Technology Assessment Report commissioned by the NHS R&D HTA Programme on behalf of the National Institute for Health and Clinical Excellence

Recombinant human growth hormone for the treatment of growth disorders in children: a systematic review and economic evaluation

Produced by Southampton Health Technology Assessments Centre

Authors Andrea Takeda
Keith Cooper
Alex Bird
Louise Baxter
Elena Gospodarevskaya
Geoff Frampton
Karen Welch
Jackie Bryant

Correspondence to

Dr Andrea Takeda
Southampton Health Technology Assessments Centre
University of Southampton
First Floor, Epsilon House
Enterprise Road
Southampton Science Park
Southampton
SO16 7NS

Date completed 28th August 2009

Source of funding: This report was commissioned by the NIHR HTA Programme as project number HTA 08/51/01.

Declared competing interests of the authors

None.

Acknowledgements

We would like to thank members of our advisory group who provided expert advice and comments on the protocol and/or a draft of this report: Dr Justin Davies, Consultant in Paediatric Endocrinology, Southampton University Hospital Trust; Dr Justin Warner, Consultant in Paediatric Endocrinology and Diabetes, University Hospital of Wales; Ms Pauline Musson, Paediatric Endocrine Nurse Specialist, Southampton University Hospital Trust and two anonymous clinical experts (one Paediatric Endocrinologist and one Adult Endocrinologist). We are also grateful to David Turner, WIHRD (Health Economist) for work on the protocol, Liz Hodson, Library Assistant, WIHRD, University of Southampton and Jill Colquitt, Senior Research Fellow, SHTAC, University of Southampton for retrieving references and reviewing a draft of this report, respectively.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

Contributions of authors

Development of protocol	Andrea Takeda, Jackie Bryant, Louise Baxter, Keith Cooper
Epidemiology / background	Andrea Takeda, Louise Baxter, Jackie Bryant, Geoff Frampton
Literature searching	Karen Welch
Inclusion screening	Louise Baxter, Andrea Takeda
Data extraction / critical appraisal	Louise Baxter, Andrea Takeda
Health economics	Keith Cooper, Alex Bird, Elena Gospodarevskaya
Drafting of report	Andrea Takeda, Louise Baxter, Keith Cooper, Alex Bird, Jackie Bryant, Geoff Frampton, Elena Gospodarevskaya
Project Co-ordinator	Andrea Takeda

TABLE OF CONTENTS

Appendices

Tables

Figures

DEFINITION OF TERMS AND LIST OF ABBREVIATIONS

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context, but a glossary is provided for the non-specialist reader.

Table 1 Abbreviations used in this report

AE	Adverse events
AH	Adult height
AO-GHD	Adult onset growth hormone deficiency
BA	Bone age – a measure of skeletal maturity evaluated on the basis of the relative positions of the bones generally in the left hand and wrist
BMC	Bone mineral content
BMI	Body mass index (kg/m^2)
BNF	British National Formulary
BSA	Body surface area
BSPED	British Society for Paediatric Endocrinology and Diabetes
CA	Chronological age
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Database of Systematic Reviews
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CGHAC	Canadian growth hormone advisory committee
CI	Confidence interval
CKD	Chronic kidney disease
CO-GHD	Childhood onset growth hormone deficiency
CRF	Chronic renal failure
CRI	Chronic renal insufficiency
CUA	Cost utility analysis
DARE	Database of Abstracts of Reviews of Effectiveness
DEC	Development and Evaluation Committee
DEXA	Dual x-ray absorptiometry
DNA	Deoxyribonucleic acid
EQ-5D	Euro-Qol quality of life measure
EMEA	European Medicines Agency

ERF	Established renal failure
ESRF	End stage renal failure
EUROCAT	European Surveillance of Congenital Abnormalities
FDA	Food and Drug Administration
FGR	Foetal growth restriction
FH	Final height
FM	Fat mass
FT4	Free thyroxine
GFR	Glomerular filtration rate
GH	Growth hormone
GHD	Growth hormone deficiency
GV	Growth velocity (generally cm/yr)
GVSDS	Growth velocity standard deviation score – growth velocity relative to distribution of growth in children of the same chronological age (or bone age if specified)
HDL-C	High density lipoprotein cholesterol
HRG	Healthcare Resource Group
HRQoL	Health related quality of life
HTA	Health technology assessment
HtSDS	Height standard deviation score – height relative to distribution of height in children of the same chronological age (or bone age if specified)
HV	Height velocity
HVSDS	Height velocity standard deviation score
ICER	Incremental cost effectiveness ratio
IGF	Insulin-like growth factor
IGFBP	Insulin-like growth factor building proteins
IQR	Interquartile range
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ISS	Idiopathic short stature
ITT	Intention to treat
IU	International Unit (3 IU = 1 mg)
IUGR	Intrauterine growth restriction/retardation
K/DOQI	Kidney Disease Outcomes Quality Initiative
KIGS	Kabi International Growth Database (now Pfizer)
KIMS	Kabi International Metabolic Database (now Pfizer)
LBM	Lean body mass

LDL	Low density lipoprotein
LWS	Léri-Weill syndrome
m^2	square meters (in this context referring to body surface area)
mg	milligram
met-GH	methionyl growth hormone
MPHD	Multiple pituitary hormone deficiency
MS	Manufacturer's submission
MTA	Multiple technology appraisal
NFH	Near final height – height measured when growth is assumed to be near completion
NHS CRD	National Health Service Centre for Reviews and Dissemination
NHS EED	National Health Service Economic Evaluation Database
NICE	National Institute for Health and Clinical Excellence
NKF	National Kidney Foundation
nr	not reported
ns	not statistically significant
OLS	Ordinary least squares
PAH	Predicted adult height – Extrapolating adult height from childhood height
PSS	Personal social services
PWS	Prader-Willi syndrome
QALY	Quality-adjusted life year
QoL	Quality of life
QoL-AGHDA	Quality of life assessment of growth hormone deficiency in adults QoL-AGHDA$_{UTILITY}$ is the utility-weighted score
RCT	Randomised controlled trial
rhGH	Recombinant human growth hormone
SAR-SR	Social Adjustment Scale-self rating
SCI	Subcutaneous injection
SD	Standard deviation
SDS	Standard deviation score
SF-36	Short form 36 questionnaire
SG	Standard gamble
SGA	Short for gestational age
SHOX	Short stature homeobox-containing gene
SHOX-D	SHOX deficiency
SHTAC	Southampton Health Technology Assessments Centres

SMR	Standardised mortality rate
TS	Turner syndrome
TTO	Time trade off
U	Unit
WtSDS	WtSDS
wk	week
yr	year

EXECUTIVE SUMMARY

Background

Recombinant human growth hormone (rhGH) is licensed for short stature associated with growth hormone deficiency (GHD), Turner syndrome (TS), Prader-Willi syndrome (PWS), chronic renal insufficiency (CRI), short stature homeobox-containing gene deficiency (SHOX-D) and being born small for gestational age (SGA). NICE guidance currently recommends rhGH treatment for children with GHD, TS, PWS or CRI, but does not cover SGA or SHOX-D.

Objectives

The aim of this report was to assess the clinical- and cost-effectiveness of rhGH for children with GHD, TS, PWS, CRI, SHOX-D and those born SGA. The report extends the previous review by actively searching for studies which report growth outcomes, body composition, biochemical markers or quality of life (QoL).

Methods

Data sources

The systematic review of clinical effectiveness used *a priori* methods described in the research protocol. We searched key databases (e.g. Medline, Embase, NHS Economic Evaluation Database and 8 others) for relevant studies from their inception to June 2009, limiting to the English language. Relevant conferences, bibliographies of included papers, our expert advisory group and manufacturers' submissions to NICE were also consulted to identify any additional published or unpublished references. We developed an economic model using the best available evidence to determine cost effectiveness in the UK.

Study selection

Two reviewers assessed titles and abstracts of studies identified by the search strategy, obtained the full text of relevant papers and screened them against the inclusion criteria defined in the research protocol. Any differences in opinion throughout the process were resolved through discussion.

Data extraction and quality assessment

Data from included studies were extracted by one reviewer and checked by a second. The quality of included studies was assessed using standard criteria. Criteria were applied by one reviewer and checked by a second, with differences in opinion resolved by discussion and involvement of a third reviewer where necessary.

Data synthesis

Clinical-effectiveness studies were synthesised through a narrative review with tabulation of results of included studies. Meta-analysis was not appropriate due to heterogeneity in study design and participants.

Economic model

A decision analytic model was developed to estimate the cost effectiveness of rhGH treatment compared with no treatment for a cohort of children with GHD, TS, PWS, SGA, CRI, and SHOX-D. The perspective of the analysis was that of the NHS and PSS. The model was informed by a systematic search of the literature to identify parameters on the natural history and epidemiology of the indicated conditions, health related QoL, and costs. The model estimated the lifetime costs and benefits of rhGH with discount rates of 3.5%. The intervention effect in terms of improvement of HtSDS was derived from the systematic review of effectiveness. The outcome of the economic evaluation is reported as cost per quality adjusted life year (QALY) gained and cost per cm gained.

Results

Number and quality of studies

Of the 674 references identified, 560 were excluded on inspection of their titles and abstracts. The full papers of 114 references were retrieved, of which 28 RCTs in 34 publications were included in the systematic review of clinical effectiveness. Overall, the studies were generally poorly reported and some were of short duration.

Summary of benefits and risks

None of the studies reported QoL measures, and reporting of adverse events was limited.

GHD (1 RCT)

Children in the rhGH group grew 2.7cm/yr faster than children in the untreated group and had a statistically significantly higher height SDS (HtSDS) after one year: -2.3 ± 0.45 vs. -2.8 ± 0.45.

TS (6 RCTs)

Girls in one study grew an average of 9.3cm more than untreated girls. In a study of younger children, the difference was 7.6 cm after two years. HtSDS values were statistically significantly higher in treated than in untreated girls.

PWS (8 RCTs)

Infants who received rhGH for a year grew significantly taller (6.2 cm more) than those in the untreated group in the only study to report change in height. Two studies reported a statistically

significant difference in HtSDS in favour of rhGH. RhGH-treated patients had statistically significantly higher lean body mass and lower body fat than untreated patients in three studies. Effects on BMI were mixed.

CRI (6 RCTs)

RhGH-treated children in a one-year study grew an average of 3.6 cm more than untreated children. Height SDS was statistically significantly higher in treated than in untreated children in two studies.

SGA (6 RCTs)

No RCTs met the original inclusion criteria for the review, so these were amended to include children from the age of 3 with no catch-up growth, with no reference to mid parental height. Only one of the six included RCTs used the licensed dose; the others used doses two or three times higher. Adult height was approximately 4cm higher in rhGH-treated people in the only study to report this outcome ($p<0.005$). Adult height gain SDS was also statistically significantly higher in this study's rhGH group. Mean HtSDS was higher in treated than untreated patients in four other studies, significantly so in two of these.

SHOX-D (1 RCT)

After two years of treatment, children were approximately 6cm taller than the control group and HtSDS was statistically significantly higher in treated than in untreated patients.

Summary of cost-effectiveness

The systematic review of published economic evaluations identified two North American studies for children with TS and GHD and no studies conducted in the UK. The results of the two identified studies produced two very different estimates of cost effectiveness, largely due to the choice of utility estimates and assumptions on the effectiveness.

The systematic review of QoL identified only six studies, mostly of poor methodological quality and for small numbers of individuals. One reasonable study was found for GHD. An additional study was found which estimated QoL utilities in the general adult population according to height, using the Health Survey for England. These studies suggested that there is likely to be a small gain in utility from rhGH.

Six of the seven manufacturers submitted evidence to be considered for this review. Five out of the six manufacturers collaborated and submitted essentially the same electronic model. The model developed was based upon the previous HTA report but was extended to consider longer term

outcomes in order to estimate cost effectiveness in terms of QALYs. In the manufacturers' base case, the cost effectiveness results for all conditions were less than £30,000 per QALY gained.

From the model we developed for this review, the incremental cost per QALY estimates of rhGH compared to no treatment were: £25,483 for GHD, £43,405 for TS, £148,860 for PWS, £43,214 for CRI, £36,392 for SGA and £44,596 for SHOX-D. A further analysis was run for PWS which included a lifelong improvement of body composition of 1.8 BMI and an associated additional utility of 0.031. Under these assumptions, the cost effectiveness of PWS reduced to £60,753 per QALY gained.

The effects of a range of parameter values for the economic model were evaluated in sensitivity analyses. The model results were found to be most sensitive to the discount rate used. When the previous NICE discount rate of 6% for costs and 1.5% for benefits was used, all conditions were cost effective for a willingness to pay threshold of £30,000 per QALY. The model results are also sensitive to treatment start age and length, compliance and utility gain. The probability sensitivity analysis estimated the probability of each of the conditions to be cost effective at £30,000 to be: 88% for GHD, 12% for TS, 0% for PWS, 11% for CRI, 28% for SGA and 15% for SHOX-D.

Discussion

The systematic review was restricted to RCTs because these provide the highest level of evidence for clinical effectiveness. However, very few of these reported either final height or QoL as outcome measures, most were only one or two years in length, and some had very few participants. We did not identify any RCTs which met the original inclusion criteria for children born SGA, so these had to be amended. Only one of the included trials used the licensed dose, so results from the other five could over-state the effectiveness of rhGH treatment for this patient group.

The QoL gains were highest for individuals with lower starting heights; for those with starting height of less than <-2 HtSDS the QoL gain was minimal. For example those with PWS had a starting height of -2 HtSDS, and so for this group of patients the health gain (in terms of height) is small and therefore rhGH has high ICER values compared to no treatment. PWS patients may experience an improvement in body composition due to rhGH treatment, and this is often the point of treatment rather than gain in height, but this was difficult to quantify, especially in the long term.

The current analysis has not considered other benefits in addition to height gain within the model, apart from as a scenario analysis for PWS. The base case does not include possible benefits from changes in body composition such as reduced risk of diabetes or cardiovascular disease, which may result in increases in life expectancy. At this stage, these health gains would be purely speculative due

to lack of data, and it is not possible to quantify them. It is also possible that there may be additional psychological benefits such as improved self esteem.

Conclusions

The included studies reported statistically significantly larger HtSDS values for rhGH-treated children than untreated children with GHD, TS, PWS, CRI, SGA and SHOX-D. RhGH-treated children with PWS also showed statistically significant improvements in body composition measures compared with controls.

The cost effectiveness estimates from our model vary between conditions. Only GHD would be considered cost effective according to a willingness to pay threshold of £20,000 to £30,000 per QALY gained. TS, CRI, SGA and SHOX-D have ICERs between £35,000 to £45,000 per QALY gained. PWS has an ICER of between £60,000 and £150,000 per QALY gained depending on assumptions.

Key research priorities

- Longer studies beyond two years reporting near-final height or final adult height.
- A standardised QoL assessment specifically designed for children and adults, to be used in future RCTs and QoL studies.
- Good quality trials of GH in children born SGA, where the children included and the dose administered match the licensing criteria.
- Good quality studies of the long term effects of rhGH on body composition, psychological benefits, long term morbidities such as diabetes or cardiovascular disease, and life expectancy, particularly for individuals with PWS.

Word count: 1685

BACKGROUND

1.1 Description of health problem

Sections 1.1.1 to 1.1.5 below describe the health problem individually for the different conditions covered in this review, in terms of their aetiology and epidemiology. Sections 1.1.7 to 1.5 are general sections covering the impact of the health problems and measurement of disease for all the conditions combined.

1.1.1 Growth hormone deficiency

Growth hormone deficiency (GHD) occurs when the pituitary gland fails to produce sufficient levels of growth hormone.

There is some debate about the diagnostic criteria for GHD: the diagnosis of GHD includes short stature, growth velocity below the 25th percentile for at least one year, and delayed bone age.[1] Rosenfeld suggests other criteria of height >3 SD below the mean, < -2 SD to -3 SD for age and deceleration in growth (such as growth velocity< 25th percentile for age), GV < 5th percentile where there is no other explanation, a predisposing condition along with growth deceleration or other signs of pituitary dysfunction.[2] Juul and colleagues found 'large heterogeneity in the current practice of diagnosis and treatment of childhood GHD'. Their survey of European paediatricians found that the cut off points of GH peak response used for diagnosis of deficiency clustered around 10ng/ml or 20mU/l.[3]

The primary goals of rhGH treatment for children with GHD are: to normalise height during childhood, for the treated child to reach a 'normal' adult height as defined by the parental target and for mature somatic development to be reached around age 25.[4] The British Society for Paediatric Endocrinology and Diabetes (BSPED) recommends three or six monthly growth monitoring, annual IGF-1/ IGFBP3 monitoring and compliance assessment at each appointment.[4]

Aetiology, pathology and prognosis

GHD can be caused by a variety of factors, but in many cases the cause is unknown. In some children, failure or reduction in growth hormone secretion is congenital, and may be accompanied by other pituitary hormone deficiencies. In others, growth hormone deficiency is acquired as a result of: trauma, either at birth or later in childhood; histiocytic infiltration (build up of tissue cells); lymphoma or leukaemia; tumours involving the pituitary gland or hypothalamus; or following radiotherapy.[5] Untreated patients have a final height of 134-146 cm in males and 128-134 cm in females.[1]

Incidence and Prevalence

The UK Child Growth Foundation estimates that growth hormone deficiency of unknown origin occurs in about one in every 3800 births,[6] but reliable figures are difficult to obtain for GHD associated with radiotherapy and other causes. Figures from a study in Belgium indicate an overall prevalence of GHD of 1 in 5600.[7] The origin of GHD was stated to be unknown in 41% of the patients in the Belgian study, congenital in 20%, and acquired in 35%.[7] While the authors of this latter study state that these yearly numbers have remained similar across the sixteen years of the study, these were not collected as part of a formal screening study, and as a result the study authors believe that this figure is an underestimation. [7]

A Danish study calculated incidence rates of childhood-onset GHD, based on 1823 patients incident during 1980-1999. The average incidences per 100,000 population were calculated to be 2.58 (95% CI 2.3-2.88) for males, and 1.70 (95% CI 1.48-196) for females. The differences between the sexes was statistically significant ($p<0.001$).[8] Other sources suggest that the disorder is two to three times more common in boys than in girls.[6] A hereditary factor may be identified in some children; about 3% of children with GHD also have an affected sibling.[6]

1.1.2 Turner syndrome

Aetiology, pathology and prognosis

Turner syndrome (TS) is caused by the complete or partial absence of the second sex chromosome in girls, with or without cell line mosaicism (the presence of two populations of cells with different genotypes in one individual) leading to the presence of characteristic physical features including, but not limited to, short stature.[9,10] Other features of TS can include skeletal abnormalities, higher risk of scoliosis, cardiovascular abnormalities, lymphoedema and higher rates of hearing problems and ear malformations.[10]

While short stature is the most common clinical feature of TS,[10] in the majority of girls with TS, the missing or abnormal second chromosome causes ovarian failure, leading to lack of pubertal progression and sexual maturation. TS girls therefore receive oestrogen replacement therapy as part of their treatment.

Untreated, the average adult height deficit in women with TS is 20cm, with the average height being 143cm, (4’ 8”).[11] Cases of reduced stature are thought to be predominantly due to haploinsufficiency of the SHOX gene.[12] Not all girls with TS will require rhGH treatment and the condition does not

necessarily involve a deficiency in natural growth hormone secretion, although there may be a relative lack of sensitivity to GH, and in some cases diminished secretion.[5,13]

Incidence and Prevalence

The European Surveillance of Congenital Abnormalities (EUROCAT) reported in 2003 that TS occurred in 2.08 per 10,000 births in the UK in 2002, [14] which equates to approximately one in 2500 live-born females.[10] A Belgian study analysed age at diagnosis of 242 TS girls who were treated with rhGH between 1991 and 2002.[15] The median age at diagnosis was 6.6 (range 0-18.3) years. Although the survey found that 22% of girls were diagnosed after the age of 12 years, there was a general increase in earlier diagnosis in infancy and childhood compared with a previous survey.

A study in Denmark identified a standardised mortality rate (SMR) of 2.89 in their TS population, which was increased compared with the general population.[16] However, this significantly decreased over the three years of the study. It is unclear if this is due to a real decrease in mortality, better care of individuals with TS, or an increase in karyotypes with lower mortality.[16]

1.1.3 Prader-Willi syndrome

Prader-Willi syndrome (PWS) is a genetic disorder characterised by short stature, abnormal body composition, hypogonadism, obesity, dysmorphic features, hyperphagia (compulsive over-eating), hypotonia (diminished muscle tone) and specific learning and behavioural issues.[17]

Aetiology, pathology and prognosis

The genetic basis of the syndrome is a deletion on the long arm of the paternally derived chromosome 15 (15q11-q13) which is found in approximately 70% of affected individuals.[18] Other abnormalities have been identified including maternal uniparental disomy (2 maternal copies of chromosome 15 and no paternal chromosome 15), imprinting mutations and translocations. Abnormalities to chromosome 15 lead to disruption of the hypothalamus which controls appetite. The combination of impaired growth, abnormal body composition and hypothalamic dysfunction (hyperphagia, hypogonadism) is suggestive of growth hormone deficiency.

Birth length and weight are normal or just below normal in PWS, but growth is slow due to poor feeding. The child is noticeably short from around the first year of life and remains short throughout childhood (mean HtSDS -2) despite normal growth rate.[19] Hypotonia at birth improves towards the end of the first year of life and developmental milestones are achieved although delayed. By 2 or 3 years of age the hyperphagic phase of the condition begins, and unless eating is controlled the child will become obese.[17]

Behavioural features include food seeking, temper tantrums, obsessive compulsive disorders, high pain threshold, sleep disturbances, and skin picking. Learning disabilities are always present to some degree.[19] Hypogonadism causes delayed but complete puberty in females, although menses are infrequent or absent. Males have cryptorchidism (undescended testis) at birth and usually require androgen replacement therapy from mid-puberty even after successful orchidopexy. [17]

During adolescence the growth rate declines as a result of the absence of pubertal growth spurt. Reported mean final heights in the UK are 155cm (-3.2 SD) for males and 147cm (-2.8 SDS) for females.[20] Body composition shows increased fat mass and reduced fat free mass resulting in a high fat to lean body mass ratio even in children with normal weight to height ratios. In addition bone mineral density is reduced. The reduced bone density is multifactorial; in older patients this is due to sex steroid deficiency (hypogonadism), whereas in younger patients this is due to hypotonia, which responds to rhGH therapy.[21]

The prognosis of the condition in adulthood can be reasonable if the person can find occupation and can live in an environment where access to food can be controlled. However, many adults with the disorder develop morbid obesity, often accompanied by type II diabetes, resulting in premature death from cardiorespiratory failure.[17]

Incidence and prevalence

One UK study estimated a birth incidence for PWS of 1:20,000, with a lower bound of 1:29000.[22] The study gave a population prevalence of 1:52,000, considered the lower bound, with county rates varying from 1:42,000 to 1:67,000.[22] The overall death rate for the PWS population aged 3.4 to 56 years was found to be around 3% in one UK study compared with the standard death rate of about 0.3% each year for people in England and Wales up to the age of 55 years.[22]

1.1.4 Chronic renal insufficiency

Chronic renal insufficiency (CRI) is defined as a persistent elevation of serum creatinine and/or urea level. It can be caused by a variety of conditions, including congenital disorders, glomerular disorders and infections. Growth failure associated with CRI can be caused by acidosis, rickets, GH resistance, inadequate nutrition and anorexia.[23] Children with CRI experience impaired growth once their glomerular filtration rate (GFR) falls to 50% of normal, with increasing problems once the GFR falls below 25%.[24] Following kidney transplantation, chronic graft rejection and treatment with steroids can restrict growth and development.[25] Patients undergoing haemodialysis or peritoneal dialysis can be considered for rhGH treatment, as well as those who have received kidney transplantations.

Aetiology, pathology and prognosis

CRI is characterised by a GFR of less than 75ml/min per 1.73 m^2 body surface area (BSA).[26] The term chronic kidney disease (CKD) is also sometimes used,[26] following guidelines developed by the National Kidney Foundation (NKF) Kidney Disease Outcomes Quality Initiative (K/DOQI).[27]

The aetiology of growth failure in children with CRI includes abnormalities in the growth hormone (GH)-insulin-like growth factor (IGF)-I axis, together with nutritional and metabolic problems.[26] Nutritional supplementation in malnourished children with CRI can improve growth.[28-30] The NKF K/DOQI guidelines recommend that patients' existing nutritional deficiencies and metabolic abnormalities should be corrected before considering treatment with rhGH.[31] However, it is estimated that growth remains suboptimal even with energy intake above 80% of the recommended daily allowance.[32]

Not all patients with CRI will be shorter than average, but figures from the UK Renal Registry indicate that 29% of transplant patients and 41% of dialysis patients are below the second percentile for height.[33] Children with congenital disorders (approximately 60% of children with CRI)[25] are usually of normal length at birth, but are below the 3rd percentile for height within their first year and remain parallel to normal percentiles throughout childhood.[25] A cohort study of CRI patients who grew up before rhGH treatment was available reported that more than two-thirds remained shorter than the average population.[34] One study reported a mean height from birth to age ten which was -2.37 SD ± 1.6 below the mean. [25] Similarly, final height is reported to be reduced to below the third percentile in patients who developed end-stage renal failure in childhood.[25] Adult final height was more than two SD below the mean for approximately 60% of boys and 41% of girls who started renal replacement therapy before they were 15 years old.[35]

Incidence and prevalence

It is difficult to find accurate figures for CRI, and these do not appear to be available nationally. The UK Renal Registry reports an incidence of established renal failure (ERF) of 8.0 per million population under the age of 15 years.[36] However, established renal failure is more severe than CRI so can only really serve as a guide to the minimum number of patients for whom rhGH might be appropriate.

The UK Renal Registry reported that in 2005 there were 748 patients under the age of 18 years who were on renal replacement therapy in the UK's 13 paediatric renal centres,[33] corresponding to a prevalence of 47.7 per million.[36] However, the number of patients with CRI will be higher than this, as not all will require renal replacement therapy. ERF is reported to be more common in males than in

females (ratio 1.54:1), due to the prevalence of males with renal dysplasia and obstructive uropathy causing ERF.[36]

1.1.5 Small for gestational age

There are various thresholds for defining a child as being born 'small for gestational age' (SGA), the most commonly used being where the birth height or weight is ≤ 2 standard deviations (SD) below the population average, or is below the tenth centile for birthweight.[37] However, this group is heterogeneous in composition. Between 50 and 70% of these babies are 'constitutionally small' but otherwise healthy. The other babies in the group are those who have not reached their height or weight potential, having possibly experienced foetal growth restriction (FGR).[37] For this reason, the terms intrauterine growth retardation (IUGR) and SGA are not synonymous: a child born SGA has not necessarily undergone IUGR or FGR, and a child who has IUGR or FGR may not necessarily be born SGA.

Aetiology, pathology and prognosis

There are several possible causes for children being born SGA. These include maternal factors, such as age, ethnicity, weight, height, parity, medical conditions, smoking, malnutrition, and alcohol abuse; placental factors, and foetal factors such as chromosomal abnormalities and genetic defects.[38] Children classified as SGA may have concurrent diagnoses, such as familial short stature, TS, GHD or skeletal dysplasia.[38]

More than 80% of babies born SGA will achieve catch-up growth (growth velocity greater than the median for chronologic age and gender[38]) during their first six months,[39] with catch-up growth completed within two years for most SGA infants.[40,41] However, babies born prematurely who are SGA may take around four years to achieve catch-up growth.[42] Around 50% of the children who do not experience catch-up growth at this stage will go on to achieve their target height. It has been estimated that approximately 10% of SGA children remain at a height below -2SD throughout their childhood.[43,44] Children who are born SGA with low birth weight and who do not achieve catch-up growth by the age of two years face a relative risk of short stature (< -2SD) at age eighteen of 5.2.

