM. Development of a questionnaire to assess the quality of life of adults with growth hormone deficiency. *Endocrinology and Metabolism* 1995; **2**:1-69 (qualitative study of quality of life in GHD adults).

Hunt SM. Developing a measure of quality of life for adults with growth hormone deficiency. *Drug Information Journal* 1994; **28**:1-11 (qualitative study of quality of life in GHD adults).

Keselman A, Martinez A, Pantano L, Bergada C, Heinrich JJ. Psychosocial outcome in growth hormone deficient patients diagnosed during childhood. *Journal of Pediatric Endocrinology* 2000;**13**:409-16 (subjects not assessed for GHD at study entry).

Nicholas LM, Tancer ME, Silva SG, Underwood LE, Stabler B. Short stature, growth hormone deficiency, and social anxiety. *Psychosomatic Medicine* 1997;**59**:372-5 (subjects not assessed for GHD at study entry).

Sandberg DE, MacGillivray MH, Clopper RR, Fung C, LeRoux L, Alliger DE. Quality of life among formerly treated childhood-onset growth hormone-deficient adults: a comparison with unaffected siblings. [see comments]. *Journal of Clinical Endocrinology & Metabolism* 1998;**83**:1134-42 (subjects not assessed for GHD at study entry).

APPENDIX 4 - AGHDA

LISTED BELOW ARE SOME STATEMENTS that people may make about themselves.

Read the list carefully and put a tick in the box marked **YES** if the statement applies to you,

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

Please answer every item. If you are not sure whether to answer YES or NO, tick whichever answer you think is most true in general

	YES	NO
I have to struggle to finish jobs	☐	☐
I feel a strong need to sleep during the day	☐	☐
I often feel lonely even when I am with other people	☐	☐
I have to read things several times before they sink in	☐	☐

	YES	NO
It is difficult for me to make friends	☐	☐
It takes a lot of effort for me to do simple tasks	☐	☐
I have difficulty controlling my emotions	☐	☐
I often lose track of what I want to say	☐	☐

	YES	NO
I lack confidence	☐	☐
I have to push myself to do things	☐	☐
I often feel very tense	☐	☐

	YES	NO
I feel as if I let people down	☐	☐
I find it hard to mix with people	☐	☐
I feel worn out even when I've not done anything	☐	☐

	YES	NO
There are times when I feel very low	☐	☐
I avoid responsibility if possible	☐	☐
I avoid mixing with people I don't know well	☐	☐

	YES	NO
I feel as if I am a burden to people	☐	☐
I often forget what people have said to me	☐	☐
I find it difficult to plan ahead	☐	☐
I am easily irritated by other people	☐	☐

	YES	NO
I often feel too tired to do the things I ought to do	☐	☐
I have to force myself to do all the things that need doing	☐	☐
I often have to force myself to stay awake	☐	☐
My memory lets me down	☐	☐

Now please go back to page 1 and make sure that you have answered "YES" or "NO" to every question, on all two pages of the questionnaire. Thank you for your help.