Incidence and Prevalence

A study of US births estimated an annual incidence of 91,000 infants born SGA, using a definition of SGA as -2SD, or equivalent to the 2.3 percentile.[38] A Swedish study of full-term births in 1973, 1974 and 1975 found that 5.4% of neonates were SGA, defined as being <-2SD for birth length and/or height.[45] However, other studies have cited an incidence of around 3% of babies being born SGA.[46,47]

1.1.6 SHOX-D

Aetiology, pathology and prognosis

The SHOX gene is located on the distal ends of the X and Y chromosomes. This gene plays a significant role in long bone growth, and normal growth requires two functional copies.[48,49] Growth impairment can result from having a haploinsufficiency of SHOX, or from mutations.[48] Clinical features associated with SHOX-D include disproportionate shortening of the middle sections of the limbs (mesomelia), bowing of the forearms and lower legs, cubitus valgus (increased carrying angle of elbow) and Madelung deformity of the wrist.[48] However, not all people with SHOX-D will have these physical characteristics. Langer syndrome is a rare homozygous (or compound heterozygous) form of SHOX-D. It is characterised by extreme dwarfism, profound mesomelia and severe limb deformity.[48,50,51]

Incidence and prevalence

SHOX-D could be the underlying cause of restricted height in some children whose short stature cannot be explained by an underlying pathology. Estimates of the prevalence of SHOX haploinsufficiency in children with short stature of unknown origin range from 1% to 12.5%. [12,52-58] Rappold and colleagues studied 900 short children and found SHOX mutations in 2.4% of the patients with short stature of unknown origin, implying a prevalence of at least 1 in 2000 children.[55] Binder and colleagues reported a lower prevalence of SHOX haploinsufficiency, estimating it to be 1:4000.[56]

SHOX-D also causes short stature in people with concurrent diagnoses. Huber and colleagues reported that 68% of 56 children with dyschondrosteosis (a rare form of dwarfism) had SHOX anomalies.[58] Other screening studies have reported it as the cause of short stature in approximately 70% of patients with Léri-Weill syndrome (LWS).[59] Girls with TS have only one copy of the SHOX gene, and this haploinsufficiency causes short stature in some girls and women with the condition.[48]

A small study[60] which compared 26 SHOX-haploinsufficient people with 45 of their relatives and general population standards found that the SHOX haploinsufficient cohort was 2.14 SDS (3.8 cm) shorter at birth and 2.1 SDS shorter throughout childhood. Females were more severely affected than males, with women's final height being 2.4 SDS (14.4 cm) shorter than unaffected siblings, and men's final height being 0.8 SDS (5.3 cm) shorter. SHOX haploinsufficiency led to short stature in 54% of the cohort, short arms in 92% and Madelung deformity in 73%. It is not clear whether the SHOX haploinsufficient cohort in this study had concurrent diagnoses.[60]

1.1.7 Impact of health problem

Severe short stature may be physically debilitating in untreated children,[61] with children being at greater risk of bullying at school and social isolation.[62] Some children with short stature may also have difficulties with emotionally immature behaviour, anxiety and poor school performance.[63] However, not all children who are shorter than their peers will experience problems. For example, the Royal College of Obstetricians and Gynaecologists state that the majority of children born SGA do not have any appreciable morbidity or mortality.[37] However, others indicate that children born SGA who remain short may suffer from alienation, low self-esteem, impaired social dynamics, behavioural problems, lower educational achievement and professional success.[38,42]

Children with short stature can also be at increased risk of morbidity and mortality in later life. For example, the risk of cardiovascular morbidity is increased in patients with GHD,[64] TS,[65] and PWS,[66] whilst some patients with growth disorders may also be at increased risk of type 2 diabetes and metabolic syndrome.[66,67] Low birth weight is also associated with future increased risk of coronary heart rate and stroke.[68]

1.1.8 Outcome measures

The main parameter used to measure the efficacy of rhGH treatment is growth. This reflects the main goals of therapy, which are: physiological catch-up growth if possible; achievement of normal height during childhood; timely and normal growth during puberty; and normal height in adulthood. In children with PWS, treatment with rhGH aims to improve body composition as well as boosting growth.

Measures of growth include:

- Final height (FH) or adult height (AH), measured either in cm or expressed as a SDS is the best measure of how rhGH treatment affects growth. Final height has been achieved when the growth rate has slowed to less than some specified amount (e.g. 1-2cm/year) and radiographs of the wrist and hand show that the epiphyses have closed (often expressed as a bone age of more than 14-15 years).[5] Ideally, FH would be calculated in comparison with an untreated control group in an RCT. Some non-RCT designs use historical controls, which may overestimate the effects of rhGH treatment. Similarly, database studies may not include all relevant factors or be representative samples of treated patients.[5]
- Near final height (NFH) is sometimes reported where it is assumed that final height has been reached using the above criteria, but it is acknowledged that growth may not yet be quite complete.[5]

- Height, usually measured standing, using a wall-mounted Harpenden stadiometer or a similar device. For very young children, supine length is measured.
- Height standard deviation score (HtSDS). This expresses height relative to norms for children of the same age, allowing comparisons independent of age or gender. The normal population mean is zero and a normal SD score will lie between –2 and +2 SD. Increased SDS implies catch-up growth and a decrease implies growth failure. Calculation of SDS depends on the reference data used, i.e. normal height for children in the same country.
- Growth velocity (GV), also referred to as height velocity, is the change in height over a specified period, e.g. cm/year. Although the overall effectiveness of rhGH in treating short stature is to be found in measures of final height, velocity may be a better interim growth measure than height attained at a particular age as it is independent of growth in previous years.
- Growth velocity standard deviation score (GVSDS). This is the growth velocity relative to norms for children of the same age.
- Bone age. A measure of skeletal maturity, usually determined by examining the relative positions of the bones in the left hand and wrist from a radiograph. The measurement of bone age relative to chronological age is important in height prediction models. In addition, bone age assessments are used to evaluate when the epiphyses have closed and growth is complete. The interim assessment of bone age is important in determining whether treatment is advancing bone maturity such that short-term growth velocity might come at the expense of early closure of the epiphyses. Clinical trials often measure bone age to monitor whether this is accelerating undesirably fast in rhGH treated patients compared with control patients. Height for bone age can also be used as an estimate of improved height potential in response to rhGH therapy, especially in short term studies.

Measures of body composition assess obesity and the amount of fat relative to other body tissues. Body mass index (BMI) calculates the ratio of body mass to the square of body height, expressed as kg/m^2. NICE recommends BMI as providing a practical estimate of overweight in children, although mentions that it needs to be interpreted with caution as it is not a direct measure of adiposity.[69] Dual x-ray absorptiometry (DEXA) can be used to measure lean mass (fat-free mass) and percentage body fat, which can be used to indicate body composition.

Physiological outcomes reported in studies of rhGH may include assessments of the concentrations of hormones, glucose, cholesterol, and markers of bone and general metabolism. Such measures are important for assessing the biochemical, metabolic and adverse effects of rhGH, and can have implications for long-term health. Insulin-like growth factor-1 (IGF-1) is an endocrine hormone produced by the liver, and its production is stimulated by growth hormone. Lower than normal levels

are therefore seen in people with growth disorders. The insulin-like growth binding proteins (IGFBP) act as carrier proteins for IGF-1. There are six IGFBP binding proteins, with IGFBP-3 being the most abundant.[70] IGF-I is monitored during rhGH therapy as there is a theoretical concern that persistently elevated levels may predispose the patient to other diseases later in life. Monitoring levels also helps to tailor the dose to the individual. Since IGFBP-3 binds IGF-I, monitoring this gives an indication of the levels of "free" IGF-I in circulation. High levels of IGF-I with low levels of IGFBP3 may be linked with breast, colorectal and prostate cancer.[71,72]

1.2 Current service provision

Management of rhGH therapy

Children who receive rhGH therapy require regular review by paediatric endocrinologists. Older children and adolescents in need of continued rhGH therapy may enter transitional care arrangements that involve consultations with both paediatric and adult growth specialists.[73] A system of shared care is sometimes employed for rhGH therapy in the UK,[1] with diagnosis and assessment of growth being carried out in hospital outpatient consultations and some GPs writing prescriptions and possibly monitoring adverse events (AE). In other areas, all care including prescriptions and monitoring of compliance and side effects takes place in secondary care.

Administration of rhGH is usually done at home by the patient or a family member, after training, by subcutaneous injection using either needled or needle-free devices, usually pharmaceutical companies' devices rather than syringe and needle. Termination of rhGH therapy is indicated if there is a poor response (<50% increment in GV within the first year) or when final height is achieved. In children with CRI, therapy with rhGH is stopped at the time of a transplant. Therapy would not resume until at least 1 year post-transplant, and is dependent upon the absence of catch-up growth.[1]

Relevant guidance

Current guidance from NICE on the use of rhGH in England and Wales for children with growth failure due to GHD, TS, PWS or CRI was published in 2002.[74] This is discussed further in Section 1.4. Since 2002, a range of guidance on the use of rhGH in children with short stature has been published by various national health agencies and clinical expert groups for GHD, TS, CRI, PWS and SGA, but guidance for children with SHOX-D is lacking.

Guidelines on the use of rhGH for the treatment of girls and women with TS (published in 2007, relevant to US practice) recommended that treatment with rhGH should be considered as soon as growth failure has been identified and its potential risks and benefits have been discussed with the

family. It also provided rhGH dosing information and a comprehensive set of recommendations for the diagnosis, evaluation, monitoring and ongoing care of children with TS.[10]

Summary guidelines[75] and detailed recommendations[26] on the use of rhGH for short stature in children with CRI (published in 2005-2006, relevant to US practice) recommended that therapy should not commence unless patients exhibit clearly defined CRI and attain appropriate phosphorus and parathyroid hormone status.[75] The detailed recommendations included rhGH dosing information and a treatment algorithm outlining appropriate steps to improve growth and overall health outcomes.[26]

Consensus statements on using rhGH therapy in children and adults born SGA (published in 2003[38] and 2007,[76] relevant to European and US practice) emphasized the need for accurate diagnosis of SGA and recommended that rhGH therapy should be considered in children who are SGA and older than 2 years of age. However, this reflects differences in licensing in Europe and America. The FDA authorisation is for children age 2 years and over with no catch up growth (no criteria specified), and no specified HtSDS at start of treatment or reference to mid-parental height.[77] By contrast, the EMEA authorisation is for children aged 4 years and over, with a HtSDS of -2.5 at start of treatment, with a GV <0 SDS and HtSDS > 1SD below mid-parental height.[78] In addition, the licensed dose is 70mcg/kg/day in the USA and 35mcg/kg/day in Europe.

For UK populations, guidelines on rhGH therapy for children with GHD, TS, CRI, PWS and SGA was published in 2006 by the BSPED.[1] This guidance provided recommendations for shared care between GPs and specialists, together with dosing information and treatment entry and exit criteria.

1.3 Description of technology under assessment

Somatropin **(**rhGH) has been available since 1985, following the withdrawal of cadaveric human pituitary GH due to possible transmission of Creutzfeldt-Jakob disease.[5] rhGH is a synthetic form of human growth hormone produced by recombinant DNA technology, having a sequence identical to that of pituitary-derived human growth hormone. Licensed dosages vary for the different indications (Table 2), depending on whether the treatment is aiming to replace growth hormone to normal levels (for children with growth hormone deficiency), or being used in supraphysiological doses where there is no hormone deficiency but some lack of sensitivity to the hormone. It is given as a subcutaneous injection, usually at night (to mimic the child's natural fluctuations in growth hormone).[5]

Seven pharmaceutical companies have UK marketing authorisations for various indications, as shown in Table 2.

Table 2 Indications for the use of rhGH in children

Indication	Dose*	Licensed drugs (manufacturers)
Growth hormone deficiency	23-39 mcg/kg daily, or 0.7-1.0 mg/m^2 daily	Humatrope (Eli Lilly & Co. Ltd) Zomacton (Ferring Pharmaceuticals UK) NutropinAq (Ipsen Ltd) Norditropin Simple Xx (Novo Nordisk Ltd) Genotropin (Pfizer Ltd) Omnitrope (Sandoz Ltd) Saizen (Merck Serono)
Turner syndrome	45-50 mcg /kg daily, or 1.4 mg/m^2 daily	Humatrope (Eli Lilly & Co. Ltd) Zomacton (Ferring Pharmaceuticals UK) NutropinAq (Ipsen Ltd) Norditropin Simple Xx (Novo Nordisk Ltd) Genotropin (Pfizer Ltd) Omnitrope (Sandoz Ltd) Saizen (Merck Serono)
Prader-Willi syndrome, with growth velocity > 1cm/year (in combination with energy-restricted diet)	35 mcg /kg daily, or 1.0 mg/m^2 daily; max 2.7 mg daily.	Genotropin (Pfizer Ltd) Omnitrope (Sandoz Ltd)
Chronic renal insufficiency in children	45-50 mcg /kg daily, or 1.4 mg/m^2 daily	Humatrope (Eli Lilly & Co. Ltd) NutropinAq (Ipsen Ltd) Norditropin Simple Xx (Novo Nordisk Ltd) Genotropin (Pfizer Ltd) Omnitrope (Sandoz Ltd) Saizen (Merck Serono)
SHOX-D	45-50 mcg /kg daily	Humatrope (Eli Lilly & Co. Ltd)
Growth disturbance (current HtSDS -2.5 and parental adjusted HtSDS, -1) in short children born SGA, with a birth weight and /or length below -2SD, who failed to show catch up growth (HV SDS<0 during the last year) by 4 years of age or later	35 mcg/kg daily, or 1.0 mg/m^2 daily	Humatrope (Eli Lilly & Co. Ltd) Norditropin Simple Xx (Novo Nordisk Ltd) Genotropin (Pfizer Ltd) Omnitrope (Sandoz Ltd) Saizen (Merck Serono)

*Dosing information from the Electronic Medicines Compendium (http://emc.medicines.org.uk/), accessed 30 April 2008.

Adverse events have been reported in patients using rhGH. For example, sleep apnoea and sudden death among PWS patients who have one or more of the following risk factors: severe obesity, history of upper airway obstruction or sleep apnoea, or untreated respiratory infection.[79,80] There are potential risks of acromegaly, hyperglycaemia and glucosuria if the recommended dosage is exceeded.[80]Patients receiving rhGH should be monitored for glucose intolerance, as the drug may induce a state of insulin resistance.[80] It is also recommended that thyroid function should be monitored.[80] Possible side effects mentioned for 1-10% of patients include: hypersensitivity to solvent, hypothyroidism, injection site pain (reaction), and oedema.[80] Treatment should be discontinued in the event of intracranial hypertension,[80] although it may be possible to restart

treatment at a lower dose for patients who develop benign intracranial hypertension. Treatment with rhGH leads to increasing sensitivity to GH, expressed as an increase in serum IGF-I.[80]

Omnitrope, marketed by Sandoz, is a biosimilar product. This means that it is an active substance that is similar, but not identical, to the other drugs considered in this review. The issue of growth hormone therapy and biosimilars in clinical practice was the subject of a recent Parliamentary Summit.[29] The current review assesses the clinical and cost effectiveness of rhGH, without reference to the brand product or manufacturer. Discussion of the comparative safety and efficacy of biosimilars compared with reference products is therefore beyond the scope of this review.

1.4 Place of the intervention in the treatment pathway

The place of rhGH in the treatment pathway depends on the child's particular condition or syndrome, and age at diagnosis. Appropriate timing of treatment with growth hormone will depend on the underlying pathology. RhGH therapy is contraindicated in cases of progressive tumour activity and should not be used for growth promotion in children with closed epiphyses.

1.4.1 GHD

Treatment with rhGH is currently recommended by NICE to help increase the growth of children with GHD.[74] For children with congenital GHD, rhGH therapy is not generally started before the child is four years old.[5] However, if there is profound growth failure or evidence of recurrent hypoglycaemia, which may occur in infants under the age of one, treatment may be started earlier. For children who acquire GHD at an older age, treatment can start at a time appropriate to their condition and stage of growth. Treatment is discontinued after the first year if there is a poor response, i.e. <50% increase in growth rate, or if compliance or growth rate remains poor thereafter. Otherwise treatment can continue until GV is < 2cm/year, assessed over 6-12 months, when final height is achieved. Other clinical advice suggests that treatment is necessary for the patient to attain peak bone mass, which may not be until the age of 25 or 26 in some people. A recent survey of paediatric endocrinologists (56 responses out of 72 questionnaires) found that 56% of clinics provide transfer clinics for patients ending paediatric treatment and transferring to the care of an adult endocrinologist. Of the 56 respondents, 80% retest for GHD prior to transfer, 55% transfer all rhGH treated patients and the remainder transfer only those who are still GHD on retesting.[73]

Transition Phase

The transition phase in GHD is defined as the period from near final height, usually around the mid to late teens, until about twenty five years of age, or when final adult height has been reached. At the

stage of near final height, it is important to re-evaluate whether the patient is still growth hormone deficient, and if they need to continue with treatment and monitoring. Some cases, such as isolated GHD with a genetically identified mutation or multiple pituitary hormone deficiency (MPHD), severe GHD due to genetic causes, pituitary abnormalities, congenital hypopituitarism, or acquired GHD from tumours or cranial irradiation, are likely to require a continuation of therapy. However, cases of unknown origin and isolated cases of GHD carry a lower likelihood of requiring continuing treatment.[4] The BSPED consensus document suggests testing IGF-1 levels: if these are lower than -2 SD then these patients require GH stimulation re-tests. A peak GHD level of <5µg/ l during the transition phase is indicative of severe GHD.[81]

During the transition phase the authors of the consensus paper recommend that monitoring of patients should include weight and BMI at least six monthly, IGF-1, QoL, waist circumference and fasting glucose annually and body composition and total and LDL cholesterol every two to five years.[81]

1.4.2 TS

Current NICE guidance recommends that rhGH treatment for girls with TS should begin at the earliest age possible, to boost growth.[74] Some patients with profound growth retardation and failure to thrive may commence treatment earlier than those who are diagnosed later. A Belgian study [15] found that median age at diagnosis of 242 girls was 6.6 (range 0-18.3) years, although the survey found that 22% of girls were diagnosed after the age of 12 years. Some clinical expert advice suggests that the mean age for starting treatment is 8-9 years of age as many girls are not diagnosed until later in childhood, although there has been a recent trend towards earlier diagnosis.

1.4.3 PWS

NICE guidance currently recommends the use of rhGH for children with PWS to improve height, body composition and bone mineral density. For children with PWS, treatment with rhGH is intended to improve body composition and metabolism as well as increase final height. Its place in the treatment pathway depends on age at diagnosis. Children with PWS are assessed for obesity, potential for obstructive sleep apnoea and ongoing respiratory illness before treatment is considered. Low muscle tone and its impact on the child's development are also considered.

1.4.4 CRI

Treatment with rhGH is currently recommended by NICE to help increase the growth of prepubertal children with CRI.[74] The guidance recommends that treatment should be stopped after a renal transplantation, and only re-established after one year if it has been ascertained that catch-up growth has not occurred.[74] The place of rhGH in the treatment pathway for children with CRI depends on age

at diagnosis, and on clinical factors related to management of the child's condition. rhGH treatment can take place either before or after renal transplant, although allograft rejection can be a concern if rhGH treatment is given post-transplant.

1.4.5 SGA

Previous NICE guidelines did not consider children born SGA, as rhGH was not licensed for this indication at the time.[82] Children born SGA but with no comorbidities may not be diagnosed until they fail to achieve catch-up height by the age of two to four years,[38] or when they start school. The International SGA Advisory Board indicated that SGA children aged two to four years who show no evidence of catch-up with a height of -2.5 SD should be eligible for growth-hormone treatment. They also recommended that treatment should be considered in children older than four years who show no catch up at a height -2 SD or less.[38] The European license for rhGH is for children aged 4 years and over.

1.4.6 SHOX-D

Currently, there is no NICE guidance available for the use of rhGH in children with SHOX-D. Initiation of rhGH treatment for children with SHOX-D depends on age at diagnosis. Clinical evaluation is used to assess growth failure, but GH provocation tests are not required once SHOX-D has been established via a positive SHOX DNA blood test.

1.5 Current usage in the NHS

According to a survey of endocrine clinics published in 2006 by the BSPED,[73] 4758 patients have been receiving rhGH in the UK, of which 4168 were in England and Wales. Responses to the survey gave a breakdown of rhGH use by diagnosis for 3951 of the 4758 patients, indicating that 57.4% of the patients on rhGH were treated for GHD, 18.7% for TS, 4.6% for PWS, 5.2% for SGA, 2.5% for CRI, and 11.6% for other diagnoses. If we assume that these 3951 patients are a representative sample of the total population of rhGH treated patients in the UK, the total numbers of rhGH treated patients with each diagnosis would be around 2731 with GHD, 890 with TS, 219 with PWS, 247 with SGA, 119 with CRI and 552 with other diagnoses. It is possible that the number of children with CRI who received rhGH in this survey was underestimated, as some patients with CRI are managed in nephrology, rather than paediatric endocrine clinics.[73] The number of patients treated with rhGH for SHOX-D was not reported in the survey and published figures are not available. Expert advice indicates that very few SHOX-deficient patients are currently receiving rhGH, for example only two of between 350 and 400 patients in one unit receiving rhGH are being treated for this. The level of service provision for SHOX deficient patients would be similar to that required for a patient with TS.

Anticipated costs associated with intervention

The costs associated with rhGH therapy interventions comprise those of:

- The drug (dose adjusted for body weight);
- Self-therapy training of the patients and their parents (involving home visits by specialist and community nurses); and
- Monitoring of treatment effectiveness (involving paediatric endocrinology outpatient visits for blood tests, a test of pituitary function, and an assessment of bone age by hand x-ray).

The costs of training patients and their parents are limited to the first year of treatment. During each year of treatment, until they stop growing, patients would typically attend two outpatient consultations. Estimates of the current costs of these components of the rhGH interventions for patients with GHD, TS, PWS, CRI and SGA are provided in Section 4.1.6.

2 DEFINITION OF THE DECISION PROBLEM

2.1 Decision problem

Recombinant human growth hormone (rhGH) is currently recommended by NICE[74] for children with a proven clinical diagnosis of GHD, TS or PWS and for pre-pubertal children with CRI. Since the last review, rhGH has received marketing authorisation for the treatment of children born SGA and for children with growth failure associated with SHOX-D. The scope of the current project is broader than that for the previous systematic review[5] in that it covers body composition as an outcome measure for all disease areas, and also includes biochemical and metabolic markers. In addition, evidence for the use of rhGH for children born SGA, or with SHOX-D (conditions not considered in the original review) are included in this report. For these reasons, the current systematic review was undertaken as a complete review not an update. The aim of this health technology assessment is to assess the clinical effectiveness and cost effectiveness of rhGH for children with GHD, TS, PWS, CRI, SHOX-D and those born SGA.

Interventions

The intervention is recombinant human growth hormone (rhGH) also known as somatropin. It is marketed as the following products: Humatrope (Eli Lilly & Co.); Zomacton (Ferring Pharmaceuticals); NutropinAq (Ipsen); Norditropin SimpleXx (Novo Nordisk); Genotropin (Pfizer); Omnitrope (Sandoz) and Saizen (Merck Serono).

Population including sub-groups

The population is children with one of the following conditions: GHD; TS; PWS; CRI; SHOX-D; being born SGA. No age-specific definition of a child was given during the scoping process for this review. Possible subgroups could be children with different causes of GHD, and children with CRI who are either pre- or post-transplant. However, analysis of the effectiveness of rhGH treatment for any of these subgroups of patients is limited by the available data and the statistical power of the identified trials.

Transition of care from paediatric to adult endocrine services of young people requires patients to have repeat testing of their growth hormone axis to be sure that they need to continue treatment. This transition period is only considered within this review where evidence from the identified studies allows, for patients whose linear growth is not complete.

Relevant comparators

The standard comparator for this review is management strategies without rhGH. This includes placebo injections and no treatment.

Outcomes

Clinical outcomes of interest include: final height gained; height standard deviation score; growth velocity; growth velocity standard deviation score; body composition; biochemical/metabolic markers; adverse effects of treatment; health-related QoL. Direct costs include estimates of all health care resources consumed in the provision of the intervention, including diagnostic tests, administration and monitoring costs – as well as consequences of those interventions, such as treatment of adverse effects.

2.2 Overall aims and objectives of assessment

The aim of this report is to assess the clinical effectiveness and cost-effectiveness of rhGH treatment for children with GHD, TS, PWS, CRI, SHOX-D and those born SGA.

The objectives are to:

- summarise the evidence of clinical effectiveness and cost-effectiveness of rhGH when compared with no treatment;
- develop, where appropriate, an economic model adapting an existing cost-effectiveness model[5] or constructing a new model using best available evidence to determine cost effectiveness in the UK
- identify priorities for future research.

3 ASSESSMENT OF CLINICAL EFFECTIVENESS

3.1 Methodology

The methods for the systematic review of clinical effectiveness were described *a priori* in the research protocol (Appendix 1), which was sent to experts for comment. We received helpful comments relating to the general content of the research protocol, but there were none that identified specific problems with the methods of the review. The methods are summarised below.

3.1.1 Search strategy

An experienced information specialist developed and tested search strategies for this review. Separate searches were carried out to identify studies reporting clinical-effectiveness, cost-effectiveness, health-related QoL, resource use and costs, and epidemiology/ natural history of the conditions. The search strategy for Medline, shown in Appendix 2, was adapted as appropriate for a number of other electronic databases. We searched: The Cochrane Database of Systematic Reviews (CDSR); The Cochrane Central Register of Controlled Trials; NHS CRD (University of York) Database of Abstracts of Reviews of Effectiveness (DARE) and the NHS Economic Evaluation Database (NHS EED); Medline (Ovid); Embase (Ovid); National Research Register; Current Controlled Trials; ISI Proceedings; Web of Science; and BIOSIS. For all disease areas we searched the databases from their inception to June 2009. This meant there was some duplication of earlier work for the previous review, but was necessary since the present review required searches for additional outcomes, such as biochemical and metabolic markers. Searches were limited to the English language.

Relevant conferences (European Society for Paediatric Endocrinology, The Endocrine Society, American Association of Endocrinologists, Paediatric Academic Societies) were searched for recent abstracts (up to June 2009) to assess against the inclusion criteria. Bibliographies of related papers were screened for relevant studies, and we contacted experts to identify any additional published or unpublished references. We also assessed the manufacturers' submissions to NICE for any additional studies which met the inclusion criteria.

3.1.2 Inclusion and data extraction process

Titles and abstracts of studies identified by the search strategy were assessed for potential eligibility by two reviewers. The full text of relevant papers was then obtained, and inclusion criteria were applied by two independent reviewers. At both stages of the screening process, any differences in opinion on inclusion of a particular study were resolved through discussion. Data from included

studies were extracted by one reviewer using a standard data extraction form and checked by a second reviewer. Any discrepancies were identified and resolved through discussion.

3.1.3 Quality assessment

The quality of included studies was assessed using NHS CRD (University of York) criteria.[83] Quality criteria were applied by one reviewer and checked by a second reviewer, with differences in opinion resolved by discussion and involvement of a third reviewer where necessary. The criteria used are shown in Appendix 3.

3.1.4 Inclusion criteria

Patients

The inclusion criteria required the patient group to be children with growth disturbance due to one of the following conditions:

- insufficient secretion of growth hormone (growth hormone deficiency)
- Turner syndrome
- Prader-Willi syndrome, confirmed by genetic testing
- chronic renal insufficiency (prepubertal children only)
- SHOX-D
- small for gestational age (see below).

The licensed indication[80] for SGA is for growth disturbance (current HtSDS -2.5 and parental adjusted HtSDS, -1) in short children born small for gestational age, with a birth weight and/or length below -2 SD, who failed to show catch-up growth (HV SDS <0 during the last year) by 4 years of age or later. However, the review group could not find any RCTs whose inclusion criteria matched these criteria exactly. Following discussions with NICE, the team amended the criteria to be: "growth disturbance (current HtSDS <-2.5, ***but with no reference to parental height***) in short children born small for gestational age with a birth weight and/or length below -2SD, who failed to show catch-up growth (***with no particular criteria specified***) by ***3*** years of age or later."

Studies which included adolescents and young adults who have completed linear growth were excluded from the systematic review of effectiveness.

Interventions

Recombinant human growth hormone (somatropin)

Comparators

Management strategies without somatropin.

Outcomes

The following outcomes were included in the review, where data were available:

- final height gained
- height standard deviation score (height relative to the distribution of height in children of the same chronological age)
- growth velocity
- growth velocity standard deviation score (growth velocity relative to the distribution of growth in children of the same chronological age or bone age)
- body composition
- biochemical and metabolic markers
- adverse effects of treatment
- health-related QoL

Types of studies

- Fully published randomised controlled trials were included in the review, and systematic reviews of RCTs were included as sources of information. Indicators of a systematic review include: explicit search strategy, inclusion criteria, data extraction and assessment of quality.
- Studies published only as abstracts or conference presentations were included in the primary analysis of clinical and cost-effectiveness if sufficient details were presented to allow an appraisal of the methodology and assessment of results.
- Non-English language studies were excluded.
- In an effort to capture all randomised evidence, all identified RCTs were included with no restriction on length of treatment, size of study population, or design (parallel group or cross-over design). Cross-over studies could potentially be problematic as children's growth continues without treatment, making comparisons between the different arms less straightforward than in a parallel-group trial. However, we have attempted to include discussion of this in the quality assessment of studies.

3.1.5 Data synthesis

- Clinical-effectiveness studies were synthesised through a narrative review with tabulation of results of included studies. Key outcome measures are reported in tables in the text, and other outcomes are shown in the full data extraction forms in Appendix 4. For conciseness, where a

study reported outcome measures after one and two years, only the final year's outcomes are included in the table since these show the longest duration of treatment effect.

- Where data were of sufficient quality and homogeneity, a meta-analysis of the clinical-effectiveness studies was considered using Review Manager 5.0 software.
- Quality of life studies were synthesised using the same methods as above, i.e. narrative review and meta-analysis only if feasible.

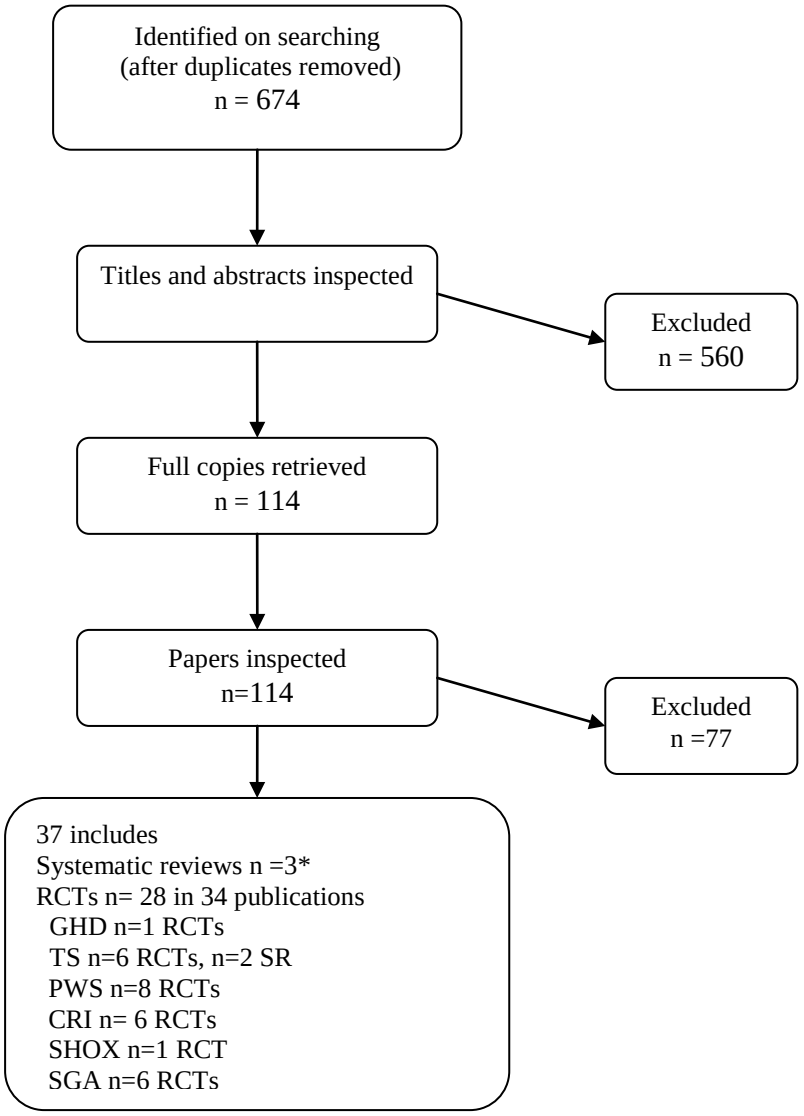
3.2 Results

A brief overview of the results of the searches is presented below. Owing to the extensive nature of this MTA, the clinical effectiveness results for the six different disease areas are presented separately in Sections 3.3 to 3.8.

3.2.1 Quantity and quality of research available

The number of references considered at each stage of the review is shown in Figure 1. Of the 674 references identified, 560 were excluded on inspection of their titles and abstracts. The full papers of 114 references were retrieved and assessed against the inclusion criteria. 77 of the retrieved full papers were rejected at this stage, mostly due to the patient group not meeting the inclusion criteria (n=40) or due to a non-RCT study design (n=27). A list of papers excluded at this stage is included in Appendix 5, together with reasons for exclusion. A total of 28 RCTs in 34 publications were included in the systematic review of clinical effectiveness. Appendix 6 lists conference abstracts which were identified as being of interest, but which contained insufficient information to be included in the review of clinical effectiveness.

Figure 1 Flowchart of identification of published studies for inclusion in the systematic review of clinical effectiveness



* one of the systematic reviews was the previous HTA report written for NICE, so this was not data extracted. It is discussed briefly in Section 3.10.

An overview of the included studies is given in Table 3. Only one SGA paper and one TS paper reported final height; none of the other conditions' studies reported final height as an outcome measure. None of the papers reported specific QoL measures. All disease areas included at least one paper which reported outcomes on height gained, body composition, biochemical markers and AE. The characteristics and quality assessment of the included studies are discussed in each of the relevant disease-specific results chapters.

Table 3 Included RCTs

Author and date	Total n	Outcomes included in the systematic review						
		Final height	Height gained/ Height SDS	growth velocity /growth SDS	Body comp	Biochem/ metab markers	QoL	AE
GHD								
Soliman [84]	19		✓	✓		✓		
TS								
Davenport 2007 [85]	89		✓	✓	✓	✓		✓
Stephure 2005 [86] + Rovet 1993[87]	154	✓	✓	✓				✓
Quigley 2002 [11]	232			✓				✓
Gravholt 2002 [88]	12				✓	✓		
Gravholt 2005 [89]	9				✓	✓		
Johnston 2001 [90]	58		✓					
PWS								
Festen 2007 [91]	20		✓		✓	✓		
Festen 2007 [92]	29		✓		✓	✓		✓
De Lind van Wijngaarden 2009 [93] and Festen 2008 [94]	42 infants; 49 children		✓		✓	✓		
Carrel 1999 [95] and Myers [96]	54		✓	✓	✓	✓		✓
Carrel 2004 [21] and Myers [97] and Whitman [98]	32		✓	✓	✓	✓		✓
Hauffa 1997 [99]	19		✓	✓		✓		✓
Lindgren [100] and [101]	29		✓	✓	✓	✓		✓
Haqq et al., 2003 [102]	14		✓	✓	✓	✓		✓
CRI								
Sanchez 2002 [103]	23		✓	✓	✓			✓
Hokken-Koelega 1991 [104]	20			✓		✓		✓
Hokken-Koelega 1996 [105]	11			✓		✓		✓
Powell 1997 [106]	69		✓		✓	✓		
Broyer [107]	203		✓	✓				
Fine 1994 [108]	125		✓	✓	✓	✓		✓
SHOX-D								
Blum 2007 [48]	52		✓	✓		✓		✓
SGA								
De Schepper 2007 [109]	40		✓		✓			✓
Lagrou 2008 [110]	40		✓		✓			✓
Carel 2003 [111]	168	✓	✓					✓
de Zegher 1996 [112]	54		✓	✓	✓	✓		✓
de Zegher 2002 [113]	13		✓	✓	✓			
Philip 2009[114]	151		✓			✓		

3.2.1.1 Comparison with previous review

The previous review by Bryant and colleagues[5] included a number of studies which were excluded from the present review. As described in Section 3.1.2 and the research protocol, the present review only included RCTs as these form the highest level of evidence in the hierarchy of clinical trial designs.[83] The previous review included two non-RCT studies for GHD,[115,116] four for TS,[117-120] two for CRI[121,122] and one for PWS.[123] In addition, the previous review included two RCTs for TS which have been excluded from the present review. The first of these, by Rosenfeld and colleagues,[124,125] was excluded from the present review as it used methionyl growth hormone (met-GH) rather than rhGH. The second TS RCT was by Ross and colleagues,[126] which reported cognitive function. This was not one of the outcome measures listed in the inclusion criteria for the present review, so this RCT was excluded. The previous review also included a PWS RCT by Whitman and colleagues[127] which was considered for the current review. However, the study reported psychological outcomes rather than a measure of health related QoL, so this study did not meet our inclusion criteria.

3.3 Growth Hormone Deficiency

3.3.1 Quantity and quality of research available

One study met the inclusion criteria for this review, and the key characteristics are presented in Table 4. The full data extraction form in Appendix 4 has further details.

Soliman and colleagues[84] recruited two groups of growth hormone deficient children and one group of children who were not growth hormone deficient. These groups were then subdivided into treatment groups: group 1a received 30 U/m^2/wk of rhGH and group 1b 15 U/m^2/wk. Group 2a received 15 U/m^2/wk and group 2b no treatment. Group 3 (non-GHD short children) was subdivided in the same way as group 2. Group 2 was the only group in this study with growth hormone deficiency and with children randomised to either rhGH or no treatment, and as such is the only group considered in this report. The treatment groups' baseline characteristics were similar. The study used a dose of 15 U/ m^2/wk, and it is not clear how this corresponds to the licensed dose as neither mg nor IU are used.

Table 4 Characteristics of GHD study

Reference	Intervention	Control group	Total randomised and withdrawals	Duration of randomised treatment
Soliman et al., 1996 [84]	GH 15 U/m^2/wk n=9 Overall mean age± SD: 6.8 ± 2.1	No treatment n=10 Overall mean age± SD: 6.8 ± 2.1	Total n=19 No withdrawals reported	1 year

Table 5: Quality assessment of included GHD study

	Soliman [84]
1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Inadequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Adequate
10. Were withdrawals and dropouts completely described?	Adequate

Overall the quality of the reporting of the included study was mixed (Table 5). No details were given on randomisation or allocation to treatment groups. For example Soliman and colleagues[84] recruited children into specified groups according to peak GH response to provocation, and these groups were then divided at random into two subgroups. No further details were given. The low patient numbers will affect interpretation of results from this trial.

The comparator group did not receive placebo: this could mean that both care providers and patients would have been aware of whether they were receiving treatment, which in turn can affect reporting of some outcomes. Soliman and colleagues[84] appear to have carried out an intention to treat analysis (ITT), which can protect against attrition bias.

3.3.2 Growth outcomes

The Soliman[84] study reported growth velocity and height standard deviation score (HtSDS), and these are presented in Table 6. The data extraction forms in Appendix 4 list further outcome measures such as bone age.

Table 6 Growth outcomes for GHD

Study	**Mean (SD)**	**GH**	**No treatment**	**P Value**
Soliman and colleagues[84] GH 15 U/m^2/week (n=9) vs. no treatment (n=10) ; 12 months	HtSDS	-2.3 ± 0.45	-2.8 ± 0.45	P<0.05
	Growth velocity (cm/yrs)	8.4 ± 1.4	5.7 ± 1.8	P<0.05

Children in the treated group in the Soliman study grew an average of 2.7 cm/year faster than those receiving no treatment in the 12 months of the study, and the difference between groups was statistically significant ($p<0.05$). Similarly children in the treated group had a statistically significantly higher height standard deviation score: -2.3 ± 0.45 vs. -2.8 ± 0.45 in the untreated group ($p <0.05$).

3.3.3 Body composition outcomes

Soliman and colleagues did not report body composition outcomes.

3.3.4 Biochemical markers

The results reported for IGF-I levels in the Soliman study[84] are shown in Table 7. Further biochemical markers, such as insulin, are included in the data extraction tables in Appendix 4.

Table 7: Biochemical markers in GHD studies

Study	Outcomes	GH	Control	P Value
Soliman[84] GH 15 U/m^2/week (n=9) vs. no treatment (n=10); 12 months	IGF-I (ng/ml)	91.2 ± 30.4	49.4 ± 19	p <0.05

The IGF-I levels at 12 months are statistically significantly higher in the treated than the untreated group, 91.2 ± 30.4 vs. 49.4 ± 19.

3.3.5 Quality of life

Soliman and colleagues did not report QoL results.

3.3.6 Adverse events

Adverse events were not reported by Soliman and colleagues.

3.3.7 Summary

- One trial examining the effectiveness of rhGH for GHD met the inclusion criteria for the review.
- The quality of the included study was mixed. It was an unblinded study, which can have an impact on outcome reporting, but did report an ITT analysis.
- Children in the rhGH group grew 2.7cm/yr faster than children in the untreated group during the one year study, and had a statistically significantly higher HtSDS: -2.3 ± 0.45 vs. -2.8 ± 0.45.
- The IGF-I levels were statistically significantly higher in the treated group than in the untreated group.
- The included study did not report QoL or AE.

3.4 Turner syndrome

3.4.1 Quantity and quality of research available

Six studies assessing the effectiveness of growth hormone for growth restriction in Turner syndrome met the inclusion criteria for the review.[11,85,86,88-90] The key characteristics of these studies are presented in Table 8 –Table 13; Appendix 4 has further details.

Two of the included studies were of a cross-over design,[88,89] and these compared doses of 0.1 IU/kg/d[88] and a mean of 1.3 ± 0.3mg/day (alone or in combination with oestradiol) [89] with placebo. The group receiving oestradiol is not discussed further here. Of the remaining studies, two compared rhGH with no treatment,[85,86] one with low dose oestrogen,[90] and one with placebo.[11] Stephure and colleagues[86] administered a rhGH dose of 0.30mg/kg/wk with a maximum weekly dose of 15mg. The dose of 50µg in the Davenport study[85] is comparable with that of Stephure and colleagues. Those in the Quigley study were slightly different: Group 1 received 0.27mg/kg/wk, Group 2 received 0.36mg/kg/wk. Johnston and colleagues gave a dose of 28-30IU/m^2/week. All studies included at least one treatment arm with a dose that was broadly comparable with the licensed dose of 45-50mcg/kg/d or 1.4 mg/m^2/d.

Four of the six included studies reported growth outcomes including height gain, and change in height standard deviation score.[11,85,86,90] The remaining two studies reported body composition and biochemical marker outcomes.[88,89]

The trials varied considerably in size. The two crossover trials were small, with 12[88] and nine[89] participants. The Stephure[86] and Quigley[11] studies were larger, with 154 and 232 participants, respectively. Johnston and colleagues[90] recruited 58 patients and Davenport and colleagues, 89.[85] The included trials also ranged in length. The groups in Quigley and colleagues[11] remained randomised for 18 months, the Davenport study[85] for two years and the Johnston study lasted for one year.[90] Protocol completion in the Stephure[86] study was defined as annualized GV less than 2cm/yr and bone age of 14 years or greater, which we have interpreted to mean final height. In contrast the two Gravholt studies[88,89] were short crossover trials, with rhGH treatment for two months.

Five of the six trials recruited broadly similar age groups, whilst the sixth by Davenport and colleagues[85] specifically targeted very young girls with Turner syndrome. As a result their girls have much younger mean ages of 1.98 ± 1.01 and 1.97 ± 1.01 for treatment and control groups, respectively.

Four of the included studies reported baseline characteristics that were similar between groups. However, none reported p values for between group differences, so there may have been small differences at baseline. For example, in the study by Stephure and CGHAC 2005,[86] girls in the rhGH group were on average 3cm shorter than those in the control group. The SD values indicate overlapping CI, suggesting there is no statistically significant difference between the two groups. However, the 3cm difference could have an impact on end of study height. The other two studies,

reported by Gravholt and colleagues, were of cross-over design. One reported baseline characteristics for the whole study group[89] and the other did not appear to report any baseline conditions.[88]

Table 8 Characteristics of Turner syndrome studies

Reference	Intervention	Control group	Total randomised and withdrawals	Duration of randomised treatment
Stephure and CGHAC 2005[86] Rovet et al, 1993 [87]	rhGH 0.30mg/kg/wk n=76 Mean age (± SD): 10.3 ± 1.8	no rhGH treatment n=78 Mean age (± SD): 10.9 ± 1.7	Total n=154 Sample attrition: rhGH: n=15 control: n=35	until HV<2cm/yr and bone age ≥14yr
Davenport et al, 2007[85]	rhGH 50 µg/kg/d n=45 Mean age (± SD): 1.98 ±1.01	No treatment n=44 Mean age (± SD): 1.97 ± 1.01	Total n=89 Sample attrition: rhGH: n=4 control: n=6	2 years
Gravholt et al., 2002 [88]	rhGH 0.1 IU/kg/d Overall age range: 9.5-14.8 years, (median 12.9)	Placebo Overall age range: 9.5-14.8 years, (median 12.9)	Total n=12 Withdrawals not reported	Cross-over RCT, 2 months in each arm
Gravholt et al, 2005[89]	rhGH (1.3 ± 0.3) mg/d Overall mean age (± SD): 15.9 ± 1.8	placebo Overall mean age (± SD): 15.9 ± 1.8	Total n=9 Sample attrition: n=1	Cross-over RCT, 2 months in each arm
Johnston et al., 2001[90]	rhGH 28-30 IU/m^2 /wk (n=22) Mean age (range): 9.0 (5.2 - 15.4)	Ethinyloestradiol [a] 50-75 ng/kg/day (n=13) Mean age (range): 9.1 (6.0 – 13.7)	Total n=58[b] Sample attrition: n=12	1 year
Quigley et al, 2002[11]	rhGH 0.27 mg/kg/wk (n=45) Mean age (± SD): 9.7 ± 2.7 rhGH 0.36 mg/kg/wk (n=49) Mean age (± SD):9.8 ± 2.9	Placebo (n=41) Mean age (± SD): 9.4 ± 2.7	Total n=232[b] Sample attrition: n=8	18 months

[a] low dose oestrogen; [b] including additional study arm(s) not relevant here.

Table 9: Quality assessment of included Turner syndrome studies

	Stephure and CGHAC[86]	Davenport et al, [85]	Gravholt et al. 2002[88]	Gravholt et al. 2005[89]	Johnston et al.[90]	Quigley et al.[11]
1. Was the assignment to the treatment groups really random?	Un	Ad	Un	Un	In	Un
2. Was the treatment allocation concealed?	Un	Ad	Un	Un	Un	Un
3. Were the groups similar at baseline in terms of prognostic factors?	Rep	Rep	Not rep	Not rep	Rep	Rep
4. Were the eligibility criteria specified?	Ad	Ad	In	In	In	Ad
5. Were outcome assessors blinded to the treatment allocation?	Un	Un	Un	Un	Un	Un
6. Was the care provider blinded?	In	In	Un	Un	Un	Un
7. Was the patient blinded?	In	In	Un	Ad	Un	Par
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Ad	Ad	Ad	Ad	Ad	In
9. Did the analyses include an ITT analysis?	In	In	In	In	In	In
10. Were withdrawals and dropouts completely described?	Ad	Ad	In	Ad	Ad	Ad

Un=unknown; ad=adequate; rep=reported; not rep = not reported; in=inadequate; par=partial

The six included trials were generally of poor methodological quality, and poorly reported (Table 9). Only one reported adequate methods of randomisation to treatment groups.[85] Davenport and colleagues stratified their participants by age and then randomised them using a blinded phone-in process. Four of the six trials did not describe randomisation techniques.[11,86,88,89] Johnston and colleagues reported that five participants were reallocated from the oestrogen group to receive rhGH: it is unclear when this occurred and therefore method of randomisation was judged inadequate.

Concealment of treatment allocation was also judged to be adequate in the Davenport trial, and 'unknown' in the remaining five. In the Gravholt[89] study it is unclear how allocation to treatment groups had taken place. The study had only nine participants, and these were simply reported to have been given the treatment regimen sequentially and in random order.

Blinding of participants, those who provide care and those who assess outcomes can protect against the reporting of some outcomes being affected by the knowledge of which treatment is being received. Blinding of outcome assessors, care providers and patients was judged 'unknown', 'inadequate' or 'partial' in five of the six trials; Gravholt and colleagues[89] adequately blinded their patients by administering placebo in place of both rhGH and the oestradiol.

None of the six studies included here employed an ITT analysis. This kind of analysis can protect the study from attrition bias, where, for example, participants withdrawing from the treatment arm could represent AE or treatment failure.

3.4.2 Growth outcomes

Four of the six included studies reported growth outcomes, and key measures are shown in Table 10. Please see Appendix 4 for additional outcomes. Neither of the studies by Gravholt and colleagues[88,89] reported growth outcomes.

Table 10 Growth outcomes for Turner syndrome studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Stephure and CGHAC [86] Protocol completion rhGH 0.30mg/kg/wk (n= 61) vs. no treatment (n= 43)	Height (cm)	147.5±6.1	141.0 ± 5.4	p<0.001
	Change in height (cm)	28.3 ± 8.9	19.0 ± 6.1	p<0.001
	Height SDS (age-specific turner)	1.4 ± 1.0	0.2 ± 0.9	p<0.001
	Height SDS (adult Turner)	0.7 ± 0.9	-0.3 ± 0.8	p<0.001
	Change in HtSDS (age-specific Turner)	1.6 ± 0.6	0.3 ± 0.4	p<0.001
Stephure and CGHAC [86] Addendum follow-up rhGH 0.30mg/kg/wk (n= 40) vs. no treatment (n= 19)	Height (cm)	149.0 ± 6.4	142.2 ± 6.6	p<0.001
	Change in height (cm)	30.3 ± 8.3	21.6 ± 6.2	p<0.001
	Height SDS (age-specific Turner)	0.9 ± 0.9	-0.1 ± 1.0	p<0.001
	Height SDS (adult Turner)	0.9 ± 0.9	-0.1 ± 1.0	p<0.001
	Change in HtSDS (age-specific Turner)	1.1 ± 0.5	0.0 ± 0.5	p<0.001
Davenport et al.[85] GH (n= 41) vs. no treatment (n=37), 2 years	Height (cm)	99.5 ± 7.6	91.9 ± 7.2	<0.0001
	Height SDS	-0.34 ± 1.10	-2.16 ± 1.22	<0.0001
	Height velocity (cm/yr)	8.4 ± 1.6	5.5 ± 1.8	<0.0001
	Height velocity SDS	0.70 ± 1.11	-1.63 ± 1.29	<0.001
Johnston et al.[90] **rhGH** 28-30 IU/m^2 /wk (n=?)* vs. oestrogen (n=?)*; 1 year	Change in HSDS in first year	+0.7 (0.7)	+0.4 (0.9)	<0.05
Quigley et al.[11] GH 1: rhGH 0.27 (n=45) GH 2: rhGH 0.36 (n=49) vs. placebo (n=41); 1 year	Height velocity 0-18 months (cm/yr)	1: 6.6 ± 1.1 2: 6.8 ± 1.1	4.2 ± 1.1	<0.001

* n unclear for this outcome

Two studies reported height at the end of the study: both found a statistically significant difference between the treated and untreated groups (p<0.0001).[85,86]

Children in the treated group in the Stephure study[86] were 6.5cm taller on average than the untreated group at protocol completion. However, there was a 3 cm difference between the groups' mean heights at baseline. Mean change from baseline was therefore 9.3cm more in the rhGH than in the untreated group at the end of protocol completion (28.3 ± 8.9 vs. 19.0 ± 6.1).

The Stephure study[86] also reported an addendum follow-up (approximately 10 years since randomisation) which included 66% of rhGH patients and 44% of the control group. The treated group's mean final height was 149.0 ± 6.4 compared with 142.2 ± 6.6 in the untreated group ($p < 0.001$), i.e. a difference of 6.8cm. Mean change from baseline to final height was 8.7cm more in the rhGH than in the untreated group.

In the Davenport study[85] the mean difference was 7.6cm (height at study end: 99.5 ± 7.6 cm in the treated group vs. 91.9 ± 7.2cm in the untreated group, $p<0.0001$).

Height standard deviation score (HtSDS) is also reported by these two studies.[85,86] Both authors report statistically significant differences between groups for this outcome, with the treated groups both achieving higher HtSDS. In the Stephure study[86] the HtSDS is reported for the age-specific Turner population and for the adult Turner population.

The difference in change in height was statistically significant between groups in the two studies that reported it. Stephure and colleagues report a change in height at protocol completion of 28.3 ± 8.9cm vs. 19 ± 6.1 in the untreated group, $p<0.001$. Davenport and colleagues[85] reported a two year height gain of 20.4 ± 3.3cm (treated group) vs. 13.6 ± 3.5cm (untreated group,) $p<0.001$, (not shown in table). Change in HtSDS in both the Stephure[86] and Johnston[90] studies was higher in the treated than untreated group: 1.6 ± 0.6 (treated) vs. 0.3 ± 0.4 (untreated) $p<0.001$ at protocol completion in the Stephure study; 0.7 (0.7) vs. 0.4 (0.9) $p <0.05$ in the Johnston study after one year.

Height velocity was statistically significantly greater in the treated groups in the Stephure,[86] Davenport,[85] and Quigley[11] studies. Davenport and colleagues reported GV at the end of the first and second year. While this was greater in the treated groups at both times, GV fell in the second year in both groups: 8.4 ± 1.6cm/ yr (treated group) vs. 5.5 ± 1.8 (untreated). Additionally, Davenport and colleagues measured GV SDS at the end of the first and second years. Again, this was greater in the treated group at the end of the first year: 1.75 ± 1.25 vs. 0.8 ± 0.95, $p <0.001$, but was reduced by the end of the second year in both groups: 0.70 ± 1.11 (treated) vs. -1.63 ± 1.29 (untreated), $p<0.001$. Quigley and colleagues reported GV after 18 months. This was broadly similar in both the lower and higher rhGH dose groups: both were significantly higher than that in the placebo group: 6.6 ± 1.1 (GH

0.27/Pla group) vs. 6.8 ± 1.1 (GH 0.36/Pla group) vs. 4.2 ± 1.1 (Pla/Pla group), p<0.001 compared with placebo.

Bone age differences for the younger participants in the Davenport study were statistically significant:[85] the growth hormone treated group at 2 years had a mean bone age of 4.24 ± 1.35 vs. 3.38 ± 1.11 in the untreated group, p=0.0033. Davenport and colleagues also reported bone age – chronological age; this is lower in the treated group, and the difference was statistically significant: 0.64 ± 0.80 vs. 0.21 ± 0.96 p<0.001.

3.4.3 Body composition outcomes

Three of the TS studies reported body composition outcomes, and these are presented in Table 11. One of the studies reported weight, WtSDS and BMI, [85] the remaining two reported fat mass (FM), bone mineral content (BMC) and lean body mass (LBM) for arms, legs, trunk, head and as a total.[88,89] Please see Appendix 4 for BMC results.

Table 11: Body composition outcomes for Turner syndrome studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Davenport et al.[85] GH (n= 41) vs. no treatment (n=37), 2 years	Weight (kg)	16.62 ± 2.86	13.81 ± 2.50	<0.0001
	WtSDS	0.20 ± 1.06	-1.37 ± 1.36	<0.0001
	BMI (kg/m^2)	16.72 ± 1.70	16.24 ± 1.29	0.1724
Gravholt et al. [88] GH 0.1 IU/kg/d vs. placebo; 2 months[a]	FM total (g/ kg)	231.0 ± 49.5	247.8 ± 58.1	0.04
	LBM total (g/ kg)	725.4 ± 44.8	710.5 ± 54.6	0.05
Gravholt et al.[89] GH 1.3mg/day vs. placebo, 2 months[b]	FM total (g/ kg)	274.5 ± 55.5	312.9 ± 74.7	nr
	LBM total (g/ kg)	692.8 ± 55.5	655.2 ± 73.7	nr

[a]cross-over study, total n=12; [b]cross-over study, total n=9

Weight and WtSDS were significantly greater in the group receiving rhGH than in the untreated group in the Davenport study,[85] reported as 16.62 kg ± 2.86 vs. 13.81 kg ± 2.50 and 0.20 ± 1.06 vs. -1.37 ± 1.36, respectively (p<0.0001 for both comparisons).

Two studies considered FM, BMC and LBM.[88,89] In both studies the total FM was greater in the untreated group than the in the treated group, and LBM was slightly higher in treated than in untreated patients (Table 11). The differences between groups were of borderline statistical significance in one study[88] but no p values were presented in the other study.[89]

3.4.4 Biochemical markers

Three of the studies[85,88,89] reported biochemical outcomes. Key results are shown in Table 12 – other outcomes are in Appendix 4.