APPENDIX 5 - SENSITIVITY ANALYSIS

Scenario	Overall	18-30			31-55			56-64			65+		
		6-10	11-15	16+	6-10	11-15	16+	6-10	11-15	16+	6-10	11-15	16+
[1a] ScHARR 'Optimistic'	£51,457	£124,941	£55,358	£40,746	£114,789	£50,884	£37,483	£94,866	£42,420	£30,971	£38,185	£27,885	£25,286
[1b] ScHARR 'RCT gains'	*	*	*	*	*	*	*	*	*	*	*	*	*
[1c] ScHARR 'No QoI benefit'	£20,090,631	£254,012,046	£193,237,345	£78,963,137	£40,085,679	£21,163,464	£18,668,317	£27,827,509	NA	£15,096,273	£1,788,694	£1,939,915	£3,569,092
[7a] Pharmacia & Upjohn 'Worst case'	£35,235	£85,425	£38,005	£28,425	£76,434	£34,593	£25,753	£56,036	£26,848	£20,256	£24,456	£18,327	£16,927
[7b] Pharmacia & Upjohn 'Intuitive'	£27,825	£54,654	£30,403	£23,927	£48,968	£27,588	£21,645	£36,556	£21,393	£16,972	£19,025	£15,108	£14,128
[8a] Eli Lilly without lost productivity	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586	£30,586
[8b] Eli Lilly with lost productivity	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705	£25,705
Sensitivity Analyses													
[2a] 1 Year	£62,011	£141,303	£62,807	£47,091	£136,178	£60,477	£45,402	£124,191	£55,224	£41,299	£48,201	£34,999	£32,251
[2b] 4 Year no discounting	£53,968	£131,272	£58,170	£42,840	£120,400	£53,377	£39,341	£99,081	£44,309	£32,369	£39,911	£29,151	£26,448
[2c] 10 Years	£44,207	£110,385	£48,809	£35,509	£99,973	£44,451	£32,256	£78,405	£35,866	£25,841	£32,056	£23,699	£21,284
[2d] 20 Years	£37,496	£96,087	£42,408	£30,081	£85,315	£38,700	£27,323	£60,767	£30,130	£21,729	£27,189	£20,657	£18,481
[2e] 20 Years no discounting	£49,862	£127,629	£56,643	£41,420	£112,273	£51,307	£37,299	£76,584	£38,464	£28,656	£33,330	£25,519	£23,411
[3] No clinical impact	£51,617	£125,006	£55,374	£40,769	£115,153	£51,024	£37,572	£95,285	£42,271	£31,067	£39,134	£28,374	£25,538
[4a] High Costs	£55,172	£133,358	£59,090	£43,504	£122,900	£54,482	£40,145	£102,406	£45,793	£33,448	£41,657	£30,425	£27,600
[4b] Low Costs	£45,337	£110,554	£48,991	£36,088	£101,260	£44,893	£33,098	£82,932	£37,087	£27,107	£32,914	£24,041	£21,828
[5] 40% Initially fail(Novo)	£57,714	£139,500	£61,799	£45,426	£128,705	£57,049	£41,978	£107,642	£48,251	£35,106	£42,940	£31,423	£28,528
[6] Productivity included	£42,003	£104,294	£46,210	£34,015	£94,275	£41,792	£30,785	£74,613	£33,354	£24,355	£28,562	£20,858	£18,899
[7c] PU correct	£36,084	£86,905	£38,772	£28,837	£78,116	£35,442	£26,231	£58,483	£28,054	£21,002	£26,258	£19,746	£18,090

Scenario	Overall	Overall across age groups		
		6-10	11-15	16+
[1a] ScHARR 'Optimistic'	£51,457	£100,750	£48,462	£36,511
[1b] ScHARR 'RCT gains'	*	*	*	*
[1c] ScHARR 'No Qol benefit'	£20,090,631	£17,868,789	£24,926,168	£18,931,277
[7a] Pharmacia & Upjohn 'Worst case'	£35,235	£68,478	£33,147	£25,237
[7b] Pharmacia & Upjohn 'Intuitive'	£27,825	£45,058	£26,515	£21,203
[8a] Eli Lilly without lost productivity	£30,586	£30,586	£30,586	£30,586
[8b] Eli Lilly with lost productivity	£25,705	£25,705	£25,705	£25,705
Sensitivity Analyses				
[2a] 1 Year	£62,011	£119,935	£57,736	£44,343
[2b] 4 Year no discounting	£53,968	£105,642	£50,822	£38,312
[2c] 10 Years	£44,207	£87,799	£42,342	£31,445
[2d] 20 Years	£37,496	£75,126	£36,883	£26,767
[2e] 20 Years no discounting	£49,862	£98,353	£48,665	£36,416
[3] No clinical impact	£51,617	£101,389	£48,580	£36,599
[4a] High Costs	£55,172	£108,035	£51,957	£39,146
[4b] Low Costs	£45,337	£88,706	£42,686	£32,198
[5] 40% Initially fail(Novo)	£57,714	£112,962	£54,386	£40,917
[6] Productivity included	£42,003	£82,147	£39,556	£29,837
[7c] PU correct	£36,084	£70,164	£34,053	£25,760

* The full conclusions on the health economics are not reported here because they were based upon the results of the Mckenna (1) trial which was submitted to NICE in confidence by Pharmacia & Upjohn.