Table 12: Biochemical markers in TS studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Davenport et al.[85a] GH (n= 41) vs. no treatment (n=37), 2 years	IGF-I SDS	1.26 ± 0.72	-0.69 ± 0.84	<0.0001
	IGFBP-3 SDS	0.97 ± 0.94	-1.12 ± 1.13	<0.0001
	ΔIGF-I SDS	1.53 ± 0.93	-0.09 ± 0.87	nr
Gravholt et al. [88] GH 0.1 IU/kg/d vs. placebo; 2 months[a]	IGF-I (µg/l)	380.5 ± 116.3	179.8 ± 79.4	<0.0005
	IGFBP-3 (µg/l)	5982 ± 1557	4344 ± 787	0.002
Gravholt et al.[89] GH 1.3mg/day vs. placebo, 2 months[b]	IGF-I (µg/l)	661 ± 192	288 ± 69	nr
	IGFBP-3 (µg/l)	5157 ± 741	4146 ± 573	Unclear

[a]baseline data missing for eight control subjects and three GH-treated subjects; endpoint data missing for four control subjects and seven rhGH subjects

Two studies reported mean levels of IGF-I at end of treatment. In both studies IGF-I levels were statistically significantly higher in the group receiving rhGH. One study[88] reported values of 380.5 ± 116.3 vs. 179.8 ± 79.4 in the treated and untreated group, respectively ($p<0.0005$). The other[89] reported 661± 192 vs. 288 ± 69 (p not reported) for treated and untreated patients, respectively.

Davenport and colleagues[85] reported that IGF-I SDS was significantly greater in the treated group (1.26 ± 0.72 vs. -0.69 ± 0.84; $p<0.0001$). Change in IGF-1 SDS from baseline to year two was 1.53 ± 0.93 vs. -0.09 ± 0.87 in the treated and untreated group, respectively.

One Gravholt study[88] reported that IGFBP3 levels were statistically significantly higher in the treated group than in the untreated group (5982 ± 1557 vs. 4344 ± 787, respectively; $p = 0.002$). The other study by Gravholt and colleagues reported higher IGFBP3 SDS values in treated patients, but no clear p value was reported.[89] Davenport and colleagues found that IGFBP3 SDS was higher in their treated group (0.97 ± 0.94 vs. -1.12 ± 1.13; $p <0.0001$).

Fasting glucose and fasting insulin were reported in the two studies by Gravholt and colleagues,[88,89] both of which were raised in the groups receiving growth hormone in each study. Mean glucose (nmol/l) were 4.28 ± 0.59[88] and 4.46 ± 0.40[89] in the treated groups, vs. 4.02 ± 0.44[88] and 4.04 ± 0.47[89] in the untreated groups. This difference reached statistical significance in the first study,[88] $p=0.046$. Mean fasting insulin levels in the first Gravholt study[88] were 17.17 ± 8.30 vs. 8.58 ± 4.27 $p=0.007$.

3.4.5 Quality of life

None of the TS studies reported QoL as an outcome.

3.4.6 Adverse events

Adverse events (AE) were only reported by four of the studies.[11,85,86,90] Details presented by three of the studies are shown in Table 13 (the fourth study did not present figures.[90])

Table 13 AE for Turner syndrome studies

Study	AE (n)	GH	Control	P Value
Stephure and CGHAC [86] **GH (n=74) vs. no treatment (n=64)**	Surgical procedures	37	17	0.005
	Otitis media	35	17	0.014
	Ear disorder	15	4	0.024
	Joint disorder	10	2	0.036
	Respiratory disorder	8	1	0.037
	Sinusitis	14	4	0.041
	Goiter	0	4	0.004
	Death (ruptured aortic aneurysm)	0	1	nr
	Elevated transamine levels	1	0	nr
	Intracranial hypertension	1	0	nr
Davenport et al.[85] GH (n= 45) vs. no treatment (n=44), 2 years	Serious AE, n (%)	4 (9)	4 (9)	nr
	Treatment emergent AE, n (%)	42 (93)	43 (98)	nr
Quigley et al.[11]	Otitis Media (occurrence/ worsening) , n (%)	54/186 (29%)	6/46 (13%)	0.037

The group receiving growth hormone in the Stephure study[86] experienced a statistically significantly greater level of all AE (where statistical significance was reported), with the exception of goiter and one instance of death from ruptured aortic aneurysm which occurred in the untreated group. The one case of elevated transamine levels in the treated group led to withdrawal from the study.

Davenport and colleagues[85] report the same level of serious AE for both the treated and untreated groups. For treatment emergent AE defined as 'events or conditions that began or worsened after study entry,' results were similar. There were 42 (93%) in the treated group and 43 (98%) in the untreated group. Most treatment emergent AE were ear disorders.

Quigley and colleagues[11] found a significant difference in levels of occurrence or worsening of otitis media between the treated group (29%) and the control group (13%), p=0.037. Ear pain and ear disorder were reported as not differing between groups. Three girls discontinued rhGH due to hypertension, ulcerative colitis and brain tumour. The authors stated that these were not directly related to GH. Overall, AE were not presented separately for the groups, however five were reported to have accidentally overdosed on the study drug. Five further events described as possibly related to the study drug were hypertension (two), surgical procedures (two), and scoliosis (one).

Five participants were reallocated from the group receiving oestrogen to rhGH after concerns over early breast development in the study by Johnston and colleagues.[90] Seven patients developed 'coincidental disorders' not severe enough to warrant treatment discontinuation. The authors reported that compliance problems led to the withdrawal of four patients, but no details were given. It is unclear which treatment groups these latter events occurred in.

3.4.7 Summary

- Six trials examining the effectiveness of growth hormone for growth disturbance in patients with TS met the inclusion criteria for the review.
- The reporting and methodological quality of the studies was poor. Of the six included studies, one reported adequate randomisation to treatment groups,[85] one study described adequate concealment of treatment allocation[85] and one adequately blinded the patient to treatment by administering placebo.[89] None of the included trials employed an ITT analysis.
- Children in the rhGH group in the Stephure[86] study grew an average of 9.3cm more from baseline than those in the untreated group. In a study of younger children[85] the difference was 7.6cm. Both of these were statistically significant results. In the same two studies[85,86] the groups receiving rhGH achieved a significantly higher HtSDS.
- Change in height, and change in HtSDS were statistically significantly greater in the groups treated with r-h GH.[85,86,90]
- Height velocity was greater in the treated groups in three studies that reported this outcome,[11,85,86] although this was greater in the first year and fell in the second year in both treatment groups where this was reported separately.[85]
- One study[86] found a significant difference in bone age between groups, being higher in the treated patients.
- Fat mass and lean body mass were reported in two studies.[88,89] In both, the total fat mass was at a lower level in the treated groups, compared with those untreated, and lean body mass was higher in the treated groups compared with untreated. There was no statistically significant difference in BMI between treated and untreated girls in one study.[85]
- The IGF-I levels were substantially higher in the treated groups in the studies reporting this outcome.[88,89]IGF-I SDS was also significantly higher in the group receiving GH.[85] Levels of IGFBP3 and IGFBP3 SDS were also found to be higher in children treated with growth hormone.[85,88,89]
- Levels of fasting glucose and fasting insulin were both raised in the treated groups in two studies.[88,89]
- There were variable levels of detail in the reporting of AE across the six studies. Two studies did not discuss these.[88,89]In those studies that did, no clear picture emerges. One found greater levels

of AE in the treated group,[86] one found similar levels across groups,[85] one found significantly higher levels of or worsening of otitis media, and one reported seven patients with 'coincidental disorders' and four withdrawals due to compliance problems, but gave no further details.

3.5 Prader-Willi syndrome

3.5.1 Quantity and quality of research available

Eight RCTs in 13 publications of the clinical effectiveness of rhGH in patients with PWS met the inclusion criteria for this review.[21,91-102] Their key characteristics are shown in Table 14– please see Appendix 4 for further details.

It was not possible to perform any meta-analysis of outcomes from the PWS studies due to variation in the trials' participants' ages, dosing calculations and methods of presenting results. The included studies had well matched patient groups, whose baseline characteristics were generally similar in the treated and untreated groups. Median baseline HtSDS was lower in the rhGH group than in the untreated group in the study reported by both Festen and colleagues[94] and by de Lind van Wijngaarden and colleagues[93], although the interquartile ranges were similar (-2.0 (-3.1 to -1.7) vs. -2.5 (-3.3 to -1.9), respectively). Other exceptions were the cross-over study by Haqq and colleagues,[102] which presented baseline characteristics for the study population as a whole, and the study by Lindgren and colleagues[100,101] which reported slightly lower baseline GV SDS in the rhGH group (-1.9 ± 2.0, range -6.4 to -0.9 vs. -0.1 (SD not reported) range -1.7 to -2.71).

Table 14 Characteristics of included PWS studies

Reference	Intervention	Control group	Total randomised and withdrawals	Duration of randomised treatment
Carrel et al. 2004[21] and Myers et al. 2007[97] Whitman et al. 2004[98]	1mg/m^2/d rhGH n=15 Mean age ± SD (months): 13 ± 8	no treatment n=14 Mean age ± SD (months): 15 ± 0	N=32 Sample attrition: n=3 [a]	1 year
Carrel et al. 1999[95] and Myers et al. 1999[96]	GH 1 mg/m^2/d n=35 Mean age (y): 9.8	no treatment n=19 Mean age (y): 10.0	N=54 no withdrawals	1 year
de Lind van Wijngaarden et al. 2009 [93]; Festen et al. 2008[94]	1mg/m^2/d Infants (<3.5 years): n=19 Children (>3.5 years): n=23	no treatment Infants (<3.5 years): n=19 Children (>3.5 years): n=21	N=104 enrolled Sample attrition: 4 infants and 5 children	1 year for infants 2 years for children

	Median (IQR) age: Infants: 2.0 (1.6-3.1) Children: 6.8 (5.4-8.8)	Median (IQR) age: Infants: 1.3 (1.0 – 2.8) Children: 5.9 (4.7 - 7.4)		
Festen et al. 2007[91]	GH 1mg/ m^2/d n=10 Median age (IQR)(yr): 6.2 (5.1-71)	no treatment n=10 Median age (IQR)(yr): 5.8 (4.9-7.8)	N=20 withdrawals: none	2 years
Festen et al., 2007 [92]	GH 1mg/ m^2/d N=15 Median (IQR) age, yr: 2.3 (1.7-3.0)	no treatment N=14 Median (IQR) age, yr: 1.5 (1.2-2.7)	N=43 Sample attrition: n=14	12 months
Haqq et al. 2003[102]	GH 0.043 mg/kg/d (n=6) Overall mean age ± SD (yrs): 9.7 ± 3.3	Placebo (n=6) Overall mean age ± SD (yrs): 9.7 ± 3.3	14 randomised Sample attrition: n=2	Cross-over RCT, 6 months in each arm
Hauffa 1997[99]	GH: 0.15 IU/kg/d n=8 Mean age ± SD (yrs): 8.25 ± 2.4	no treatment n=9 Mean age ± SD (yrs): 7.56 ± 2.0	N=19 Sample attrition: n=3	1 year
Lindgren et al. 1998;[101] 1997[100]	GH 0.1 IU/kg/d n=15 Mean age (range) (yrs): 6.8 (3.6-11.9)	no treatment n=14 Mean age (range) (yrs): 6.4 (3.3-11.7)	Total n=29 Sample attrition: n=2	1 year

[a] difference between patient numbers in Whitman[97] and Carel[21]=3

Five of the studies were RCTs which compared 1mg/m^2/day rhGH with no treatment for one[21,92-98] or two[91,93,94] years. The study by Haqq and colleagues[102] was a cross-over RCT which compared 0.043 mg/kg/day rhGH with placebo injections, with patients spending 6 months in each treatment arm. There does not appear to have been a wash-out phase between the two treatment phases, which could affect the generalisability of results.

The doses used in the included studies reflect the various marketing authorisations for this drug (0.035 mg/kg body weight or 1.0mg/m^2 BSA), with 1 IU of rhGH being equivalent to approximately 0.33mg/kg. The study reported by both de Lind van Wijngaarden and colleagues[93] and Festen and colleagues[94] reported results separately for infants and children. Two RCTs reported results for infants and toddlers aged between one and two and a half years.[21,92,97,98] The five remaining trials were in children aged between approximately 6 and 10 years old. The studies were generally small, randomising between 14[102] and 54[95,96] children. The study reported by both de Lind van Wijngaarden and colleagues[93] and Festen and colleagues[94] had a total of 91 participants, but since children and

infants were randomised separately, the randomised comparisons were of rhGH vs. no treatment within two smaller groups (42 infants and 49 children). This was the only study to report a sample size/power calculation,[93,94] and it is not clear whether the other studies were adequately powered to detect a difference between treatment groups.

With the exception of the two RCTs by Festen and colleagues,[91,92] the studies did not clearly state which of their reported outcomes were primary or secondary measures of effect. Seven of the eight trials reported measures of body composition. The two RCTs by Festen and colleagues[91,92] focussed on body composition and biochemical markers, and did not report any measure of change in height. The other six studies all reported GV SDS or an indicator of linear growth velocity.[102] IGF-1 and other biochemical markers were reported by five RCTs.[21,92-99]

One RCT was reported in three papers, by Carrel and colleagues,[21] Myers and colleagues[97] and Whitman and colleagues.[98] The most complete data was reported by Carrel and colleagues,[21] and this data is included in the tables in this section.

Table 15 Quality assessment of included PWS studies

	Carrel et al.;[21] Myers et al.[97] Whitman et al.[98]	Carrel et al.;[95] Myers et al.[96]	De Lind van Wijngaarden et al.[93] Festen et al.[94]	Festen et al.[91]	Festen et al.[92]	Haqq et al.[102]	Hauffa et al.[99]	Lindgren et al.;[101] Lindgren et al.[100]
1. Was the assignment to the treatment groups really random?	un	un	un	un	un	un	un	un
2. Was the treatment allocation concealed?	un	un	un	un	un	un	un	un
3. Were the groups similar at baseline in terms of prognostic factors?	rep	rep	rep	rep	rep	not rep	rep	rep
4. Were the eligibility criteria specified?	ad	ad	ad	ad	ad	ad	ad	ad
5. Were outcome assessors blinded to the treatment allocation?	un	un	un	un	un	un	un	un
6. Was the care provider blinded?	in	in	in	in	in	un	in	in
7. Was the patient blinded?	in	in	in	in	in	ad	in	in

8. Were the point estimates and measure of variability presented for the primary outcome measure?	ad	ad	ad	ad	ad	ad	in	ad
9. Did the analyses include an ITT analysis?	in	ad	in	ad	in	in	in	in
10. Were withdrawals and dropouts completely described?	in	ad	ad	ad	in	in	in	ad

Un=unknown; ad=adequate; rep=reported; not rep = not reported; in=inadequate

The included studies were generally poorly reported (Table 15) and lacked information on method of randomisation or concealment of allocation. It is possible that selection bias could have affected the trials if they were not properly randomised, but there is insufficient information provided on which to make such a judgement. The trial by Haqq and colleagues[102] was a cross-over study, and did not report baseline characteristics separately for the two groups. The other studies reported baseline characteristics which indicated that patients in the two treatment groups were comparable at the start of the study. With the exception of the cross-over trial by Haqq and colleagues,[102] which had a placebo injection group, the studies were open-label, with the comparator groups receiving no treatment. Whilst this could have allowed a degree of bias in reporting and assessing results, measurement of objective outcomes such as height gained is less likely to be open to bias. Only two of the studies reported results on an ITT basis,[91,96] so attrition bias could have affected the remaining studies.

The outcome measures for the included studies are shown in Table 16 to Table 18 below. P values in the tables refer to between-group differences, since this is the comparison of interest for this report. Some of the studies reported statistical significance in change from baseline for each of the treatment groups individually, but not for between-group comparisons. To avoid confusion with the between-group comparison p values, such results have not been included in the tables below and are not discussed in the text. The full data extraction tables in Appendix 4 include any statistical significance for change from baseline for individual treatment groups without between-group comparisons.

3.5.2 Growth outcomes

Changes in height and other growth outcome measures are shown in Table 16. The infants in the study by Carrel and colleagues[21] who received rhGH for a year grew an average of 6.2 cm more than those in the untreated group ($p<0.001$). None of the other studies reported change in height as an outcome measure.

Two studies reported a statistically significant difference in HtSDS at end of treatment between treated and untreated patients.[93-96] Treated patients in the study reported by both Carrel and colleagues[95] and Myers and colleagues[96] had a mean HtSDS of -0.6 ± 1.2 compared with -1.6 ± 1.2 in the untreated group (p<0.01). The study reported by de Lind van Wijngaarden and colleagues[93] and by Festen and colleagues[94] also reported statistically significant improvements in height for rhGH treated infants and children compared with unmatched controls. The rhGH treated infants in their study had a median HtSDS of -0.9 compared with -1.8 in the untreated patients (p=0.003). This reflected a change from baseline HtSDS of +1.2 for treated infants and -0.2 for untreated infants (p<0.0001). After two years of treatment with rhGH, children had a median HtSDS of -0.5 compared with -2.6 in untreated children (p<0.001).[93]

Festen and colleagues[91] reported that the difference between the two groups was statistically significant at year one (year 1 HtSDS -1.3 vs. -2.8; p<0.01). At year two, the difference between the two groups was even greater (-0.6 compared with -3.0 in the treated and untreated groups, respectively), but no p value was reported.[91] The other five studies all reported that HtSDS values were higher in treated than in untreated children, but did not report whether or not differences between groups were statistically significant.

Table 16 Growth outcomes for PWS studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Carrel et al. [21] 1mg/m^2/day rhGH (n= 15) vs. no treatment (n=14), 1 year	Change in height (cm)	15.4 ± 2.3	9.2 ± 3.2	P<0.001
	Height SDS	-0.2 ± 1.5	-1.5 ± 0.7	NR
	Growth velocity SDS	5.0 ± 1.8	1.2 ± 1.4	NR
Carrel et al. [95] and Myers et al. [96] GH 1mg/m^2/d (n=35), vs. no treatment (n=19) 1 year	Height SDS	-0.6 ± 1.2	-1.6 ± 1.2	p < 0.01
	Mean GV (cm/y)	10.1 ± 2.5	5.0 ± 1.8	p < 0.01
	Mean GV SDS	4.6 ± 2.9	-0.7 ± 1.9	p < 0.01
de Lind van Wijngaarden et al. [93] Festen et al. [94] (infants) 1mg/m^2 rhGH (n=19) vs. no treatment (n=19); 1 year	HtSDS median (IQR)	-0.9 (-1.6 to -0.1)	-1.8 (-3.5 to -1.4)	0.003
	ΔHtSDS median (IQR)	1.2 (1.0 to 1.6)	-0.2 (-0.6 to 0.3)	<0.0001
de Lind van Wijngaarden et al. [93] Festen et al. [94] (children) 1mg/m^2 rhGH (n=23) vs. no treatment (n=21); 2 year	HtSDS median (IQR)	-0.5 (-0.8 to 0.0)	-2.6 (-3.4 to -2.3)	<0.0001
	ΔHtSDS median (IQR)	1.4 (1.3 to 1.8)	-0.1 (-0.4 to 0.1)	<0.0001
Festen et al. [91] 1mg/m2/day rhGH (n=10) vs. no treatment (n=10) 2 years	Height SDS median (IQR)	-0.6 (-0.9 to -0.3)	-3.0 (-3.5 to -1.8)	NR
Festen et al. [92] rhGH 1mg/m^2/day (n=15) vs. no treatment (n=14), 1 year	Height SDS median (IQR)	-1.6 (-2.1 to -0.8)	-2.3 (-3.9 to -1.5)	NR
Haqq et al. [102] rhGH 0.043 mg/kg.d (n=12) vs. placebo	HtSDS	-1.2 ± 1.1	-1.3 ± 1.3	NR
	Growth velocity	7.5 ± 3.5	4.5 ± 2.7	P<0.05

(n=12), 6 months	(cm/yr)			
Hauffa [99] rhGH 0.15 IU/kg/day (n=7) vs. no treatment (n=9), 1 year	Height SDS	1.07	-0.25	NR
	HV SDS	5.5	-2.3	P=0.0012
Lindgren et al.[101] **and Lindgren et al.**[100] 0.1 IU/kg/day rhGH (n=15) vs. no treatment (n=12)	HtSDS mean (range)	-0.4 (-2.7 - 1.9)	-1.8 (-5.1 - 0.2)	NR
	Height velocity (SDS) mean ± SD (range)	6.0 ± 3.2 (1.4-11.9)	-1.4 (-3.2 - 0.3)	NR

The five studies which used growth velocity as an outcome measure all reported faster growth in the treated group compared with the untreated group, although statistical significance for differences between groups was only reported for three of these. The mean growth velocity in the study reported by Carrel and colleagues[95] and by Myers and colleagues[96] was twice as fast in the treated group as in the untreated group (10.1 vs. 5.0; $p<0.01$). The corresponding mean growth velocity SDS values were 4.6 in the treated group and -0.7 in the untreated group ($p<0.01$), indicating faster than average growth in the treated group and slower than average growth in the untreated patients. Similarly, Hauffa and colleagues reported a positive growth velocity SDS for treated patients and a negative one for untreated children (5.5 vs. -2.3; $p=0.0012$). Haqq and colleagues[102] calculated growth velocity that was 3cm/year faster in patients receiving rhGH than in patients in the placebo arm (7.5 vs. 4.5, $p<0.05$).

Two of the included studies reported bone age as an outcome measure. There was no statistically significant difference in bone age at follow-up between patients in the treated and untreated groups in the study reported by both Carrel and colleagues[95] and by Myers and colleagues.[96] Lindgren and colleagues[100,101] reported similar change from baseline in both groups (1.4 in the treated group, 1.5 in the untreated group), but did not report whether or not there was any statistical significance to their results.

3.5.3 Body composition

Seven of the trials reported changes in body composition, as shown in Table 17. [21,91-94,100-102] The trial by Hauffa and colleagues[99] did not report any results but stated that there were no significant within- or between group changes for BMI, skinfold thickness, waist or hip circumference.

Four of the trials reported a statistically significantly lower percentage of body fat in patients treated with rhGH compared with no treatment or placebo. In the trial reported by Carrel and colleagues [21] mean percentage body fat was 10% lower for treated patients than for untreated patients ($p=0.03$). On average treated patients in this trial experienced an approximately 5% reduction in body fat, compared with an average 4% increase in the untreated patients' body fat ($p=0.001$). The other two trials which

found a statistically significant difference reported that treated patients had approximately 4% (Haqq and colleagues[102]) or 7% (Carrel[95] and Myers[96]) less body fat than those in the comparator group. De Lind van Wijngaarden and colleagues[93] did not report percentage body fat for infants, but did report this outcome for the children in their study who were over 4 years of age (n=unclear). Children who received rhGH for a year had a median percentage body fat of 1.5%, compared with 2.3% in the control group (p<0.001). After two years of treatment, the values were 1.9% vs. 2.4% for the treated and untreated groups respectively (p<0.001).

Table 17 Body composition outcomes for PWS studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Carrel et al. [21] 1mg/m^2/day rhGH (n= 15) vs. no treatment (n=14), 1 year	Mean % body fat	23.2 ± 8.9	32.7 ± 8.8	0.03
	Change in body fat	-4.8% ± 5.7%	+4.1% ± 4.6%	P=0.001
	Change in LBM (kg)	3.6 ± 0.5	1.8 ± 0.7	P<0.001
Carrel et al. [95] and Myers et al. [96] GH 1mg/m^2/d (n=35), vs. no treatment (n=19) 1 year	Body fat (%)	38.4 ± 10.7	45.8 ± 8.8	p < 0.01
	Lean mass (kg)	25.6 ± 4.3	21.7 ± 5.0	p < 0.01
	BMI (kg/m^2)	23.7 ± 6.3	25.2 ± 8.9	n/s
de Lind van Wijngaarden et al. [93] Festen et al. [94] (infants) 1mg/m^2 rhGH (n=19) vs. no treatment (n=19); 1 year median (IQR)	BMI (kg/m^2)	16.3 (15.7 to 18.2)	16.4 (15.4 to 19.8)	nr
	BMI (SDS)	0.3 (-0.1 to 1.6)	0.3 (-0.6 to 1.6)	0.72
de Lind van Wijngaarden et al. [93] Festen et al. [94] (children)* 1mg/m^2 rhGH vs. no treatment; 2 year median (IQR)	BMI (kg/m^2)	17.5 (16.1 to 21.1)	19.1 (17.8 to 20.8)	
	BMI (SDS)	1.1 (-0.2 to 1.7)	1.4 (1.1 to 1.6)	0.19
	Fat % (SDS)	1.9 (0.7 to 2.3)	2.4 (2.1 to 2.7)	P<0.001
	Fat (SDS)	1.1 (0.6 to 2.0)	4.5 (0.9 to 2.0)	P<0.01
	LBM_{age} (SDS)	-0.1 (-1.3 to 0.6)	-2.5 (-3.8 to -1.4)	P<0.001
	LBM_{HtSDS}	-1.9 (-2.4 to -1.4)	-2.3 (-2.7 to -1.3)	P<0.05
Festen et al. [91] 1mg/m2/day rhGH (n=10) vs. no treatment (n=10) 2 years median (IQR)	BMI (kg/m^2)	16.3 (15.8 – 19.0)	18.5 (17.5-20.6)	P<0.05
	BMI SDS	0.4 (-0.3 to 1.1)	1.2 (0.9-1.5)	P<0.05
	LBM SDS	-1.2 (-1.7 to -1.1)	-2.8 (-3. to 1.9)	nr
	Percent fat SDS	1.7 (0.9 to 1.9)	2.1 (1.9 to 2.4)	nr
Festen et al. [92] rhGH 1mg/m^2/day (n=15) vs. no treatment (n=14), 1 year median (IQR)	BMI (kg/m^2)	16.4 (15.2 – 18.5)	15.5 (14.9-17.6)	nr
	BMI SDS	0.3 (-0.9 – 1.8)	-0.4 (-0.8-1.3)	nr
	Body fat (%)	22.5 (11.3 – 33.2)	22.8 (19.5-32.9)	nr
	LBM (%)	74.8 (63.7 – 82.3)	73.6 (61.6-75.9)	nr
Haqq et al. [102] rhGH 0.043	BMI (kg/m^2)	31.2 ± 8.9	32.8 ± 9.7	P<0.05

mg/kg.d (n=12) vs. placebo (n=12), 6 months	BMI (SDS)	2.4 ± 0.5	2.5 ± 0.6	nr
	Body fat (%)	49.7 ± 5.8	54.1 ± 5.6	P<0.05
	Fat mass (kg)	26.1 ± 12.8	29.1 ± 14.1	P<0.05
	Lean mass (kg)	24.1 ± 8.8	22.4 ± 8.5	P<0.05
Lindgren et al.[101] and Lindgren et al.[100] 0.1 IU/kg/day rhGH (n=15) vs. no treatment (n=12)	BMI (SDS)	2.0 (-2.4 -6.7)	2.5 (0.1-6.1)	nr
	Body fat (%)	30.9± 11.4	38.2± 9.1	nr

* n= unclear for many of these outcomes

Four trials reported that patients treated with rhGH had statistically significantly higher lean body mass [93,95,96,102] or a larger improvement in lean body mass than untreated patients.[21] In the trial reported by Carrel and colleagues,[21] treated patients' lean body mass increased by 1.8kg more than the improvement seen in the untreated group (3.6 vs. 1.8kg; p<0.001). Treated patients in the other two studies had approximately 2kg[102] or 4kg[95,96] more lean body mass than their untreated counterparts (p<0.05 and p<0.01, respectively). De Lind van Wijngaarden and colleagues[93] reported that change in trunk LBM was statistically significantly better for treated than for untreated infants (1.7 vs. 0.7, respectively). For children, they reported SDS for LBM adjusted for age and height, as well as change in trunk LBM. All of these outcomes were statistically significantly better for treated children than for untreated children after both one and two years of treatment.

Six of the studies reported BMI, with mixed results. Festen and colleagues reported a BMI of 16.1 at year one for treated patients and 18.5 for untreated patients (p<0.05) with similar results at year 2.[91] Haqq and colleagues also reported a statistically significant difference of 1.6 in BMI (31.2 vs. 32.8 for treatment phase vs. placebo phase in a small cross-over RCT; p<0.05). By contrast, the RCTs reported by Carrel[95] and Myers[96] and by de Lind van Wijngaarden[93] found no statistically significant difference between treated and untreated patients. Neither of the other RCTs which reported BMI gave a value for between-group statistical significance, and both treated and untreated patients had similar values.[92,100,101]

There was no statistically significant difference in bone mineral density between treated and untreated patients in the study reported by Carrel and colleagues.[21] No statistically significant differences in progression of scoliosis or onset of scoliosis in either infants or children were reported by de Lind van Wijngaarden.[93]

3.5.4 Biochemical and metabolic markers

The included studies reported a range of biochemical and metabolic markers, and key results are included in Table 18 – please see Appendix 4 for further outcomes. For conciseness, only the key outcomes of IGF-1, IGFBP-3, insulin and glucose are discussed in the narrative summary below.

All of the RCTs reported IGF-1 values or IGF-1 SDS as an outcome measure, and found that levels were higher in rhGH treated patients than in untreated children. Three studies reported that IGF-1 values were statistically significantly higher in rhGH treated patients than in untreated patients.[21,95,96,102] Three studies reported that IGF-1 SDS values were statistically significantly higher in treated than in untreated patients.[91-94]

The included studies had well matched patient groups, whose baseline characteristics were similar in the treated and untreated groups. The only exception was the cross-over study by Haqq and colleagues,[102] which presented baseline characteristics for the study population as a whole, and the study by Lindgren and colleagues[100,101] which reported slightly lower baseline GV SDS in the rhGH group (-1.9 ± 2.0, range -6.4 to -0.9 vs. -0.1 (SD not reported) range -1.7 to -2.71).

Table 18 Biochemical and metabolic markers for PWS studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Carrel et al. [21] 1mg/m^2/day rhGH (n= 15) vs. no treatment (n=14), 1 year	IGF-1 ng/mL	231 ± 98	51 ± 28	P<0.001
Carrel et al. [95] and Myers et al. [96] GH 1mg/m^2/d (n=35), vs. no treatment (n=19) 1 year	IGF-1 (ng/mL)	522 ± 127	121 ± 52	p < 0.01
	IGFBP-3 (mg/L)	3.5 ± 0.73	2.07 ± 0.45	p < 0.01
de Lind van Wijngaarden et al. [93] Festen et al. [94] (infants) 1mg/m^2 rhGH (n=19) vs. no treatment (n=19); 1 year median (IQR)	IGF-I (ng/ml)	179.0 (119.5 to 241.0) (n=12)	33.0 (22.5 to 47.8) (n=15)	nr
	IGF-I SDS	2.5 (1.4 to 2.9)	-2.6 (-4.1 to -0.7)	<0.0001
	IGFBP-3 (ng/ml)	2.2 (1.6 to 2.4) (n=12)	0.9 (0.7 to 1.3) (n=12)	nr
	IGFBP-3 SDS	0.5 (0.0 to 1.2) (n=12)	-2.4 (-3.5 to -1.2) (n=12)	nr
de Lind van Wijngaarden et al. [93] Festen et al. [94] (children)* 1mg/m^2 rhGH vs. no treatment; 2 year median (IQR)	IGF-I (ng/ml)	424.0 (313.0 to 570.0) (n=20)	92.0 (61.8 to 130.0) (n=16)	nr
	IGF-I SDS	2.4 (2.1 to 2.8) (n=20)	-1.6 (-2.5 to -1.0) (n=16)	<0.0001
	IGFBP-3 (ng/ml)	2.8 (2.6 to 3.2) (n=20)	1.5 (1.2 to 1.8) (n=16)	nr
	IGFBP-3 SDS	0.6 (0.3 to 1.1) (n=20)	-1.7 (-2.3 to -1.2) (n=16)	P<0.001
Festen et al. [91] 1mg/m2/day rhGH (n=10) vs. no treatment (n=10) 2 years median (IQR)	IGF-1 SDS year 2	2.3 (2.1-2.9)	-2.0 (-2.7 to 1.0)	P<0.001
	IGFBP-3 SDS year 2	0.6 (0.4-1.1)	-1.8 (-2.7 to -1.5)	P<0.001
Festen et al. [92] rhGH 1mg/m^2/day (n=15) vs. no	IGF-1 SDS	1.7 (0.1 – 2.5)	-2.6 (-4.1 to -0.4)	p<0.001

treatment (n=14), 1 year median (IQR)	IGFBP-3 SDS	0.4 (-0.3 to 1.1)	-3.1 (-4.0 to -2.2)	P<0.05
Haqq et al. [102] **rhGH** 0.043 mg/kg.d (n=12) vs. placebo (n=12), 6 months	IGF-1 (ng/ml)	720 ± 379	232 ± 182	P<0.001
	IGFBP-3 (ng/ml)	6029 ±1311	4247 ± 1209	P<0.01
Lindgren et al.[101] **and Lindgren et al.**[100] 0.1 IU/kg/day rhGH (n=15) vs. no treatment (n=12)	IGF-1 SDS	1.8 (-0.1 -4.1)	-1.4 (-2.9 to -0.3)	nr

Three of the RCTs reported IGFBP-3 values,[93,95,96] and these were higher in treated patients than in untreated patients. In the trial reported by Carrel[95] and Myers,[96] patients treated with rhGH had a mean level of 3.5 mg/ml compared with 2.07 in the untreated patients (p<0.01). Haqq and colleagues reported mean values of 6029 ng/ml in the treated patients and 4247 ng/ml in the untreated patients (p<0.01).[102] Treated children and infants in the study reported by de Lind van Wijngaarden and colleagues[93] had higher IGFBP-33 values than untreated children, although no p values were reported for between group comparisons.

The three studies which reported IGFBP-3 SDS found positive values in the treated children, with SDS of 0.4[92,93] and 0.5 (year 1) or 0.6 (year 2).[91,93] In comparison, untreated patients' median scores were between -2.4[91,93] and -3.1[92] in year one, and -1.7[93] to -1.8[91] in year two. Differences between treated and untreated patients were statistically significant in all three studies (P<0.05[92]; P<0.001[93]; P<0.001[91]).

The RCT reported by Carrel and colleagues[21] reported that there was no statistically significant difference in fasting insulin levels between the treated and untreated infants in their study (5.6 vs. 5.7 μIu/mL, respectively). Two other studies[91,95,96] reported slightly higher insulin levels in treated patients, but did not report p values. The study by Haqq and colleagues[102] reported very similar levels in both treated and untreated patients. Glucose levels appeared to be similar in both treated and untreated patients in the two studies which presented this as an outcome, but neither study reported any p values.[91,102]

3.5.5 Quality of life

None of the included studies reported a measure of health-related QoL.

3.5.6 Adverse events

None of the studies reported AE in any detail. Neither the study reported by de Lind van Wijngaarden and colleagues[93] and by Festen and colleagues[94] nor the one reported by Festen and colleagues[91] reported on AE at all. In the other study by Festen and colleagues,[92] the paper stated that rhGH treatment did not induce disadvantageous effects on carbohydrate metabolism, sleep-related breathing

disorders or thyroid hormone levels. Hauffa and colleagues[99] reported that one patient in the rhGH group developed pseudotumour cerebri after increasing the starting dose to the final dose, but their symptoms resolved on discontinuation. No abnormalities of glucose regulation were observed in either group. None of the patients in the study reported by Carrel and others[95,96] experienced pseudotumour cerebri. Two of their patients who received rhGH experienced headaches within the first 3 weeks, but these resolved with temporary stoppage and gradual re-institution of treatment.

Carrel and colleagues commented that there was no evidence of changes in the prevalence of scoliosis with rhGH treatment,[21] although another paper reporting the same study reported that there was progression of scoliosis in one patient.[97] Lindgren and colleagues[100,101] and Haqq and colleagues[102] reported that there was no severe progression of scoliosis (angle $\geq 20^{o}$) during their RCTs.

Lindgren and colleagues[100,101] noted that one child in their study developed low levels of thyroxine without any change in TSH levels. He received substitution with L-thyroxine during the rhGH treatment. Carrel and colleagues[21] commented that no child in their RCT required thyroid hormone therapy. Haqq and colleagues[102] reported that only one patient required thyroid hormone replacement while receiving rhGH treatment.

3.5.7 Summary

- The evidence for the clinical effectiveness of HGH as a treatment for PWS comes from eight small RCTs (one cross-over trial and 7 parallel group trials), reported in 13 publications. The included studies were generally poorly reported and only two[91,96] presented results on an ITT basis.
- Only one of the studies reported changes in height. Infants who received rhGH for a year grew an average of 6.2 cm more than those in the untreated group ($p<0.001$).[21] Two studies reported a statistically significant difference in HtSDS between treated and untreated patients. The difference was 1 SDS (favouring rhGH treatment) in one study, [95,96] and >2 (year 2) in the other.[93]
- Treated patients grew 3cm/year faster than untreated patients in one RCT[102] and 5cm/year faster in another.[95,96] Another study reported a positive growth velocity SDS for treated patients and a negative one for untreated children (5.5 vs. -2.3).[99] The differences between groups were statistically significant in all three studies.
- Two of the included studies reported bone age as an outcome measure, and this was similar in both treatment groups.[95,96,100,101]
- Four trials reported a statistically significantly lower percentage of body fat (between 1%[93] and 10%[21] lower) in patients treated with rhGH compared with no treatment or placebo.

- Three trials reported that patients treated with rhGH had statistically significantly higher lean body mass[95,96,102] or a larger improvement in lean body mass than untreated patients.[21] One study reported that LBM SDS was significantly better in treated than in untreated children.[93]
- Two studies found that BMI was statistically significantly lower in treated patients than in untreated patients.[91,102] However, another RCT [95,96] found no statistically significant difference between the two groups, and three more studies did not report a p value for between-group statistical significance.[92,93,100,101]
- IGF-1 values were statistically significantly higher in patients treated with rhGH than in untreated patients in three studies.
- Two RCTs reported IGFBP-3 values that were statistically significantly higher in treated patients than in untreated patients.[95,96,102] Three studies[91-93] reported positive IGFBP-3 SDS values in treated patients and negative values in untreated children; differences between the groups were statistically significant.
- Four of the studies reported insulin levels, with varying results. One study[21] reported that there was no statistically significant difference between treated and untreated infants. Insulin levels in another study[95,96] appeared to be considerably higher in treated patients than in untreated patients. Another study[91] reported higher insulin levels in treated patients at year one but lower levels than in untreated patients at year two. Similar values in both groups were also reported.[102]
- None of the included studies reported a measure of health-related QoL.
- None of the studies reported AE in any detail.

3.6 Chronic Renal Insufficiency

3.6.1 Quantity and quality of research available

Six RCTs of patients with CRI met the inclusion criteria for this review,[103-108] and their key characteristics are shown in Table 19 – further details are shown in Appendix 4. The inclusion criteria for this systematic review specified that children should be prepubertal. Five of the studies stated in their inclusion criteria that patients should be prepubertal/ Tanner stage 1, but one study included both prepubertal and pubertal patients.[107] However, we have included outcome measures from this study where data were presented separately for prepubertal children and pubertal children.

The included RCTs were of different designs (two cross-over and four parallel-group). Three of the parallel-group RCTs were open label, with the comparator groups receiving no treatment,[103,106,107] and one was placebo-controlled.[108] The two cross-over studies[104,105] had placebo and treatment phases. There does not appear to have been a wash-out phase in either of the cross-over trials, so a carry-over effect could have affected results. The doses all appeared to correspond to those specified in the

marketing authorisation, but dosages were reported differently, with some using IU and others mg, and some using doses based on weight and others surface area. Randomised treatment duration was six months in the two cross-over trials,[104,105] two years in one study[108] and 12 months in the other studies.

Three of the studies investigated rhGH treatment in children who had received a kidney transplant at least one year before starting the study[103,105,107] and the other three studied children who had CRI.[104,106,108] There was considerable variation in the age of children in the included studies, ranging from 5.6[106] to 12.6[107] years old. Two of the studies were relatively large (n=203[107] and n=125[108]), one was of medium size (n=69[106]), and the remaining three were rather small (n=23[103], n=20[104] and n=11[105]).

Only one study[107] specified a primary outcome. Broyer and colleagues[107] designed their study to test glomerular filtration rate, with growth velocity and HtSDS being used as secondary outcomes. The other studies reported various outcomes relating to growth, body composition and biochemical/metabolic markers, but did not specify which were primary outcomes. Only Sanchez and colleagues[103] mentioned a power calculation, and this appears to have been based on bone formation rates in a previous study so it is not clear what the primary outcome was for the included study. The lack of clarity around primary outcomes and power calculations, together with the small size of three of the studies[103-105] suggests that the trials may have been underpowered to detect differences in outcomes relating to growth and body composition.

The included studies had well matched patient groups, whose baseline characteristics were similar in the treated and untreated groups.

Table 19 Characteristics of CRI studies

Reference	Intervention	Control group	Total randomised and withdrawals	Duration of randomised treatment
Broyer et al., 1996 [107]	rhGH 1 IU/kg/wk n=106 Mean ±SD age (yrs): 12.6 ± 3.4	no treatment n=97 Mean ±SD age (yrs): 12.1 ± 3.1	Total n=203 Sample attrition: n=49	1 year
Fine et al., 2004 [108]	rhGH 0.05 mg/kg/d n=82 Mean ± SD age (yrs): 6.0 ± 3.9	Placebo n=43 Mean ± SD age (yrs): 5.7 ± 3.6	Total n=125 Sample attrition: rhGH: 26 placebo: 15	2 years
Hokken-Koelega et al., 1991 [104]	4 IU/m^2/d rhGH, then placebo n=8	Placebo, then 4 IU/m^2/d rhGH n=8 Median (range)	Total n=20 Sample attrition: n=4	6 months in each arm

	Median (range) age (yr): 8.7 (4.4 to 11.3)	age (yr): 8.6 (4.4 to 16.0)		
Hokken-Koelega et al., 1996 [105]	4 IU/m^2 rhGH / placebo daily s.c.i. n=6 Median (range) age (yr): 12.1 (9.1 to 18.7)	placebo / 4 IU/m^2 rhGH daily s.c.i. n=5 Median (range) age (yr): 11.1 (8.3 to 14.9)	Total n=11 No withdrawals	6 months in each arm
Powell et al., 1997 [106]	0.05 mg/kg/d rhGH n=30 Mean age (yrs) ± SD: 5.6 ± 2.0	no treatment n=14 Mean age (yrs) ± SD: 5.7 ± 2.6	Total: n=69 Sample attrition: 20 withdrew; 4 rhGH pts and 1 control pt excluded from analyses	1 year
Sanchez et al, 2002 [103]	0.05 mg/kg/d rhGH n=12 Mean age (± SD) 9.7 ± 4.5	no treatment n=11 Mean age (± SD) 11 ± 1.8	Total: n=23 Sample attrition: rhGH: n=1 control: n=1	12 months

Table 20 Quality assessment of CRI studies

	Broyer et al. [107]	Fine et al. [108]	Hokken-Koelega et al. [104]	Hokken-Koelega et al. [105]	Powell et al. [106]	Sanchez et al [103]
1. Was the assignment to the treatment groups really random?	un	un	un	un	un	un
2. Was the treatment allocation concealed?	un	un	un	un	un	un
3. Were the groups similar at baseline in terms of prognostic factors?	rep	rep	rep	rep	rep	rep
4. Were the eligibility criteria specified?	ad	ad	ad	ad	ad	ad
5. Were outcome assessors blinded to the treatment allocation?	un	un	un	un	un	par
6. Was the care provider blinded?	in	un	un	un	in	in
7. Was the patient blinded?	in	ad	ad	ad	in	in
8. Were the point estimates and measure of variability presented for the primary outcome measure?	ad	ad	ad	ad	ad	ad
9. Did the analyses include an ITT analysis?	in	in	in	ad	in	in
10. Were withdrawals and dropouts completely described?	ad	ad	ad	ad	ad	ad

Un=unknown; ad=adequate; rep=reported; not rep = not reported; in=inadequate; par=partial

None of the included RCTs provided clear information on method of randomisation or concealment of allocation (Table 20), so it is not possible to say whether or not selection bias may have affected these

studies. The studies all reported eligibility criteria, and presented baseline characteristics which indicated that groups (within trials) were similar at the start of the studies.

The studies gave little information on whether or not outcome assessors were blinded to patients' treatment groups, although Sanchez and colleagues did comment that skeletal radiographs were reviewed by a single observer who had no information about patients' clinical condition or treatment status.[103] In addition, three of the trials gave patients in the comparator group no treatment, so it would have been clear to patients and their care providers whether or not they were receiving rhGH. In three trials, patients in the comparator group had placebo injections. It is not clear whether or not their care providers were also blinded to treatment group. Lack of blinding could have led to performance bias in measuring treatment effect, but the objective nature of outcomes such as height change and growth velocity would have protected against bias to a certain degree.

All the studies presented results as mean values with standard deviations or standard errors to give a measure of variability. The studies all provided adequate details of any patients who withdrew from the study, but only one study[105] presented results on an ITT basis (no patients withdrew from this study). Attrition bias could therefore have affected the results of the non-ITT studies, i.e. if there had been unbalanced and selective withdrawal from different treatment groups within a study, or if particular patients were more likely to withdraw or be excluded from the analysis.

There was a statistically significant difference between treated and untreated children's birth length SDS in one study,[111] but baseline height was the same in both groups. The very small study by de Zegher and colleagues[113] reported slightly lower baseline growth velocity in treated compared with untreated children (5.1 (range 4.0-6.8) vs. 6.4 (range 5.3 – 7.5) cm/yr, respectively). Otherwise, the studies' treatment groups were generally comparable at baseline, with no discernible differences between treated and untreated patients.

The outcome measures for the included studies are shown in Table 21 to Table 23 below. P values in the tables refer to between-group differences.

3.6.2 Growth outcomes

Key growth outcome measures are shown in Table 21 – please see Appendix 4 for other outcome measures. Only one of the included studies reported height gain. Powell and colleagues found that treated children grew an average of 3.6 cm more than their untreated counterparts after a year of treatment (9.1 cm vs. 5.5 cm, $p<0.0001$). All children in the study by Broyer and colleagues experienced an improvement in HtSDS, but this was statistically significantly higher in the children treated with rhGH than in the untreated children (0.6 vs. 0.1; $p<0.0001$).[107] RhGH-treated children in

the study by Powell and colleagues had a statistically significantly higher HtSDS at end of 12 months than untreated children (0.8 vs. 0.0; p<0.0001).[106]

Table 21 Growth outcomes for CRI studies

Study	Outcomes (mean± SD)	rhGH	Control	P Value
Broyer et al.[107] 1 IU/kg/week rhGH (n=30) vs. no treatment (n=28) 1 year	Change in HtSDS	+0.6 ± 0.3	+0.1 ± 0.3	P<0.0001
	Change in growth velocity (cm/yr)	3.7 ±1.6	0.3± 1.6	P<0.0001
Fine et al.[108] rhGH 0.05 mg/kg/day (n=82) vs. placebo (n=43) 2 years	HtSDS	–1.6	–2.9	nr
	GV (cm/yr)	7.8 ± 2.1 (n = 55)	5.5 ± 1.9 (n = 27)	p< 0.00005
Powell et al. [106] 0.05 mg/kg/day rhGH (n=30) vs. no treatment (n=14) 1 year	Height gain (cm)	9.1 ± 2.8	5.5 ±1.9	p < .0001
	Height SDS change from baseline	0.8 ± 0.5	0.0 ± 0.3	P<0.0001
Sanchez et al.[103] 0.05 mg/kg rhGH (n=12) vs. no treatment (n=11) 1 year	Height SDS	-1.1 ± 1.0	nr	nr
	Annual growth velocity (cm/yr)	8.0 ± 2.1	4.8 ± 1.7	P<0.01
Hokken-Koelega et al.[104] 1: 4 IU/m^2rhGH then placebo (n=8) 2: placebo then 4 IU/m^2rhGH (n=8) 6 mths each arm	growth velocity (cm/6mo)	1: 5.2 (1.2) 2: 4.4 (1.6)	1: 1.5 (0.4) 2: 2.4 (1.0)	p<0.0001
	HV SDS	1: 6.9 (2.4) 2: 5.0 (4.5)	1: -3.0 (1.6) 2: -0.5 (3.2)	p<0.0001
Hokken-Koelega et al.[105] 1: 4 IU/m^2rhGH then placebo (n=6) 2: placebo then 4 IU/m^2rhGH (n=5) 6 mths each arm	growth velocity (cm/6mo)	1: 5.3 (1.0) 2: 3.9 (1.3)	1: 1.5 (0.9) 2: 1.9 (0.7)	p<0.0001
	HV SDS	1: 9.1 (2.9) 2: 5.3 (4.0)	1: -1.3 (2.9) 2: -0.4 (1.7)	p<0.0001

One of the six studies reported change in growth velocity, and this was statistically significantly faster in treated than in untreated children.[107] Four studies reported growth velocity at end of treatment, all reporting statistically significantly faster growth in children who received rhGH treatment than in untreated children.[103-105,107,108] The two-year study by Fine and colleagues reported that rhGH-treated patients' growth velocity in the first year was 4.2 cm/year faster than the untreated patients' (p<0.00005). The difference between the two groups was less in the second year (2.3 cm/year faster in rhGH-treated children) but the difference between groups was still statistically significant (p<0.00005) when comparing the difference in change from baseline in those patients who completed two years of the study.[108] A statistically significant difference in growth velocity between groups of just over 3cm/year was reported by both Broyer and colleagues (3.4 cm/year difference, p<0.0001) and by Sanchez and colleagues (3.2 cm/year difference, p<0.01).

The two cross-over studies by Hokken-Koelega and colleagues also reported statistically significantly faster growth velocities in patients during the rhGH phase compared with the placebo phase, with an average of 2.9cm/6 months difference in velocity.[104,105] In the study of children with CRI, patients who received rhGH followed by placebo grew at an average velocity of 5.2cm/6 months during treatment compared with 1.5cm/6 months in the placebo phase. Patients who received placebo followed by rhGH grew 2.4cm/6months during the placebo phase compared with 4.4cm/6months in the treatment phase. The overall mean effect of rhGH was statistically significant ($p<0.0001$). Statistical tests showed that there was no significant carry-over effect (-0.04cm/6 months, $p=0.94$). The cross-over study in children who had received a renal transplant had similar results. Patients grew on average 3.8cm/6months faster during the active treatment phase in the group who received rhGH followed by placebo, and 2cm/6months faster in the active treatment phase for patients who received placebo followed by rhGH ($p<0.0001$ for overall effect of rhGH vs. placebo).[105] Hokken-Koelega and colleagues reported that there was no significant carry-over effect (0.5cm/6months, $p=0.30$).

The two cross-over trials,[104,105] but none of the parallel group RCTs, reported GVSDS. Both trials reported positive SDS values during the active treatment phases and negative scores during the placebo phases. The reported difference in scores between active treatment and placebo phases in the trial of children with CRF was 7.7 ($p<0.0001$)[104] and in the trial of children who had received a renal transplant the difference was 8.0 ($p<0.0001$).[105]

Bone age was reported by five of the six studies. The studies by Powell and colleagues[106] and Sanchez and colleagues[103] reported that there was no statistically significant difference in bone age between the treated and untreated patients. The two cross-over studies by Hokken-Koelega and colleagues reported small differences with slightly lower mean ages for rhGH overall compared with placebo (mean differences -0.01 years[104] and -0.5 years[105]), but did not present any p values for these comparisons. Fine and colleagues[108] reported that the change in bone age between baseline and two years was greater in patients treated with rhGH than in untreated patients for those who completed both years of the study (2.3 vs. 1.6 years; $p=0.0001$).

3.6.3 Body composition

Measures of body composition were reported by three of the studies, and selected outcomes are shown in Table 22.[103,106,108] Other outcomes are tabulated in the data extraction forms in Appendix 4. Children treated with rhGH gained statistically significantly more weight than those in the control groups in the studies reported by Fine and colleagues[108] (2.1kg more in two years, $p=0.0004$) and by Powell and colleagues[106] (1.3kg more in one year, $p=0.007$). However, there was no statistically significant difference between groups in change in weight for HtSDS. Sanchez and colleagues did not report actual weight gain, but reported a statistically significant difference in change in SDS for

weight that favoured treatment with rhGH (0.2 vs. -0.3, p<0.01). Although Powell and colleagues reported a statistically significantly greater weight gain in treated patients, the weight for HtSDS was the same for both groups (0.4, p=0.8703).

Table 22 Body composition outcomes for CRI studies

Study	Outcomes (mean± SD)	GH	Control	P Value
Fine et al.[108] rhGH 0.05 mg/kg/day (n=82) vs. placebo (n=43)	Weight gain after 2 years (kg)	6.7 ± 2.2	4.6 ± 2.7	p = 0.0004
Powell et al. [106] 0.05 mg/kg/day rhGH (n=30) vs. no treatment (n=14)	Weight gain (kg)	3.5 ± 1.5	2.2 ± 1.0	p = 0.007
	Change in weight for HtSDS	0.4 ± 0.7	0.4 ± 0.5	P=0.8703
Sanchez et al.[103] 0.05 mg/kg rhGH (n=12) vs. no treatment (n=11)	Change in SDS for weight	0.2 ± 0.3	-0.3 ± 0.3	P<0.01

3.6.4 Biochemical markers

The included studies reported a range of biochemical and metabolic markers, and these are included in Table 23. For conciseness, only the key outcomes of IGF-1, IGFBP-3, insulin and glucose are discussed in the narrative summary below. In addition, the studies reported a range of markers related to liver function. These are not reported in Table 23 or discussed in the narrative summary below, but are included in the data extraction forms in Appendix 4. No data from Sanchez and colleagues are included in Table 23 as their results focussed on liver function and they did not report IGF, insulin or glucose.

Table 23 Biochemical and metabolic markers from CRI studies

Study	Outcomes (mean ± SD)	GH	Control	P Value
Fine et al.[108] rhGH 0.05 mg/kg/day (n=82) vs. placebo (n=43)	IGF-I (μg/L)	244 ± 128 (n=47)	135 ± 80 (n=20)	P=0.0001
Powell et al. [106] 0.05 mg/kg/day rhGH (n=30) vs. no treatment (n=14)	IGF-I SDS change from baseline	0.2 ± 1.0	nr	P<0.006
	IGFBP-3 SDS change from baseline	4.0 ± 3.2	nr	P<0.011
Hokken-Koelega et al.[104] 1: 4 IU/m^{2}rhGH then placebo (n=8) 2: placebo then 4 IU/m^{2}rhGH (n=8)	IGF-I ng/ml	1: 264 ± 168 2: 268 ± 120	1: 160 (104) 2: 160 (95)	nr
	IGF-I SDS for bone age	1: 2.6 ± 2.0 2: 2.9 ± 2.0	1: -0.2 ± 1.5 2: 0.3 ± 1.6	P<0.0001
	IGFBP-3 ng/ml	1:7708 ± 2323 2: 8706 ± 2275	1: 6102 ± 1892 2: 6501 ± 1988	nr
	IGFBP-3 SDS for bone age	1: 5.0 ± 1.3 2: 5.2 ± 1.4	1: 3.7 ± 1.3 2: 3.9 ± 1.4	p<0.0001

Hokken-Koelega et al.[105] 1: 4 IU/m^2rhGH then placebo (n=6) 2: placebo then 4 IU/m^2rhGH (n=5)	IGF-I ng/ml	1: 594 ± 180 2: 488 ± 237	1: 240 ± 143 2: 321 ± 94	nr
	IGF-I SDS for bone age	1: 5.4 ± 2.8 2: 3.4 ± 0.5	1: 1.0 ± 2.5 2: 6.4 ±1.9	p<0.0001
	IGFBP-3 ng/ml	1: 7457 ± 2088 2: 8495 ± 2921	1: 5681 ± 1588 2: 6228 ± 2193	nr
	IGFBP-3 SDS for bone age	1: 4.5 ± 1.5 2: 3.9 ± 1.5	1: 3.7 ± 2.9 2: 5.3 ±1.5	nr

Four studies reported IGF-1 as an outcome measure,[104-106,108] and levels were higher in treated patients than in untreated patients. IGF-1 values were statistically significantly higher in treated patients at both years one and two in the study by Fine and colleagues (p=0.0004 and p=0.0001, respectively), but only approximately half of the randomised patients were included in this analysis. Powell and colleagues also reported that IGF-1 and IGF-1 SDS values were statistically significantly higher for treated patients than untreated patients (p<0.006).[106] The two cross-over studies by Hokken-Koelega and colleagues reported that IGF-1 SDS for bone age was statistically significantly higher for treated than for untreated patients (2.7 higher in treated children with CRF[104] and 3.7 higher in treated children who were post-transplant,[105] p<0.0001 for both).

Three studies reported IGFBP values,[104-106] and in all three IGFBP-3 was higher in the treated patients. Powell and colleagues reported that IGFBP-3 and corresponding SDS values were statistically significantly higher in treated patients than in untreated patients (p<0.011). Hokken-Koelega and colleagues[104] reported that the IGFBP-3 SDS for bone age, was statistically significantly higher for treated patients (p<0.0001).

Fine and colleagues[108] reported that fasting insulin levels were statistically significantly higher in rhGH patients than in untreated patients after 2 years (p=0.03). Similarly, Hokken-Koelega and colleagues[105] reported slightly higher insulin values in treated children, but did not present p values.

3.6.5 Quality of life

Five of the included studies did not report QoL as an outcome measure. One study[107] reported QoL but did not present data for prepubertal patients (the licensed patients) separately from pubertal patients, so it is not discussed here.

3.6.6 Adverse events

Hokken-Koelega and colleagues[105] reported that no patients in their study had an acute rejection episode, and that there were no serious AE. Sanchez and colleagues[103] reported that two patients with normal rates of bone formation experienced acute rejection episodes after 3 and 12 months of rhGH therapy. One of these episodes was associated with non-compliance to immunosuppressive

medications and both reversed after treatment with methylprednisolone. There were no rejection episodes in untreated patients.

Fine and colleagues[108] reported that there were no differences between groups in year 1. In the second year, eight of 55 rhGH patients experienced asthma or wheezing, but all episodes were preceded by upper respiratory tract infections. Fine and colleagues reported that there were no clinically significant side effects associated with rhGH treatment. Hokken-Koelega and colleagues [104] reported that serum alkaline phosphate was significantly increased during rhGH treatment, but returned to pre-treatment levels when rhGH therapy was replaced by placebo ($p<0.0001$). There was no significant change in parathyroid hormone concentration during either treatment schedule, and thyroid function was reported to have been normal. Broyer and colleagues[107] did not present AE separately for prepubertal and pubertal children, so no data are reported here. Powell and colleagues did not report AE from their study.[106]

3.6.7 Summary

- The evidence for the clinical effectiveness of rhGH as a treatment for short stature owing to CRI comes from six RCTs, two of which were cross-over trials. The trials were generally poorly reported, and only one[105] presented ITT results. Three of the studies had fewer than 25 participants, which suggests that the trials may have been underpowered to detect differences in outcomes relating to growth and body composition.
- One study reported that rhGH treated patients grew an average of 3.6 cm more than their untreated counterparts after a year of treatment. Two studies reported that HtSDS was statistically significantly better in treated children than in untreated children.
- Five studies reported that change in growth velocity or growth velocity SDS was statistically significantly faster for children who received rhGH treatment than for untreated children, with between-group differences in velocity ranging from 3.2cm/year[103] to 4.2 cm/year[108] in the parallel-group trials.
- Two studies reported that there was no statistically different difference in bone age between the treated and untreated patients. Two reported small differences with slightly lower mean ages for rhGH overall compared with placebo, but did not present any p values for these comparisons. One study reported that the change in bone age between baseline and two years was greater in patients treated with rhGH than in untreated patients for those who completed both years of the study.
- IGF-1 levels were statistically significantly higher in treated patients than in untreated patients in two of the four studies which reported this outcome.
- Three studies reported that IGFBP-3 values were higher in the treated patients. Only one of these reported that differences between groups were statistically significant.

- Insulin levels were statistically significantly higher in children receiving rhGH than in those receiving placebo injections or no treatment.
- Four studies presented data on AE. Two rhGH-treated patients in one study experienced acute rejection episodes (one associated with non-compliance to immunosuppressive medications) but both reversed after treatment with methylprednisolone. There were no serious AE reported.

3.7 Children born short for gestational age (SGA)

3.7.1 Quantity and quality of research available

In the UK, rhGH is licensed for use in children born SGA who are over four years of age, have a current HtSDS of <2.5, with a parental adjusted HtSDS -1, had a birth weight and/or length SDS of <-2, and have failed to show catch up growth during the previous year (HV SDS <0). No RCTs meeting these criteria were identified. Following discussion with NICE, the criteria were amended in order to include evidence from RCTs on rhGH. As discussed in Section 3.1.4, the following amended criteria were agreed: growth disturbance (current height <-2.5, no reference to parental height), birth weight and/ or length <-2 SD and failure to show catch up growth (no stated criteria) by the age of three.

Six studies[109-114] met the amended inclusion criteria for this review, and their key characteristics are shown in Table 24 - please see Appendix 4 for further details. In the UK, the licensed dose of rhGH for SGA children is 0.035mg/kg/day, which equates to 0.105 IU/kg/day. Only the study by Phillip and colleagues[114] included a treatment arm with the licensed dose; the other studies all used approximately two or three times the UK licensed dose.

Table 24 Characteristics of SGA studies

Reference	Intervention	Control group	Total randomised and withdrawals	Duration of randomised treatment
Phillip et al. 2009[114]	1. rhGH 0.033mg/kg/d (n=51) mean age (± SD): 5.5 ± 1.5 2. rhGH 0.1mg/kg/d (n=51) mean age (± SD): 5.5 ± 1.4	No treatment (n=47) mean age (± SD): 5.6 ± 1.4	Total n=151 Sample attrition: 2	1 year
Carel et al., 2003 [111]	rhGH: 0.2 IU/kg/d N=112 Mean age (±	No treatment n=56 Mean age (±	Total n=168 Sample attrition: For treatment:	Until adult height reached (mean= 2.7 ±

	SD):12.7 ± 1.4	SD): 12.8 ± 1.6	rhGH: n=21 control: n=23 For analysis: rhGH: n=10 control: n=9	0.6 yrs)
De Schepper et al., 2007 [109]	High dose rhGH: 66 ± 3 µg/kg/d N=11 Mean age (± SD): 5.1 ± 1.6	no treatment n=14 Mean age (± SD): 5.1 ± 1.4	Total n=40 Sample attrition: n=15	2 years
de Zegher et al., 1996 [112]	1. rhGH 0.2 IU/kg/d n=20 2. rhGH 0.3 IU/kg/d n=21 mean age (± SD): 1. 5.4 ± 0.5 2. 5.1 ± 0.4	no treatment n=13 mean age (± SD): 4.9 ± 0.5	Total: n=54 Sample attrition: rhGH 1: n=2 rhGH 2: n=1 control: n=1	2 years
de Zegher et al., 2002 [113]	High dose rhGH 100 µg/kg/d n=9 mean age (range): 6.3 (4.0-8.0)	No treatment n=4 mean age (range): 4.7 (2.3 -6.3)	Total n=13 Sample attrition: Not reported	2 years
Lagrou et al., 2008 [110]	rhGH 0.066mg/kg/d N=20 mean age (± SD): 5.5 ± 1.6	no treatment n=20 mean age (± SD): 5.1 ± 1.3	Total n=40 Sample attrition: 1	2 years

Licensed dose = 35mcg/kg/day = 0.035mcg/kg/day = 0.105 IU/kg/day

Treatment duration was comparable across five of the six included studies. Four of the trials stated a treatment duration of two years.[109,110,112,113] Carel and colleagues[111] administered growth hormone for an average of 2.7 ± 0.6 years, until the participants reached adult height. The children in the study by Phillip and colleagues[114] received treatment for two years, but only the first year allowed a randomised comparison between growth hormone and no treatment.

The mean age of participants was similar both across groups within studies and across five of the six trials included.[109,110,112-114] The mean ages of groups in these trials ranged from 4.7 (2.3-6.3)[113] to 6.3 (4.0-8.0) years. The Carel study[111] included older children with mean ages of 12.7 ± 1.4 in the rhGH and 12.8 ± 1.6 in the control group.

Table 25: Quality assessment of included SGA studies

	Carel et al.[111]	De Schepper et al.[109]	de Zegher et al.[112]	de Zegher et al.[113]	Lagrou et al.[110]	Phillip et al.[114]
1. Was the assignment to the treatment groups really random?	Un	Un	Un	Un	Un	Ad
2. Was the treatment allocation concealed?	In	Un	Un	Un	Un	Un
3. Were the groups similar at baseline in terms of prognostic factors?	Rep	Rep	Rep	Rep	Rep	Rep
4. Were the eligibility criteria specified?	Ad	Ad	Ad	Ad	Ad	Ad
5. Were outcome assessors blinded to the treatment allocation?	Un	Un	Par	Un	Un	Pa\r
6. Was the care provider blinded?	In	Un	Un	Un	Un	In
7. Was the patient blinded?	In	In	In	In	In	In
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Ad	Ad	Ad	Ad	Ad	Ad
9. Did the analyses include an ITT analysis?	In	In	In	Ad	In	In
10. Were withdrawals and dropouts completely described?	Ad	Ad	Ad	Ad	In	In

Un=unknown; ad=adequate; rep=reported; not rep = not reported; in=inadequate; par=partial

The six included trials were generally of poor methodological quality (Table 25).

Phillip and colleagues reported that a centralised computer-controlled system was used to randomly assign children to groups. In the other five trials it was unclear whether the assignment to treatment groups was really random. This was reflected in the assessment of whether treatment allocation was concealed, with one exception being the study by Carel and colleagues,[111] which reported that group assignment was not masked and this was therefore judged to be inadequate.

The blinding of outcome assessors can defend against bias affecting the measurement of some outcomes. In two trials[112,114] outcome assessors for bone age were blinded to chronological age and treatment allocation. It was not stated whether this extended to assessors of other outcomes. In the remaining four trials it was not stated whether the outcome assessors were blinded.

Performance bias, where knowledge of treatment can potentially lead to differences in care provided can be protected against by blinding care givers and patients. The care provider was not blinded to treatment in the studies by Carel and colleagues[111] or Phillip and colleagues,[114] and in the four remaining trials this was unknown. In each of the six trials blinding of the patient was inadequate as

no placebo was used. Only one of the trials conducted an intention-to-treat analysis.[113] This guards against bias arising where for example only the results of patients who did not experience AE or compliance issues are included in the analysis.

3.7.2 Growth outcomes

All six studies[109-114] reported growth outcomes, and these are presented in Table 26.

Table 26 Growth outcomes for SGA studies

Study	Outcomes (mean± SD)	rhGH	Control	P Value
Phillip et al. [114] 1: rhGH 0.033mg/kg/day (n=51) 2: rhGH 0.1mg/kg/day (n=51) vs. untreated (n=47) 1 year	HtSDS	1. -2.3 ± 0.6 2. -1.8 ± 0.8	-3.0 ± 0.6	nr
	Change in HtSDS	1. 0.8 ± 0.3 2. 1.4 ± 0.4	0.1 ± 0.3	nr
	Additional height gained[b] (cm)	1. 3.3 ± 0.2, 95% CI 2.9-3.7 2. 6.5 ± 0.2, 95% CI 6.0-6.9	n/a	nr
Carel et al.[111] 0.2 IU/kg·d (n=91) vs. untreated (n=33)	AH total height gain (cm)	26 ± 7	22 ± 6	0.005
	End of treatment: HtSDS	-2.1 ± 1.0	nr	nr
	AH HtSDS	-2.1 ± 1.0	-2.7 ± 1.0	0.005
	AH total height gain SDS	1.1 ± 0.9	0.5 ± 0.8	nr
	AH difference from target HtSDS	-0.9 ± 1.2	-1.7 ± 1.2	0.005
De Schepper et al.[109] High dose rhGH (n=11) vs. untreated (n=14); 2 years	HtSDS year 2	-1.7 ± 0.7	-3 ± 1	<0.0001
de Zegher et al.[112] 1: rhGH 0.2 IU/kg/day (n=20) 2: rhGH 0.3 IU/kg/day (n=19) vs. untreated (n=13) 2 years	Gain in HtSDS	1: 2.1 ± 0.1 2: 2.5 ± 0.1	0.2 ± 0.1	<0.001 [a]
	Gain in HtSDS for bone age	1: 1.0 ± 0.2 2: 1.2 ± 0.4	0.0 ± 0.3	<0.05 [a]
	GV (cm/yr)	1: 10.2 ± 0.2 2: 11.0 ± 0.4	5.7 ± 0.3	<0.001
	GV SDS	1: 4.3 ± 0.3 2: 5.2 ± 0.4	-0.9 ± 0.3	<0.001 [a]
de Zegher et al.[113] High dose rhGH (100 µg/kg/d) (n=9) vs. no treatment (n=4), 2 years	HtSDS	-1.8 (-3.9 to -0.5)	-3.0 (-3.3 to -2.5)	nr
	GV (cm/yr)	8.5 (6.3 to 10.2)	5.6 (4.4 to 6.8)	nr
Lagrou et al.[110] rhGH 0.066mg/kg·day (n=20) vs. untreated (n=19)	HtSDS	-1.9 ± 0.7	-3.1 ± 0.9	<0.001

[a] untreated vs. treated; [b]compared with untreated controls.

Carel and colleagues[111] reported a mean gain in adult height of 26 ± 7cm in their treated group compared with 22 ± 6cm in their untreated group (p=0.005). They also reported AH SDS, which was statistically significantly higher in the rhGH treated group (-2.1 ± 1.0) compared with the untreated group (-2.7 ± 1.0), p=0.005. Similarly, the SDS for AH total gain was statistically significantly higher in treated patients compared with untreated patients (1.1 ± 0.9 vs. 0.5 ± 0.8; p=0.002). Carel and colleagues[111] also reported the difference from target HtSDS. This was statistically significantly lower in the group receiving growth hormone, compared with the control group (-0.9 ± 1.2 vs. -1.7 ± 1.2; p = 0.005).

Children who received the licensed dose of 0.033mg/kg/d for one year in the study by Phillips and colleagues[114] gained an average of 3.3 ± 0.2cm in height compared with children in the untreated control group. Those receiving the higher dose of 0.1 mg/kg/d rhGH gained an average of 6.5 ± 0.2cm compared with untreated children. No p values were presented for between group comparisons.

De Zegher and colleagues[112] found that gain in HtSDS at the end of the study was higher in the group receiving a higher dose (2.1 ± 0.1 (0.2 IU/kg/day) vs. 2.5 ± 0.1 1 (0.3 IU/kg/day) vs. 0.2 ± 0.1 (untreated), p< 0.001 treated vs. untreated groups). The other study by De Zegher and colleagues[113] reported higher HtSDS in treated patients, but did not present p values.

Phillips and colleagues[114] found that HtSDS was higher in the two rhGH treated groups than in the untreated groups (-2.3 ± 0.6, -1.8 ± 0.8 and -3.0 ± 0.6 for the 0.033mg/kg/d (licensed dose), 0.1mg/kg/d and untreated groups, respectively). These scores reflected a change of 0.8 and 1.4 in SDS for the licensed dose and high dose groups respectively, compared with a change of only 0.1 in the untreated patients' mean SDS value.

Three[109,110,113] of the included studies which used higher doses of rhGH reported that HtSDS was higher in the treated groups than in the untreated groups. De Schepper and colleagues[109] and de Zegher and colleagues[113] reported HtSDS at the end of the first and second years of treatment. In each of these studies, at both time points, the SDS was higher in the treated group, and this difference between groups increased in the second year. In De Schepper and colleagues'[109] study at the end of year one, HtSDS in the treated group was -2.1 ± 0.7 vs. -3.1 ± 1 in the untreated group (p<0.0001). In year two, HtSDS in the treated group was -1.7 ± 0.7 compared with 3.1± 1 in the untreated group (p<0.0001). At the end of two years' treatment, the treated group in the Lagrou[110] study had a statistically significantly higher mean HtSDS (-1.9 ± 0.7) compared with the untreated group (-3.1 ± 0.9), p<0.001.

Two studies[109,110] were suitable for meta analysis of the HtSDS outcome because they were sufficiently homogeneous in terms of dose, duration of treatment, and the children's mean age at start of treatment. However, both trials were small (≤ 20 girls in each treatment group), which affects the validity of tests for heterogeneity, and both used twice the licensed dose, so a meta-analysis of these was considered unlikely to add to the evidence base.

GV (cm/year) was greater at the end of year two in the groups receiving rhGH, in the two studies that presented results for this outcome.[112,113] de Zegher and colleagues 1996[112] found an increased GV in their group receiving a higher dose of growth hormone, and a greater GV for their treated participants overall: 10.2 ± 0.2 (0.2 IU/kg/day) vs. 11.0 ± 0.4 (0.3 IU/kg/day) vs. 5.7 ± 0.3 (untreated), $p<0.001$ untreated vs. treated. The de Zegher 1996 study[112] also found that GV SDS was statistically significantly higher at the end of treatment in the treated groups (4.3 ± 0.3 (0.2 IU/kg/day) and. 5.2 ± 0.4 (0.3 IU/kg/day)) compared with -0.9 ± 0.3 in the untreated group ($p<0.001$ for untreated vs. treated groups).

De Zegher and colleagues 1996[112] reported bone age. The gain in bone age (years) was statistically significantly greater in the groups receiving growth hormone than in those who were untreated. The 0.2 IU/kg/day rhGH group had a mean gain of 1.35 ± 0.16, compared with 1.33 ± 0.24 in the 0.3 IU/kg/day rhGH group and 0.84 ± 0.07 in the untreated group ($p<0.001$ treated vs. untreated groups). This is reflected in the gain in HtSDS for bone age: 1.0 ± 0.2 (0.2 IU/kg/day) vs. 1.2 ± 0.4 (0.3 IU/kg/day) vs. 0.0 ± 0.3 $p<0.05$, treated vs. untreated groups.

3.7.3 Body composition outcomes

Four of the included studies reported body composition outcomes.[109,110,112,113] These results are shown in Table 27. It should be noted that all of these studies used higher doses of rhGH than the UK licensed dose.

Table 27: Body composition outcomes for SGA studies

Study	Outcomes (mean± SD)	rhGH	Control	P Value
De Schepper et al.[109] High dose rhGH (n=11)* vs. untreated (n=14) 2 years	WtSDS	-1.8 ± 1	-3.4 ± 1.6	<0.0001
	Lean mass (kg)	15.5 ± 3.4	12.2 ± 2.5	<0.0001
	Fat mass (kg)	2.9 ± 1	3.1 ± 1.1	Ns
	Lean mass (%)	82 ± 3	77 ± 5	<0.05
	Fat mass (%)	15 ± 2	20 ± 5	<0.05
de Zegher et al.[112] 1: rhGH 0.2 IU/kg/day (n=20) 2: rhGH 0.3 IU/kg/day (n=19) vs. untreated (n=13) 2 years	Weight gain (kg)	1: 6.9 ± 0.6 2: 7.8 ± 0.5	3.6 ± 0.4	<0.001 [a]
	Gain in WtSDS	1: 1.3 ± 0.1 2: 1.8 ± 0.1	0.4 ± 0.1	<0.001[a]
de Zegher et al.[113] High dose rhGH (100 µg/kg/d) (n=9) vs. no treatment (n=4), 2 years	WtSDS (mean and range)	-2.1 (-3.6 to -0.9)	-3.8 (-4.8 to -3.2)	Nr
	BMI SDS (mean and range)	-1.2 (-3.4 to -0.4)	-2.1 (-2.9 to -1.4)	nr
Lagrou et al.[110] rhGH 0.066mg/kg·day (n=20) vs. untreated (n=19)	WtSDS	-2.3 ± 1.2	-3.7 ± 1.5	<0.01
	BMI (SDS)	-1.5 ± 1.1	-2.0 ± 1.5	ns

[a] untreated vs. treated

De Schepper and colleagues reported a WtSDS for treated patients that was almost half that for untreated patients (-1.8 vs. -3.4; p<0.0001). Lagrou and colleagues[110] found that WtSDS at the end of year two was statistically significantly higher in their treated group (-2.3 ± 1.2) than in their untreated group (-3.7 ± 1.5; p<0.01). Similar values were reported by de Zegher and colleagues,[113] although no p values were given.

De Zegher and colleagues 1996[112] also reported gain in WtSDS and weight gain (kg). For both of these outcomes the difference was statistically significant and higher in the groups treated with growth hormone. Mean weight gain (kg) was 6.9 ± 0.6 (0.2 IU/kg/day) vs. 7.8 ± 0.5 (0.3 IU/kg/day) vs. 3.6 ± 0.4 in the untreated group (p<0.001 treated vs. untreated groups). This pattern was reflected in the gain in WtSDS, which was 1.3 ± 0.1 in the 0.2 IU/kg/day group, 1.8 ± 0.1 in the 0.3 IU/kg/day group and 0.4 ± 0.1 in the untreated group (p<0.001 untreated vs. treated groups).

Lean mass and fat mass were reported in kilograms and as a percentage by De Schepper and colleagues.[109] Lean mass (kg) increased from year one to year two in both groups, and was greater in the group receiving growth hormone at both times (13.2 ± 3.4 vs. 10.9 ± 2.4 and 15.5 ± 3.4 vs. 12.2 ± 2.5 for years one and two, respectively). The p value was reported as p<0.0001, but it is unclear at

which time point this p value refers to. Lean mass (%) remained virtually unchanged from year one to year two, but was higher in the rhGH group (82 ± 3 vs. 77 ± 5 at year two). The difference between treated and untreated groups was statistically significant ($p<0.05$), but it is unclear whether this refers to the year one or year two data.

The difference in fat mass (%) between the two groups was statistically significant: 15 ± 2 vs. 20 ± 5, $p<0.05$. Two studies reported BMI SDS.[110,113] One of these reported that there was no statistically significant difference between treated and untreated children[110], and the other reported similar values but gave no p value.[113]

3.7.4 Biochemical markers

Two of the included studies, both of which used higher doses than the UK licensed dose, reported biochemical markers.[112,114] These results are shown in Table 28.

Table 28: Biochemical markers in SGA studies

Study	Outcomes (mean± SD)	rhGH	Control	P Value
de Zegher et al.[112] 1: rhGH 0.2 IU/kg/day (n=20) 2: rhGH 0.3 IU/kg/day (n=19) vs. untreated (n=13) 2 years	Serum IGF-I (µg/L)	1: 332 ± 29 2: 655 ± 69	168 ± 46	<0.01 untreated vs. group 1
	Serum IGFBP-3 (mg/L)	1: 6.10 ± 0.35 2: 6.50 ± 0.52	4.00 ± 0.58	<0.001 untreated vs. group 1
Phillip et al. [114] 1: rhGH 0.033mg/kg/day (n=51) 2: rhGH 0.1mg/kg/day (n=51) vs. untreated (n=47) 1 year	IGF-I, ng/ml	1. 345.6 ± 177 2. 594.3 ± 221	176 ± 107	nr
	IGF-I SDS	1. 0.9 ± 1.9 2. 3.3 ± 2.1	-0.9 ± 1.2	nr
	IGFBP-3, µg/L	1. 4.8 ± 1.1 2. 6.1 ± 1.4	3.9 ± 1.1	nr

Serum IGF-I levels were statistically significantly higher in rhGH treated groups at the end of treatment. In one study,[112] children receiving 0.2 IU/kg/day rhGH had values of 332 ± 29, compared with 655 ± 69 in the 0.3 IU/kg/day group and 168 ± 46 in the untreated group ($p<0.01$ 0.2 IU/kg/day vs. untreated group) after two years' treatment. Phillip and colleagues reported similar IGF-I values as de Zegher and colleagues[112] after a year's treatment, and in addition reported that IGF-I SDS was higher in rhGH treated patients than in untreated patients. Values were 0.9 ± 1.9 and 3.3 ± 2.1 in the low and high dose groups, respectively and 0.9 ± 1.2 in the untreated group.

Serum IGFBP3 levels were also greater in the groups receiving rhGH. In the one year study[114] values were lowest in untreated patients (3.9 ± 1.1µg/L) and higher in the two rhGH groups (4.8± 1.1 and 6.1± 1.4 for the low and high dose groups, respectively). No p values were reported. At the end of year two in the second study, mean values were 6.10 ± 0.35 in the 0.2 IU/kg/day rhGH group, 6.50 ± 0.52 in the 0.3 IU/kg/day rhGH group and 4.00 ± 0.58 in the untreated group ($p<0.001$ untreated vs. 0.2 IU/kg/day rhGH group).[112]

3.7.5 Quality of life

None of the included studies reported QoL outcomes.

3.7.6 Adverse events

Four of the included studies discussed AE in varying detail.[109,111,112,114]

Carel and colleagues[111] found that 44% of patients reported AE, with 10% of these reporting four or more. It was not stated whether these patients were from the treated or untreated group. The authors described two AE they believed to be causally related to treatment; one slipped capital epiphysis after 1.5 years of treatment and one simple seizure episode 10 minutes after first injection. The authors do not state if these led to withdrawal. Sixteen severe AE in 14 patients were reported. These were not thought by the authors to be related to treatment, and included trauma, psychiatric symptoms, abdominal symptoms, otitis, asthma, variocele, striae and migraine. De Schepper and colleagues[109] stated only that no participants 'had a noteworthy adverse event during the two years of study'. No further details were given.

de Zegher and colleagues 1996[112] reported four serious AE. The authors suggested that these might not be linked to growth hormone, but gave no further details. The authors described two treated children vs. one untreated child hospitalised as a result of viral disease (group/ dose not reported). There was one case of aggravated cutaneous eczema reported in group one (0.2 IU/kg/day). Three treated children (group/ dose not given) reported possible increase in size or number of pigmented nevi. Treatment was not interrupted in any of these cases.

Phillip and colleagues[114] only reported AE for the two year study overall, so it was not possible to compare the treated and untreated children. The majority (349/358) of AE in the study were of mild to moderate severity, the most common events (57%) being childhood infections. Of 16 serious AE reported, three were described as likely to be related to rhGH. Two of these (convulsions and papilloedema) resolved on discontinuation of treatment, and the third (epilepsy) stabilised when treatment was withdrawn.

3.7.7 Summary

- Six[109-113] trials examining the effectiveness of growth hormone in children born SGA met the inclusion criteria for the review. The quality of the included studies was generally poor, and only one employed an ITT analysis.[113] All but one[114] of the trials used higher than licensed doses of rhGH.
- One trial reported total gain in adult height, and found this was approximately 4cm higher in people who had received rhGH. The difference between groups was statistically significant ($p<0.005$).[111] Adult height gain SDS was also statistically significantly higher in people who had received rhGH.[111] However, the study used a dose which was approximately twice the licensed dose, and it was carried out in children with a mean age of 12.7 years at start of treatment. This may limit the generalisability of the trial.
- One study[114] reported that patients who received 0.033mg/kg/d rhGH (the licensed dose) gained an additional 3.3 cm height compared with untreated children, and those who received 0.1 mg/kg/d gained 6.5 cm of additional height after one year's treatment.
- Height SDS was found to be statistically significantly higher in children treated with growth hormone in two studies,[109,110] and higher but with no reported p value in two others.[113,114]
- Growth velocity (cm/yr) was greater in the treated groups at the end of year two in the two studies that reported this outcome,[112,113] but the difference was only reported to be statistically significant in one.[112]
- WtSDS was statistically significantly higher in children treated with rhGH in one[110] of the three studies reporting this outcome.
- Lean mass was reported in one study,[109] and was statistically significantly greater in the treated group. Two studies reported BMI SDS.[110,113] One of these reported that there was no statistically significant difference between treated and untreated children,[110] and the other reported similar values but gave no p value.[113]
- One study[112] reported that serum IGF-I and IGFBP-3 levels were statistically significantly higher in patients treated with rhGH, and another[114] reported similar results but did not present p values.
- Reporting of AE was limited in detail, and only reported by four of the trials.[109,111,112] One trial[111] reported two events in treated children that may have been linked to growth hormone. They did not discuss if these led to discontinuation of the drug. A second trial[109] reported only that there were 'no noteworthy' AE recorded. A third trial[112] reported four serious AE, which were not linked to the study drug. Three of 16 serious AE in another trial[114] were linked with rhGH, and these resolved/stabilised once treatment was discontinued.

3.8 SHOX-D

3.8.1 Quantity and quality of research available

Only one study of SHOX patients met the inclusion criteria for this review,[48] and its key characteristics are shown in Table 29. The two-year multicentre RCT by Blum and colleagues[48] compared a daily injection of 50 µg rhGH with no treatment in 52 pre-pubertal children with confirmed SHOX-D. The manufacturer's recommended dose is 45-50mcg/kg body weight,[80] but since the study did not report mean baseline weight of participants it is not possible to comment on whether or not the study reflects the licensed dose. The study also included a non-randomised rhGH-treated group of patients with Turner syndrome, but this group will not be discussed further in this report.

Table 29 Characteristics of SHOX-D study

Reference	Intervention	Control group	Total randomised and withdrawals	Duration of randomised treatment
Blum et al., 2007 [48]	50 µg/d rhGH n=27 Mean age ± SD (yr): 7.5 ± 2.7	no treatment n=25 Mean age ± SD (yr): 7.3 ± 2.1	Total n=52 Sample attrition: 1	2 years

Table 30 Quality assessment of SHOX-D study

1. Was the assignment to the treatment groups really random?	unknown
2. Was the treatment allocation concealed?	unknown
3. Were the groups similar at baseline in terms of prognostic factors?	reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	partial*
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	inadequate

* blood analyses were carried out at a central facility

The included study was generally poorly reported (Table 30), with little information on method of randomisation or concealment of allocation. Patients in the comparator arm received no treatment, so the patients themselves and their care providers would have been aware of whether or not they were receiving the study drug. The patients in the two groups had similar baseline characteristics, and the trial's inclusion criteria were clearly stated. However, the analysis was not reported on an ITT basis as

one discontinuing patient was excluded from the analysis. The study did not include discussion of sample size or a power calculation, so it is not possible to determine whether or not it was adequately powered to detect a difference in the primary outcome (first year GV).

3.8.2 Growth outcomes

Table 31 shows growth outcomes at the end of two years' treatment. Children treated with rhGH gained approximately 6cm more height than those in the control group ($p<0.001$). Although all children remained below average height, the HtSDS was statistically significantly lower in the untreated group (-3.0 ± 0.2 vs. -2.1 ± 0.2; $p<0.001$). Blum and colleagues also commented that 41% of rhGH treated patients reached a height within the normal range for age and gender (>-2.0 SDS), compared with only one patient in the untreated group.[48] There was no statistically significant difference between the groups in catch up of bone age.

Table 31 Growth outcomes for SHOX-D study

Study	Outcomes (mean ± SD)	rhGH	Control	P Value
Blum et al.[48] 50 µg rhGH (n=27) vs. no treatment (n=24); 2 years	ht gain (cm)	16.4 ± 0.4	10.5 ± 0.4	<0.001
	ht SDS	-2.1 ± 0.2	-3.0 ± 0.2	<0.001
	HV (cm/yr)	7.3 ± 0.2	5.4 ± 0.2	<0.001
	HV SDS	2.3 ± 0.3	-0.4 ± 0.1 (n=22)	<0.001

The difference in GV (1.9 cm/yr) between the two groups during the second year of the study was statistically significant ($p<0.001$). Children in the rhGH group had a positive HV SDS, i.e. their growth velocity was above average for their age group. By comparison, those in the untreated group had a negative score, indicating slower growth than normal for their age group. Again, the difference between the groups was statistically significant ($p<0.001$).

3.8.3 Body composition

The included study did not report body composition as an outcome measure.

3.8.4 Biochemical markers

Blum and colleagues did not report biochemical outcomes in any detail. However, they did state that IGF-I SDS values were in the low-normal range for both groups at baseline but increased to the upper-normal range in the rhGH treated group. In ten (37%) of the rhGH treated children, IGF-I concentrations exceeded +2 SDS at least once during treatment, whereas none of the untreated patients experienced this. Similarly, IGFBP-3 SDS values were close to the normal mean in both groups at baseline, but increased to the upper-normal range in the treated group.

3.8.5 Quality of life

The included study did not report QoL as an outcome measure.

3.8.6 Adverse events

The rate of treatment-emergent AE was higher in the rhGH group than in the no-treatment arm (Table 32), but these were reported to have mostly been common childhood illnesses.

Table 32 Adverse events for SHOX-D study

Study	Outcomes (mean± SD)	rhGH	Control	P Value
Blum et al.[48] 50 µg rhGH (n=27) vs. no treatment (n=24); 2 years	At least 1 treatment-emergent AE	85%	68%	nr
	Arthralgia	3	2	nr
	Increased number of cutaneous nevi	2	0	nr
	Recurrent otitis media	1	1	nr
	Scoliosis	1	0	nr

There were no significant changes in thyroid function reported during the study, and no serious AE occurred in the SHOX-deficient patients.

3.8.7 Summary

- The evidence for the clinical effectiveness of rhGH as a treatment for short stature owing to SHOX-D comes from the single RCT which met the inclusion criteria for this review. The study was unblinded and did not report an ITT analysis.
- By the end of the second year, children treated with rhGH had gained statistically significantly more height than those in the control group (approximately 6cm more), with no statistically significant difference in catch up of bone age. Height SDS was statistically significantly higher in treated than in untreated patients.
- Treatment with rhGH led to a statistically significantly greater growth velocity in both years one and two (3.5cm/yr greater than untreated patients in year one, and 1.9cm/year greater in year two). The HV SDS was positive, i.e. above the average for chronological age, during both years of rhGH treatment whereas untreated children had negative HV SDS.
- Treatment with rhGH raised IGF-I and IGF-BP-3 levels to the upper normal range.
- Treatment of the SHOX-deficient children in this RCT was not associated with any serious AE.

3.9 Transition phase in Growth Hormone Deficiency

The scope for this review requested that, if evidence allows, the assessment report should consider the transition of care from paediatric to adult endocrine services of young people whose linear growth is not complete. Although a number of 'transition phase' studies were assessed for inclusion in the review of clinical effectiveness, these included patients who had completed linear growth. They therefore did not meet the inclusion criteria for this review.

Once a patient's linear growth has ceased, he or she may still not have reached peak bone mass, which would increase the risk of osteoporosis later in life. Continued rhGH treatment in these patients beyond completion of linear growth can be beneficial for improving bone mass. For example,

Conway and colleagues[128] randomised 160 18-25 year olds with severe GHD who had received rhGH during childhood to continued treatment (n=109) or no treatment (n=51). They reported that two years of continued treatment was associated with approximately 3.5% greater increase in bone mineral density of the lumbar spine compared with those who had discontinued treatment.[128]

Continued rhGH treatment can also improve body composition in young adults whose linear growth is complete. Five papers[129-133] were identified that reported changes in body composition, biochemical markers, QoL or AE for this patient group. However, since the patients had completed linear growth they did not meet the inclusion criteria for this review and are therefore beyond the scope of this review.

3.10 Summary of previous systematic reviews

The searches for this systematic review identified three systematic reviews. One of these was the previous HTA report,[5] discussed in Section 3.2.1.1, and another was a Cochrane review related to that work.[134] The third reference was a new systematic review of growth hormone in TS,[135] and this is discussed below.

The new systematic review was conducted in Canada in 2007 by the Canadian Agency for Drugs and Technologies in Health (CADTH).[135] The quality of the systematic review was good. Inclusion and exclusion criteria relating to the primary studies were reported. The review included RCTs or comparative observational studies that compared rhGH with placebo or no treatment, included females with TS, measured growth (final height, interim height, growth velocity), AE and QoL. Those studies which included fewer than 20 patients, or administered rhGH for less than one year were excluded. Jadad and Hailey scales were used in quality assessment, but no further details were reported.

The CADTH included 19 studies, ten of which reported data from six RCTs.[135] Three of the six RCTs included in the CADTH review were excluded from the present systematic review. One was excluded as it was a conference abstract from 1991, another was excluded because its outcome measures did not match our inclusion criteria, and the third was excluded because it did not compare rhGH with a treatment arm that did not contain somatropin.

The CADTH authors judged the RCTs to be of good quality, and the observational studies of fair quality, using the Jadad scale. However, they not describe this in detail in their report. The present systematic review used the CRD quality assessment criteria[83] rather than the Jadad scale. This, along

with the difference in included studies, may explain this discrepancy in judgement of quality between the two reports.

The CADTH systematic review found that growth was accelerated and height increased in girls taking rhGH for TS. There were no serious AE reported in the included studies. The cost effectiveness and cost utility analyses in the CADTH study are discussed in Section 4.1.3. The CADTH study[135] concluded that the evidence suggested that rhGH is effective in improving growth and final height in girls with TS, but found no evidence available to suggest rhGH improves QoL.

4 ASSESSMENT OF COST-EFFECTIVENESS

Introduction

The aim of this section is to assess the cost effectiveness of growth hormone treatment in children with GHD, TS, PWS, CRI, SGA, SHOX-D compared to no treatment. The economic analysis comprises:

- a systematic review of the literature on the cost effectiveness of growth hormone treatment (Section 4.1)
- a review of the health related QoL (HRQoL) of people with GHD, TS, PWS, CRI, SGA, SHOX-D (Section 4.2)
- a review of the manufacturers' submissions to NICE (Section 4.3)
- a de novo SHTAC economic model and cost effectiveness evaluation (Section 4.4).

A previous HTA report has estimated the cost effectiveness of growth hormone treatment.[5] In that report, a cost effectiveness model was constructed that estimated lifetime treatment costs and benefits in terms of cost per cm gained. Those analyses are extended in the present report by including QoL factors in the economic modelling.

4.1 Systematic review of existing cost-effectiveness evidence

4.1.1 Methods for the systematic review of cost-effectiveness

A systematic literature search was undertaken to identify economic evaluations for rhGH in children. The details of the search strategy for the cost effectiveness studies are in Appendix 2. The manufacturers' submissions were reviewed for any additional studies. Titles and abstracts of studies identified by the search strategy were assessed for potential eligibility by two health economists. Full text versions of relevant papers were retrieved and checked by two health economists. Any differences in judgement were resolved through discussion. The quality of the cost effectiveness

studies was assessed using a critical appraisal checklist based on that by Drummond and Jefferson,[136] the ISPOR checklist[137] and the NICE reference case. [138]

4.1.2 Results of the systematic review of cost-effectiveness

A total of 220 potentially relevant studies were identified in the cost effectiveness searches and one in the QoL (QoL) searches. Five full papers were retrieved with only two economic evaluations meeting the inclusion criteria. The characteristics and results of the evaluations are discussed below.

4.1.3 Description of the identified studies

The literature search did not identify any economic evaluations conducted across the entire range of conditions of interest or any for the population of England and Wales. Table 33 provides a summary of the characteristics and base case findings for the two published North American economic evaluations for human growth hormone for children with TS[135] and GHD.[139]

Table 33 Characteristics of economic evaluations of rhGH treatment in children

Author	CADTH [135]	Joshi et al [139]
Publication year	2007	2006
Organisation	Canadian agency for drugs and technologies in health	Novo Nordisk
Country	Canada	USA
Study type	CEA and CUA	CEA and CUA
Study perspective	Canadian health care system	The USA health care payers' perspective
Study population	Female population aged 10 at baseline with TS receiving treatment for 5 years until 15 years old.	i) cohort of 5 years old at baseline with GHD receiving treatment for 11 years until 16 years old. ii) cohort of 3 years old at baseline with GHD receiving treatment for 15 years until 18 years old.
Intervention	rhGH	rhGH (Norditropin)
Model type	Deterministic decision analytic model	Deterministic decision analytic model
Time horizon	Lifetime (assumed to be until age 81)	Lifetime (assumed to be age 78 for males and age 80 for females)*.
Discounting	5% applied to both costs and benefits (QALYs)	3% applied to both costs and benefits (QALYs)
The primary clinical treatment effects modelled/assessed	147.5 cm was the final height in the intervention group 141 cm was the final height in the control group	The "success" of treatment is defined as achieving "normal height", ie final height within 2SD of the gender specific population mean.
Source of clinical evidence for the primary effect	Stephure and colleagues [86]	Not indicated. Appears to be an assumption. The probability of "success" was assumed to be 90% if treatment started at age 3 and continued until age 18. The probability of "success" was assumed to be 75% if treatment started at the age of 5 and continued until age 16
Health benefit	QALY	QALY

outcome		
QoL gain, per year	0.042	0.189
Results	Individuals with rhGH treatment had an additional discounted cost of C$153,593 and an additional discounted benefit of 0.63 QALY. The cost effectiveness was estimated as C$243,078 per QALY gained.	For the cohort of 5-16 years, individuals with rhGH had an additional discounted cost of US$155,005 and an additional discounted benefit of 4.2 QALY. The cost effectiveness was US$36,995 per QALY gained. For the cohort of 3-18 years, cost per QALY was US$42,556.

SD=Standard deviation; QALY=quality adjusted life year
* The authors did not report a gender distribution at baseline and whether all-cause mortality rates were used in the calculations.

The cost effectiveness studies were assessed against the critical appraisal checklist (Table 34). Generally, the CADTH study[135] was of a higher quality; the effectiveness of the treatment had been established through a systematic review, and the estimates for parameter values are more appropriate than the study by Joshi and colleagues.

Table 34 Critical appraisal checklist of economic evaluation

	Item	CADTH [135]	Joshi et al[139]
1	Is there a well defined question?	Yes	Yes
2	Is the patient group in the study similar to those of interest in UK NHS?	Yes	Yes
3	Is the correct comparator used that is routinely used in UK NHS?	Yes	Yes
4	Is the study type and modelling methodology reasonable?	Yes	Yes
5	Is an appropriate perspective used for the analysis?	?	?
6	Is the health care system or setting comparable to UK?	?	?
7	Is the effectiveness of the intervention established based on a systematic review?	Yes	No
8	Is the model structure appropriate and does it fit with the clinical theory of the disease process?	Yes	Yes
9	Are assumptions reasonable and appropriate?	Yes	No
10	Are health benefits measured in QALYs using a standardised and validated generic instrument from a representative sample of the public?	?	No
11	Are the resource costs used reasonable and appropriate for the UK NHS?	Yes	Yes
12	Are the health states and parameters used in the model described clearly and are they reasonable and appropriate for the UK NHS?	Yes	No
13	Is an appropriate discount rate used?	Yes	Yes
14	Has the model been validated appropriately?	?	?
15	Is sensitivity analysis undertaken and presented clearly?	?	?

Yes / No / ? (unclear or partially true)

4.1.4 Modelling approach

Both economic evaluations presented cost effectiveness analyses using simple deterministic decision analytic models. Both assumed that the clinical benefit achieved as a result of the rhGH treatment in the patients' early years will last through their lifetime. Joshi and colleagues[139] assumed that age-adjusted normal height was achieved after the first year of treatment. Subsequently, the benefits in terms of "normal height years" and associated utility gain were assigned from the second year of treatment. Conversely, the CADTH study[135] did not assume that patients experienced any improvement in health related QoL during the treatment. The utility gain is associated with the completion of treatment rather than with achieving normal height, as normal height was not achieved in the review of clinical effectiveness.

The cohorts differed with respect to age at baseline, duration of treatment and probability of achieving normal height at the end of treatment (see Table 33 above). The CADTH study[135] used the characteristics and clinical effectiveness data from the TS RCT[86], whilst Joshi and colleagues [139] did not provide any clinical evidence for either the baseline characteristics of the two cohorts of patients with GHD or the assumed clinical effectiveness estimates.

Joshi and colleagues[139] assumed a 20% dropout rate after 12 months of treatment and related it to the slight pain experienced by patients, although no clinical evidence was presented to support this assumption. The CADTH study[135] did not adjust the final outcomes for the dropout rate, effectively assuming it to be zero. Since none of the TS patients achieved normal height, the CADTH study[135] did not differentiate between partial and complete success of rhGH treatment. In contrast, Joshi and colleagues[139] assumed that those patients who completed treatment but did not achieve normal height still acquire a partial utility gain. However, no justification for this assumption is provided.

Discounting was appropriately applied to costs and benefits in both studies, although the discounting rates were different from the 3.5% recommended by NICE[138] (3% in the study by Joshi and colleagues[139] and 5% in the CADTH study[135]).

4.1.5 Estimation of final outcomes (QALYs)

Both studies highlighted the difficulty of translating intermediate (clinical) outcomes to final outcomes (QALYs). There is an apparent paucity of utility-based estimates of health related QoL in rhGH patients and an absence of such estimates obtained from children eligible for rhGH treatment (section 4.2). Therefore the authors chose alternative utility estimates that, in spite of acknowledged shortcomings, were judged to meet the requirements of their economic models. The utility increment associated with rhGH treatment reported in the two studies ranged from 0.04 [135] to 0.189.[139]

Joshi and colleagues[139] adapted the QoL indexes presented in the Wessex Development and Evaluation Committee (DEC) report.[140] The indexes estimated in the report were not derived using one of the methodologically rigorous techniques for obtaining utility estimates, such as time trade off (TTO) or standard gamble (SG)[141] and cannot therefore be interpreted as "utilities". Furthermore the utility element of that report was a set of scenarios not based on primary or secondary data sources and thus could not be considered reliable or valid.[5] Joshi and colleagues used utility estimates of 0.781 for the pre-treatment and no treatment groups, although this is different to the value 0.884, reported in the DEC report. Those patients who achieved success, i.e. normal height, had a utility of 0.97 applied from the start of the second year of treatment. Patients with partial success were assumed to acquire a partial utility gain defined as 35% less than the full utility gain associated with achieving normal height. The value was stated to be between 0.884 and 0.940.

The CADTH study[135] did not use absolute utility values associated with each health state but applied an incremental utility value of 0.04 for patients receiving treatment with rhGH. The utility increment was estimated from a TTO survey in a small sample of adults with TS[142] (see section 4.2 for details of this study). The patients in the QoL study were asked how many years they would be willing to lose from their life to attain an average stature. The answers were translated into the incremental utility estimate of 0.04. The CADTH study[135] stated that TS patients do not attain an average stature, and so this estimate is likely to be an overestimate and bias the result of economic evaluation in favour of rhGH treatment.

4.1.6 Estimation of costs

Joshi and colleagues[139] included costs for paediatric consultations and rhGH treatment. The CADTH study[135] also included costs for X-ray examination. The unit costs reported in the economic evaluations reflect the difference in clinical practices in Canada and the USA, the price difference of the unit of resources expressed in Canadian and US dollars, and the difference in methodological approach adopted in the two studies. For example, the CADTH study[135] excluded the specialist visits as these do not differ between the intervention and the control groups. The total incremental cost reported varies according to the length of treatment but is consistent between the two studies.

4.1.7 Model results

The cost effectiveness analysis in the CADTH study[135] used an incremental difference of 6.5 cm in final height between the intervention and control groups, based on their clinical review. They calculated the undiscounted cost effectiveness as C$26,529 per centimetre of improved final height and the discounted ICER was C$23,630 per centimetre of improved final height. They estimated an ICER of C$243,078 per QALY gained. The authors concluded that for an average patient with TS,

rhGH treatment is unlikely to be cost effective unless the payer is willing to pay more than C$200,000 to obtain a QALY.

Joshi and colleagues[139] calculated the difference in "normal height years" between the intervention and the control groups to estimate the incremental cost per normal height year. It was assumed that normal height was achieved by patients in the intervention group, but not in the control group. The incremental gain in "normal height years" in the cohort of 5-16 year olds was 17.4 (discounted). The corresponding value in the cohort of 3-18 year olds was 21.1 (discounted), which translated into an incremental cost per additional year of normal height of $8,900 (discounted) in the cohort of 5-16 year olds and an incremental cost per additional year of normal height of $9,300 (discounted) in the cohort of 3-18 year olds. They estimated an ICER of about $37,000 per QALY gained for treating children with GHD from ages 5 to 16 years and an ICER of about $42,600 per QALY gained for treating children with GHD from ages 3 to 18 years. The authors concluded that the cost effectiveness of rhGH compares favourably to accepted threshold values and represents reasonable value for money.

In both studies the deterministic one-way analyses indicated that the results were sensitive to variations in the utility estimate, the starting age of treatment, the duration of treatment and the daily dosage. The results were also sensitive to assumptions about clinical effectiveness[139] and to variations in the price of rhGH.[135]

The two economic evaluations arrived at opposite conclusions about the value for money of the rhGH treatment in children. The economic evaluation conducted for the CADTH study[135] may provide a more reliable estimate of the cost effectiveness as it has used clinical data from a reasonable quality RCT and TTO utility estimates. In contrast, the assumptions about clinical effectiveness of rhGH treatment by Joshi and colleagues[139] did not seem to be supported by clinical evidence. Furthermore they also used indexes, interpreted as utility weights, that do not appear to be reliable or valid.

4.1.8 Summary and conclusion of the systematic review of cost effectiveness studies

We undertook a systematic review of the literature in order to identify existing models in this area. The systematic review of published economic evaluations identified two North American studies relevant to the target population and no studies conducted in the UK. The results of the two identified studies produced two very different estimates of cost effectiveness. This difference is largely due to the choice of utility estimates and assumptions on the effectiveness. As discussed in section 4.2, there is a paucity of reliable estimates of utility gains associated with growth hormone treatment. Therefore the results of both studies should be treated with caution. In particular, Joshi and colleagues[139] adapted QoL indexes that were not derived according to the NICE reference case and could not be considered

reliable or valid.[5] The literature study did not identify studies which we could use for this review and so a de novo independent economic model was required.

4.2 Review of research on Quality of life

4.2.1 Systematic review of Health Related Quality of life studies

A systematic review was undertaken to identify HRQoL studies for rhGH for children. The HRQoL searches were undertaken to populate a lifetime economic model with utilities to calculate QALYs, so studies with adults and children were eligible for inclusion. Titles and abstracts of studies identified by the search strategy were assessed for potential eligibility by two health economists. Full text versions of relevant papers were retrieved and checked by two health economists. Any differences in judgement were resolved through discussion. The details of the search strategy for QoL are in Appendix 2.

The titles and abstract of the studies identified by the search strategy were assessed on the basis of the following criteria:

- Disease condition as defined in Table 2 in Section 3 of this report.
- Primary research using a preference/utility based measure for the conditions interest.
- Primary research using a generic measure (i.e SF-36) that can be translated into a utility-based estimate.
- Primary research using a condition/disease specific QoL measure and an algorithm that allowed disease specific QoL to be converted into utility values.

Exclusion criteria for the systematic literature search

- Primary research reporting QoL that could not be converted into utility values using a validated mapping algorithm.
- Background or discussion papers that do not report a QoL measure for the conditions of interest.
- Papers reported in language other than English.

The search strategy identified 391 articles that were potentially relevant. After the abstracts had been screened, 24 articles were identified and full papers were retrieved for these articles. After checking the retrieved studies, 6 papers met the inclusion criteria. These are summarised in Table 35. A further targeted search linking height to HRQoL is reported in section 4.2.2.

Table 35 Characteristics of included QoL studies

Author	Bannink et al[143]	Bertella et al[144]	Busschbach et al [145]	Carel et al[146]	Koltowska-Haggstrom et al [147]	Sandberg et al [148]
Publication Year	2006	2007	1998	2005	2008	1998
Country	The Netherlands	Italy	The Netherlands	France	England and Wales	USA
Study type	QoL observational cohort study matched to normal population	QoL observational cohort study	QoL observational case-control study	QoL observational cohort study matched to normal population	Estimated utilities from a survey of general population in England and Wales, and mapped to an observational cohort study	QoL observational case-control study
Study population	49 participants with TS.	13 participants with PWS	17 participants with CO Renal failure 25 with TS, 25 with GHD	568 participants with TS	894 participants with CO and AO GHD. CO onset GHD occurred in 21.6%	140 participants with GHD 53 participants with GHD that had siblings
Study population age	19.6 ± 3.0 years (14.8-25.8 years)	27.08 ± 4.55 years (20-33 years)	Between 24 years ± 4.1 (ISS) and 28 years ± 4.9 (TS)	22.6 ± 2.6 years	40 ±16.5 years	26.1 ±6.5 years (18.8-46.9 years)
Comparator population	Dutch general population	No comparator	44 normal short participants (not diagnosed with ISS)	French general population	E&W General population	53 Controls (unaffected siblings)
Intervention(s)	GH treatment was for 7.1 ± 2.7 years.	GH treatment in 5 participants, but had ceased treatment 1 to 4 years before being enrolled in the study.	GHD treated with rhGH during childhood.	GH treatment for 4.8 ± 2.2 years. 72% received oestrogen treatment	GH treatment	Pituitary derived rhGH and recombinant GH. GHD treatment was for 4.5 ± 3.1 years (0.9-14.3)
Included QoL instrument used	SF-36	SF-36	Time Trade Off	SF-36	QoL-AGHDA with utility weights from EQ-5D	SF-36
Time period where HRQoL instruments administered	HRQoL evaluation occurred 2.8 (1.6) years after rhGH discontinuation.	HRQoL evaluation at the beginning, during and after rhGH discontinuation	HRQol evaluation in adulthood after rhGH discontinuation if applicable	HRQoL evaluation occurred 6 years after rhGH discontinuation.	HRQoL evaluation at baseline and last reported visit follow up for 1 to 6 years	After rhGH discontinuation

Methodology of collecting QoL data	The SF-36 was administered after rhGH treatment had been discontinued for at least 6 months and final height had been reached.	The SF-36 was administered at the beginning of the treatment and then again at intervals of 6, 12 and 24 months to patients and parents.	TTO asked the participants the maximum number of years they were willing to give up in order to obtain average stature.	A postal survey including the SF-36 and GHQ-12 sent to participants	Both the EQ-5D and QoL-AGHDA were completed by general population. A regression model was used to estimate utility weights for QoL-AGHDA items in an observational study.	Eligible GHD subjects completed SF-36 questionnaire over the telephone, in addition to same sex siblings.
Results	Women with TS treated with rhGH reported significantly better HRQoL in social functioning, role limitations-emotional and bodily pain domains compared with normal population. Other domains were roughly equal to normal population.	PWS showed significant improvement during rhGH therapy on SF-36 in vitality, physical functioning, general health, social functioning, role limitation because of emotional problems, general mental health and total scale.	The GHD patients were hardly prepared to make a trade off. Participants with TS or renal failure had an estimated reduction in QoL of 2-4%.Women with TS made an average TTO for their infertility of 9%.	HRQoL was not statistically different from the reference values obtained for young French women from the general population.	Qol-AGHDA $_{utility}$ scores were higher in patients with CO than with AO. both at baseline 0.75 (SD 0.173) vs 0.64 and at the last reported visit) 0.82 (SD0.167) vs 0.76. Patients with CO-GHD gained less than AO patients with regard to the total gain 0.18 (SD 0.488) vs 0.35.	The GHD sample had only a significantly lower score from the sibling control group on general health scale ($P<0.05$) The rest of the QoL domains showed not significant difference.

GHD

Three relevant studies were identified that met the inclusion criteria.[145,147,148] Sandberg and colleagues[148] used the SF-36 in participants with GHD. The study reported no baseline data, and only reported SF-36 after rhGH treatment had finished compared with non-GHD siblings and the general population. Therefore, the study was of no value in investigating the gain in HRQoL from rhGH treatment.

The second study by Busschbach and colleagues[145] used the time trade off (TTO) method; a preference based approach that asks people to quantify the numbers of years of life they would be willing to give up to overcome a particular state of health. The participants were asked the number of years they were willing to trade off at the end of their life in order to obtain average stature. The TTO was completed by people with GHD, TS and CRI (see below for TS and CRI). There were 25 adults with isolated GHD included in the study. The sample of GHD men made only a negligible trade off (less than 2%) while the sample of GHD women were willing to make a slightly larger trade off of around 2% of their expected length of life to reach average height. The major drawbacks with this study were the small sample of between 17 and 25 people with each condition of interest, the retrospective design and the lack of a control group. Also it is unlikely that gaining average stature is a realistic possibility for most people with the conditions of interest. Furthermore, for one of the conditions of interest (GHD) the patients had received rhGH treatment, and for another condition (CRI) it was unclear whether they had or had not received rhGH treatment as children. It is likely that any rhGH treatment will underestimate the TTO made to gain average stature, as these participants have already benefited from an increase in extra height. It was decided that this study did not provide a robust enough estimate of preference of health states to be used in the model.

The third study, by Koltowska-Haggstrom and colleagues,[147] mapped EQ-5D values to a disease specific Qol assessment of GHD (QoL-AGHDA) instrument from a survey. This was then used to transform QoL-AGHDA scores from a cohort of patients from the KIMS (Pfizer International Metabolic) database into utility weighted QoL-AGHDA scores ($QoL\text{-}AGHDA_{UTILITY}$). A good response rate of 84% was achieved, and 921 individuals from the general population of England and Wales responded to the survey. A regression model was used to estimate utility weights for QoL-AGHDA ($R^2 = 0.42$). The EQ-5D responses were used as the dependent variable and the QoL-AGHDA responses were used as independent dummy variables with age as a covariate.

The patient cohort from the KIMS database consisted of 894 patients from England and Wales. However, only 21.6% had childhood onset GHD (applicable to the scope). The study was carried out in adults and it is unclear whether the child onset GHD (CO-GHD) group had had prior rhGH treatment. This may undervalue gain in HRQoL if this is the case. An inclusion criterion for the study was no treatment for rhGH for a minimum of 6 months prior to entry. The mean age for the whole cohort was 40 (SD16.5) at diagnosis and 45 (SD 14.3) years old at entry into KIMS. The study reported that CO-GHD patients had a QoL-AGHDA$_{UTILITY}$ value of 0.75 (SD 0.173) at baseline compared to the last reported visit score of 0.82 (SD 0.166). The study reports mean gain in QoL-AGHDA$_{UTILITY}$ per year of 0.05 (SD 0.117). They also reported a total gain of 0.18 (SD 0.488), and it is assumed that this is the QALY gain over the study duration worked out using trapezoid formula compared to the baseline QoL-AGHDA$_{UTILITY}$ values. A last observation carried forward (LOCF) method was used. The average length of follow up in the study for the CO-GHD was not reported and so is not possible to verify the QALY gain or gain per year.

In the combined cohort of adult onset GHD (AO-GHD) (78%) and CO-GHD (22%) the greatest improvement in utility occurred within the first year of rhGH treatment. Subsequently, the QoL improvement is maintained when compared to the general population over a 6 year follow up. It is unclear whether this benefit from rhGH treatment is maintained after treatment has stopped.

The limitations of this study were that it was observational with no control, and that the EQ-5D had not been conducted amongst the participants of the KIMS database. Furthermore, the regression model used to translate EQ-5D scores to disease specific measure explained less than half the sample variation of the EQ-5D values. Nevertheless, the study provided an estimate of utility at baseline and at the last reported visit in one of the conditions of interest. The study's generalisability to the other conditions of interest is unclear and it was felt that any attempt to link utilities in this study to the other conditions of interest was difficult due to the difference in height outcomes.

TS

There were three studies that met the inclusion criteria for people with TS.[143,145,146] Two[143,146] of these were not useful as they only reported SF-36 scores after rhGH treatment had been completed compared to a cohort of women from the general population. Therefore they could not be used to investigate the gain in HRQoL from rhGH treatment. Busschbach and colleagues[142]

used a TTO method (described above) to 25 TS women who had not received rhGH treatment as children. Their average time trade off was small in the region of 4% of their life years to reach an average height for the general population.

PWS

One study met the inclusion criteria.[144] This was potentially useful as it shows the gain in HRQoL from rhGH treatment over a 24 month period. However, the study had several limitations that make its results highly uncertain. It was a small study with only 13 Italian adult PWS participants, of whom 5 had previously undergone rhGH treatment. There was no control group. At the last recorded observation (24 months) there were only 9 participants left in the study. A new study mapping from SF-36 to a UK based EQ-5D preference based utility index has recently been published that provided an algorithm for this to be done.[149]

However the PWS QoL study is for adults who have received rhGH and it is unclear how this relates to the QoL gain for a group of children and whether this QoL benefit would be maintained throughout their lifetime.

CRI

One study was identified that met the inclusion criteria.[142] Busschbach and colleagues used a TTO approach for 17 adults who had childhood onset renal failure. It is unclear whether the participants received any rhGH treatment prior to the TTO assessment. The participants were asked what percentage of the years of their expected life they were willing to trade to reach normal height and to not experience health states involving a kidney transplant and dialysis. The resulting time trade off associated with renal failure was 4% to reach normal height.

SGA

There were no relevant HRQoL studies that were identified that met the inclusion criteria

SHOX-D

There were no relevant HRQoL studies that were identified that met the inclusion criteria.

4.2.2 Height and health related QoL.

The NICE reference case clearly states that the measure of health outcome used in the cost effectiveness analysis should be QALYs calculated with utilities derived from a validated generic, preference based measure of HRQoL.[138] The clinical effectiveness review in Section 3 found no RCTs that reported HRQoL measures as an outcome and the additional search for HRQoL studies (above) only located one relevant study by Koltowska-Haggstrom in one of the conditions of interest (GHD) that was strictly applicable to the NICE reference case.[138]

Therefore, a targeted search was conducted to identify publications that reported gains and losses in utility in relation to variation in height, as height is one of the primary outcome measures of growth hormone treatment. Details of the search are in Appendix 2. One full paper by Christensen and colleagues[150] was identified.

The study used the 2003 Health Survey for England with 14,416 observations for adults (aged >18 years).[151] HRQoL was measured using the EQ-5D with the UK tariff. Height was converted from centimetres to HtSDS using a UK population algorithm. Inter-relationships between variables were assessed using ordinary least squares (OLS) linear regressions, controlling for age, weight and gender. All OLS analyses were controlled for multicollinearity (close interaction between explanatory variables). Where there were any highly correlated variables (weight and BMI) then one variable was omitted from the regression. The regression analyses included two-level categorical variables ('sex', 'limiting long standing illness' and 'social class') to explore the relationship between height and HRQoL while controlling for these confounding factors.[150]

There was a positive correlation between an increase in height and a participant's EQ-5D score. The mean EQ-5D scores were lower in the shorter compared with taller subjects, as well as lower than the overall population mean. The authors' report an ANOVA combined with *post hoc* Tukey HSD test for homogeneous subgroups which showed that the sample could be split into three meaningful subgroups each significantly different ($P<0.05$) from each other in terms of their EQ-5D scores. The first subgroup 'HtSDS$\leq$-2.0' had significantly lower EQ-5D scores compared with the second group '-2.0>HtSDS$\leq$0' and the third group 'HtSDS >0'. The second subgroup had significant lower scores than the third group. A multivariate linear analysis using the previously identified subgroups was undertaken to predict the variation in HRQoL. The full model predicted only one-third of the sample variation in EQ-5D (R^2=0.318, 0.343 and 0.290) based on 11946 observations.[150]

The model predicted that for those people shorter than –2.0 HtSDS, an improvement of 1 HtSDS will result in a change in EQ-5D score of 0.061. However, for the subgroup between -2.0 and 0 HtSDS the gain in EQ-5D is much reduced (a 1 HtSDS improvement only increases EQ-5D score by 0.010). One drawback to the Christensen study is that the population used to elicit QoL values are not from the conditions of interest but from the general population.

4.2.3 Summary and conclusions of the QoL review

The systematic review of QoL identified 6 studies that met the inclusion criteria. None of the studies were in a childhood population. Three studies reported the SF-36 but were not useful on further examination as they only reported SF-36 scores after rhGH treatment. One poor quality study reported SF-36 at baseline, 6 months, 12 months and 24 months for a small cohort of adult PWS participants and the scores from this study were mapped to a UK based EQ-5D preference based utility index by a subsequent study.

There were only two studies that reported change in QoL using preference based measures in the conditions of interest.[145,147] The first study [145] used TTO methodology for people with GHD, TS, and CRI. The number of years they would be willing to trade to reach average height was in the range of 0-4. However there were several limitations to this study and it was felt that it did generally not provide a robust estimate of utility gain from rhGH treatment. The second study[147] used a regression model to give utility weights (based on the EQ-5D from a UK population) to the disease specific QoL-AGHDA. The KIMS database was then used to transform patients QoL-AGHDA values into QoL-AGHDA$_{\text{UTILITY}}$ values. However, it was in an adult population and it is unclear whether they had previously had rhGH treatment as children. This study was specific to GHD patients and is unlikely to be generalisable to the other conditions of interest.

An additional targeted search was undertaken for QoL in relation to height. One study was identified by Christensen and colleagues, which provided utility estimates based on the EQ-5D for different HtSDS from the Health Survey for England for an adult general population. The study provides a common utility gain that could be compared across all the conditions of interest that could be used with the clinical effectiveness outcomes from the RCTs.

Based on the review of the QoL literature, there is likely to be a small gain in utility for individuals receiving growth hormone treatment. However, this is based on a proxy measure of gain in height from shorter people in the general population. This excludes many relevant potential benefits and disadvantages of rhGH treatment that it is not possible to capture without good quality evidence from the conditions of interest. This is especially true for PWS as additional HRQoL gain from improved body composition is unlikely to be captured with this method. Furthermore, there is also uncertainty over the impact of extrapolating back into childhood with adult utility data.

4.3 Review of the manufacturers' submissions

Six of the seven manufacturers submitted evidence to be considered for this review. Five out of the six manufacturers' submissions (MS) consisted of a written report and an electronic model supporting the cost-effectiveness analyses. The sixth MS by Sandoz did not comply with the NICE template for MTA and presented a description of the product (Omnitrope) and what appears to be a cost-minimisation analysis using Genotropin as a comparator (defined as a reference product). The collaborative submission is appraised below and a critique of the Sandoz submission is presented in section 4.3.6.

A de novo economic model has been used by the five collaborating manufacturers involved in the submission to the MTA of rhGH. Under Pfizer's leadership, a common modelling framework was developed and used in the cost-effectiveness analysis of treatment in children with GHD, TS, PWS, CRI, SGA. Each of the collaborating manufacturers presented essentially the same model with some minor modifications, for example changes in the unit price of rhGH. The model developed was based upon the previous HTA report[5] but has been extended to consider longer term outcomes in order to estimate cost effectiveness in terms of QALYs. One manufacturer, Merck Serono, produced their own version of the model and so the health benefits differ slightly to the other models.

The manufacturers' submissions also included a rapid review on QoL that was undertaken by Eli Lilly on behalf of the collaboration of manufacturers. The aim of the main review was to provide a rapid search to identify the key papers that explored the impact of short stature in childhood and the impact of short stature in transition to adulthood and as adults. The overall conclusion from this review highlighted the inconsistent findings relating to the role of short stature in QoL and psychosocial functioning in both childhood and adulthood.

4.3.1 Modelling approach

In the manufacturers' submissions, the base case analyses estimated the incremental cost of rhGH per cm of height gained relative to no treatment (in order to compare with previous HTA report[5]) and the incremental cost of rhGH per QALY gained relative to no treatment. The utility scores used in the model in children with GHD, TS, CRI, and SGA were based upon the study by Christensen and colleagues[150], discussed here in Section 4.1.3. A gain in height was assumed to be associated with QoL improvements, which was assessed using the EQ-5D utility scale. In PWS patients QoL gain was based upon a small study of adult PWS patients, together with an estimation of the benefits associated with a reduced risk of diabetes. The assumptions used to derive QoL utility improvements are discussed in section 4.2.

The economic evaluation of rhGH treatment in GHD, TS, SGA and CRI is based on a single clinical effect of additional height gained as a result of treatment. This clinical effect and many of the other parameters used in the model are estimated from the Kabi International Growth Study (KIGS) database,[152] which is a large scale collaborative database developed by Pfizer for the safety and efficacy of treatment with rhGH. It includes data from more than 60,000 treated patients in over 50 countries for all licensed indications, i.e. GHD, TS, PWS, SGA and CRI. Table 36 shows the input parameters used in the manufacturers' model that have been derived from the KIGS database. The costs used in the manufacturers' model were based upon those used in the previous HTA report and inflated to current prices where appropriate.[5]

Table 36 Input parameters from manufacturers' submission from KIGS database (from Pfizer MS)

Parameter	GHD	TS	PWS	CRI	SGA
Number of patients (start of treatment)	7036	2749	485	806	990
Number of patients (near adult height)	2547	1349	75	157	127
Start age	9.14	9.3	7.42	9	8.18
End age	16.37	16.45	15.21	13.95	14.18
Drop out rate (% at 1 year)	0.04	0.0273	0.02	0.117	0.03
Dose (mg/kg/day) 0-17 years of age	0.03	0.04	0.03	0.04	0.04
Utility: Treated	0.83	0.8	0.76	0.8	0.81
Utility: Untreated	0.69	0.69	0.67	0.69	0.69
Height SDS: Treated	-1.17	-2.24	-1.36	-2.17	-2.01

Height SDS: Untreated	-2.99	-3.18	-2.22	-2.99	-3.23

SDS, standard deviation score

The cost-effectiveness analysis of rhGH treatment in PWS is based on an alternative structure of the model which estimates the utility gain based on a small study of 13 adult PWS patients[144] (see Section 4.2) who received rhGH for two years and a further utility gain for reduced diabetes risk. However the PWS QoL study is for adults who have received rhGH and it is unclear how this relates to the QoL gain for a group of children and whether this QoL benefit would be maintained throughout their lifetime. Furthermore the two methods[153,154] used by the Pfizer submission to translate SF-36 scores into utilities were not based on choice based methods like TTO or SG that produce utilities more rigorously.[141] The model assumes that individuals with PWS and diabetes would have a 10% lower QoL than those without. Based on Pfizer's submission to the Pharmaceutical Benefits Advisory Committee in Australia, it was assumed that the prevalence of diabetes in PWS patients would reduce from 8% to 2% although it was not possible to verify these assumptions in the reference provided

An alternative model structure that allowed for the second clinical effect (a reduction in the risk of osteoporosis) was also presented in a scenario analysis for GHD. In this model it was assumed that a proportion of GHD children continue treatment until they reach the age of 25.

The manufacturers' model makes the following assumptions:

i) patients with conditions of interest have the same life expectancy as the general population of England and Wales in the treated and untreated groups,
ii) patients can continue rhGH treatment or discontinue treatment at the end of one year
iii) untreated children do not gain any utility benefit throughout the course of the lifetime of the model
iv) treatment costs and monitoring costs are applied over the treatment years. Health benefits, as measured by QoL associated with particular attained heights, are maintained over patients' lifetimes. The full utility value is applied after two years of treatment.
v) Compliance is assumed to be 90% in the base case analysis and this was assumed to not impact efficacy.
vi) Adverse events are not considered in the model for both the treated and non-treated patients.

vii) In the base case, for all conditions except PWS, rhGH treatment only affects final height and does not affect the risk of morbidities, such as osteoporosis fracture or diabetes

viii) The MS estimated the average height at the end of treatment for the control group from the previous HTA report.

4.3.2 Appraisal of the manufacturer cost effectiveness analysis

A summary of the manufacturer's submission compared with the NICE reference case requirements[138] is given in Table 37 and indicates that the submission meets most of the requirements. See Appendix 9 for a tabulation of the critical appraisal of the submission against the Drummond and colleagues' checklist.[136]

Table 37 Assessment of manufacturers' submission against NICE reference case requirements

NICE reference case requirements	Included in submission
Decision problem: as per the scope developed by NICE	✓
Comparator: no treatment alternative	✓
Perspective on costs: NHS and PSS	✓
Perspective on outcomes: all health effects on individuals	✓
Type of economic evaluation: cost-effectiveness analysis	✓
Synthesis of evidence on outcomes: based on a systematic review	No evidence synthesis
Measure of health benefits: QALYs	✓
Description of health states for QALY calculations: use of a standardised and validated generic instrument	✓
Method of preference elicitation for health state values: choice based method (e.g. TTO, SG, not rating scale)	✓
Source of preference data: representative sample of the public	✓
Discount rate: 3.5% p.a. for costs and health effects	✓
Notes (✓=yes; × = no; ? = uncertain; N/A=not applicable):	

4.3.3 Cost effectiveness results

The mean daily per patient cost for each of the manufacturer's growth hormone treatments was based upon the unit cost shown in Table 38. Merck Serono stated that there will be a reduced cost of £20.87 through the use of the Merck Serono Easypod™ which they report will reduce vial wastage and increase compliance.

Table 38 Unit cost of rhGH for different manufacturers

Manufacturer / product	Unit cost, £ / mg
Genotropin (Pfizer)	£23.19
Humatrope (Eli Lilly)	£18
NutropinAq (Ipsen)	£20.70
Saizen (Merck Serono)	£23.19
Norditropin SimpleXx (Novo Nordisk)	£21.39

The basecase analyses for Pfizer, Eli Lilly, Ipsen and Merck Serono are shown in Table 39. Merck Serono produced their own version of the model and so the health benefits differ slightly from the other models'.

Table 39 Base case results for Pfizer, Eli Lilly, Ipsen and Merck Serono

		GHD continued*	GHD	TS	PWS	CRI	SGA
	Incremental QALY	3.483	3.483	2.825	2.3	2.526	2.98
	Height gain (cm)	32.24	32.24	7.95	25.59	4.48	21.92
Pfizer	Incremental Cost	£72,003	£61,124	£84,078	£74,849	£40,325	£54,088
	ICER (£/QALY)	£20,673	£17,552	£29,757	£32,540	£15,962	£18,167
	Cost per cm gain	£2,233	£1,896	£10,576	£2,925	£9,001	£2,467
Eli Lilly	Incremental Cost		£57,043	£65,654		£31,574	£42,340
	ICER (£/QALY)		£16,176	£36,237		£12,498	£14,221
	Cost per cm gain		£1,747	£8,258		£7,048	£1,932
Ipsen	Incremental Cost	£65,198	£54,779	£75,243		£36,129	
	ICER (£/QALY)	£18,721	£15,730	£26,630		£14,301	
	Cost per cm gain	£2,022	£1,699	£9,464		£8,065	
Merck Serono†	Incremental Cost	£72,719 *£65,711*		£84,077 *£75,847*		£40,325 *£36,416*	£54,087 *£48,839*
	ICER (£/QALY)	£20,881 *£18,869*		£29,757 *£26,844*		£15,962 *£14,414*	£18,167 *£16,404*
	Cost per cm gain	£2,256 *£2,038*		£10,576 *£9,540*		£9,001 *£8,129*	£2,467 *£2,228*

† Figures in italics for EasyPod device

* GHD continued is the scenario with rhGH treatment during childhood and a transition period

The base case results for the Novo Nordisk model using KIGS data are shown in Table 40. They also reported alternative ICERs using patient level data.

Table 40 Base case results for Novo Nordisk using KIGS database

	GHD continued*	GHD	TS	CRI	SGA
Incremental QALY	3.7	3.7	2.89	2.9	2.77
Height gain (cm)	27.45	27.45	7.95	3.65	5.67
Incremental Cost	£71,264	£58,637	£79,976	£41,388	£51,745
Cost per QALY	£19,276	£15,861	£27,720	£14,254	£18,655
Cost per cm gain	£2,596	£2,136	£10,060	£11,345	£9,123

*GHD continued is the scenario with rhGH treatment during childhood and a transition period

4.3.4 Manufacturers' conclusions

The authors suggested that many of the health benefits associated with rhGH treatment are not quantifiable and cannot be modelled easily. Many of these benefits would improve overall patient QoL and possibly duration of life. These benefits include self esteem, improvements in sleep and concentration and increased appetite as well as increases in lean body mass, total bone mass and increases in muscle strength. These benefits may lead to reduced risk of diabetes, obesity and cardiovascular diseases.

The manufacturers concluded that their economic analyses demonstrated that rhGH is cost effective for the treatment of short children with GHD, CRI and those born SGA and borders on cost effectiveness for the treatment of TS and PWS. They stated that the values for cost per cm compared favourably to those reported in the previous NICE assessment[5] and supported the recommendation of rhGH for children with GHD, TS and CRI, plus its extension to include SGA children.

4.3.5 Summary of general concerns

- Clinical effectiveness estimates for height gain were taken from an observational cohort rather than an RCT. It is not clear whether the subset of the KIGS database chosen was representative of the UK patient population or, for example whether the subset chosen may be more severe.

- For three of the conditions (GHD, PWS and SGA) the estimates of height gain, in cm, were considerably higher than those shown in the trials due to the estimates used for end height in the control group.
- All conditions, except PWS, used mortality rates from the general population. It is likely that individuals with these conditions, in particular CRI, will have increased mortality compared to the general population.
- The manufacturers have used the Christensen study[150] for their HRQoL utility values but have not taken these from the regression analysis from this study. Instead they have used the relationship between EQ-5D and height without controlling for other factors. Utility gain attributed to height is likely to be capturing the combined effects of other (unobserved) variables, such as age, longstanding illness and gender. For example, older generations generally have lower QoL because of their age. Not controlling for other factors, in particular age, results in the overestimation of the utility values. Furthermore the group with the lowest height and QoL (<-3 SDS) had few observations and individuals in this group were generally elderly (mean age > 70 years).
- Treatment cost is calculated by rounding up to the nearest whole year of treatment.
- There is high uncertainty associated with the assumptions and sources used to estimate QoL gain in the PWS model. These were based on a small study of adult PWS patients and it is unclear how this relates to the QoL gain for a group of children and whether this QoL benefit would be maintained throughout their lifetime. The methods used to derive values from the SF-36 for utilities were based on rating scales and therefore did not use choice-based methods like the SG and TTO. QoL gain also estimated utility gain from reduced diabetes prevalence but this evidence could not be verified. There are considerable difficulties extrapolating the benefit from treating children with rhGH to their health benefits as adults.

4.3.6 Sandoz submission to NICE

Sandoz presented an analysis comparing Omnitrope with Genotropin. The MS contained a comparison of the annual cost of treatment with omnitrope and with genotropin in patients with GHD and Turner syndrome. However the MS did not comply with NICE guidance for an MTA,[138] as QALYs were not estimated and a cost-effectiveness analysis was not presented. The MS attempted a cost-minimisation analysis implicitly suggesting that treatment with Omnitrope is equally effective as treatment with Genotropin (in terms of additional height in children with

GHD and TS) but is associated with less cost to the NHS. A critical appraisal of the Sandoz MS is given in Appendix 10.

4.4 SHTAC Independent economic assessment

4.4.1 Overview

A comparison of the costs and benefits of rhGH compared with no treatment, in cohorts of children with GHD, TS, PWS, CRI, SGA and SHOX-D was made using decision analytic models. Models were constructed in Microsoft Excel according to standard modelling methods.[138] To identify data to populate the model, systematic searches were conducted to locate studies on the natural history and epidemiology of the indicated conditions, health related QoL, and costs.

Costs were derived from published studies (where available), and from national and local NHS unit costs. The model was from the perspective of the NHS and Personal Social Services (PSS), since only these direct costs were included. The model estimates the lifelong costs and benefits from rhGH treatment. The costs and benefits were discounted at 3.5%, as recommended by NICE.[138] The base year for the costs was 2008. The intervention effect in terms of improvement in HtSDS was derived from the systematic review of effectiveness reported in Section 3. The outcome of the economic evaluation is reported as cost per QALY gained and cost per cm gained.

4.4.2 Description of the model

A decision analytic model was designed for the economic evaluation of rhGH for treatment of GHD, TS, PWS, CRI, SGA, and SHOX-D and was based upon one developed in the previous HTA report.[5] The current model compares a cohort of patients receiving rhGH during their childhood with a cohort of patients who were not treated with rhGH. The state transition Markov model has a cycle length of one year and a life-time horizon. A Markov model was used as these are suitable for lifetime analyses with few health states.[155] The base case decision analytic model includes health states for alive and dead. The England and Wales population mortality rates are applied in each cycle for patients with an adjustment using the standard mortality rates for each of the conditions.

The model assumes that a daily subcutaneous injection of rhGH is administered for the duration of treatment, unless a patient from the treatment cohort drops out of treatment or dies. The parameters of the model that determine the age at the start of treatment, the duration of treatment

and the annual drop-out rates are estimated from the KIGS database described in the MS or based upon clinical opinion and vary between conditions. A daily dose is calculated according to the child's weight. The dose regimen corresponds to the licensed indication of rhGH in children (and adults, in a scenario analysis of the GHD cohort).

Health care resources included for the cost of patient monitoring apply to both the treatment and no treatment cohorts. The cost categories and unit costs are consistent with the costs used in the previous HTA report for rhGH.[5] The discount rate of 3.5% is applied to both costs and final outcomes.

Patients from the treatment cohort who stay in treatment receive a benefit of an additional height gain relative to patients in the no treatment cohort. Patients who drop out of treatment stop accumulating height gain so their growth progression is no different from the height gain in the no treatment cohort. In each yearly cycle, individual HRQoL is estimated based upon their height gain. Individuals are assumed to maintain the same HRQoL after treatment has stopped for the rest of their lifetime. In each cycle the total costs and QALYs are calculated by multiplying the individual costs and HRQoL by the number of people in the cohort still alive for the treatment and no treatment cohorts. The total lifetime costs and QALYs are calculated for the treated and non treated groups by aggregating the costs and QALYs in each cycle. The total discounted QALY gain, and cost of treatment for the treatment and no treatment cohorts are calculated. Thus the cost-effectiveness of rhGH is calculated,

$$Cost\ effectiveness = \frac{Cost\ \ for\ treatment\ cohort - Cost\ \ for\ no\ treatment\ cohort}{QALYs\ \ for\ treatment\ cohort - QALYs\ \ for\ no\ treatment\ cohort}$$

Parameters used in the model and the data sources used to derive them are described in more detail in Section 4.4.2.2 .

A list of the model assumptions is given below. Assumptions are applied to all conditions unless explicitly stated otherwise. All assumptions were tested in sensitivity analyses.

- The diagnostic costs were not included in the analysis as they were assumed to be the same for both rhGH treated and no treatment patients.

- The base case assumes no drop out or discontinuation of treatment. This was based upon clinical opinion that this was likely to be a relatively rare occurrence. The base case model therefore evaluates just rhGH treatment versus no treatment.
- There are two health states for alive or dead in the model and the transition between them is based on age related mortality data.
- The mortality rates were assumed to be higher than for the England and Wales general population estimates for untreated and treated cohorts for all conditions.
- It was assumed that there would be no reduction in mortality as a result of rhGH treatment. There is a lack of data to assume otherwise.
- The model time horizon is 100 years and all individuals are assumed to die by this age.
- Effectiveness estimates for the conditions were based on selection of the best quality evidence from the clinical effectiveness review in section 3. RCTs were only selected if the follow up length was at least 2 years after the start of treatment. Where there were no appropriate RCTs, long term observational studies were considered. In the case of SGA, the most appropriate RCT was only for one year.
- Compliance was assumed to be 85% in the base case with no loss of efficacy for rhGH treatment.[156]
- An additional scenario was undertaken for the GHD condition where treatment continued for a transition phase into adulthood to age 25. This was only applicable for 34% of the GHD population,[157] No additional benefit, in terms of height gained, was assumed from this additional treatment.
- In the treatment and no treatment cohorts, all children are monitored until they reach adulthood, assumed to be age 17.

4.4.2.1 Evaluation of uncertainty

The evaluation of the cost-effectiveness of growth hormone treatment is based on uncertain information about variables such as clinical effect, health related QoL and resource use. This uncertainty was evaluated using deterministic and probabilistic sensitivity analyses. One-way deterministic sensitivity analyses were conducted to evaluate the influence of individual parameters on the model results and test the robustness of the cost-effectiveness results to variations in the structural assumptions and parameter inputs (section 4.6).

Multi-parameter uncertainty in the model was addressed using probabilistic sensitivity analysis (PSA) (section 4.6.2).[158] In PSA probability distributions are assigned to the point estimates used in the base case analysis. The model is run for 1000 iterations, with a different set of parameter values for each iteration, by sampling parameter values at random from their probability distributions. The uncertainty surrounding the cost-effectiveness of the growth hormone treatment is represented on a cost-effectiveness acceptability curve (CEAC) according to the probability that the intervention will be cost effective at a particular willingness to pay threshold. Appendix 12 reports the parameters included in the PSA, the form of distribution used for sampling each parameter, and the upper and lower limits assumed for each variable.

4.4.2.2 Model validation

The SHTAC model was validated by checking the model structure, calculations and data inputs for technical correctness. The completed cost-effectiveness model was verified by another health economist. The SHTAC model was checked for internal consistency against the MS economic models by running the SHTAC model with the inputs used in MS models to ensure similar results. The robustness of the model to changes in input values was tested using sensitivity analyses to ensure that any changes to the input values produced changes to the results of the expected direction and magnitude. Finally, the model results were compared with those from previous studies including the previous HTA report and this is discussed in more detail in Section 6.

4.4.3 Data sources

4.4.3.1 Life expectancy

Several studies have attempted to assess the mortality rate of adults with the conditions of interest. Nielsen and colleagues[159] conducted a meta-analysis to assess overall standard mortality rate (SMR) for men and women with benign pituitary disease. Six studies were included in the meta-analysis of sex-specific mortality. Studies (total 5412 patients) reported SMR for men of 2.06 (CI 1.94-2.2) and women 2.8 (CI 2.59-3.02). However these analyses were for hypopituitarism rather than GHD.

Shoemaker and colleagues[160] followed up 3439 women in the UK diagnosed with TS between 1959 and 2002 to the end of 2006. Mortality in women with TS is three times higher than in the

general population, is raised for almost all major causes of death, and is raised at all ages. SMR was 3.9 in women aged 15 to 44 years old and 2.6 in women age 45 to 84 years.

Population-based morbidity and mortality data for PWS are not available except from regional cross-sectional surveys.[161] A recent regional survey in England indicates high morbidity and mortality rates. Lifetime mortality rates were roughly three times higher than the general population. Within these studies the data is insufficient to construct survival curves.

Mortality and causes of death in treatment for children with end-stage renal disease was estimated in a Dutch cohort study between 1972 and 1992.[162] Of all 381 patients, 85 had died. The standardized mortality rate (SMR) was 31.0 over this period and 21 in the last cohort between 1992 and 2002.

Kajantie and colleagues [163] studied the relationship between small size at birth and all-cause and non-cardiovascular mortality in 13,830 individuals born between 1924 and 1944 in Helsinki, Finland. They found that small size at birth is associated with increased all-cause mortality at all ages among adult women but only with premature death in adult men.

We were unable to find any information on mortality rates for SHOX-D.

Using UK life tables, we estimated the life expectancy of adults with these conditions using the standard mortality rates described above. Normal adult life expectancy was estimated to be 75 years for men and 79 years for women. Life expectancy for patients with hypopituitarism was reduced to 68 years for men and 70 years for females. Life expectancy with TS was reduced to 70 years for females. We estimated the life expectancy with CRI to be reduced to 35 years for men and 42 years for females, using the end stage renal disease mortality rates as a proxy in the absence of any available data for CRI. This may underestimate life expectancy as not all CRI patients will go on to develop end stage renal disease.

In the base case model, we assume that for all conditions the life expectancy is lower than the general UK population and investigate general population life expectancy in sensitivity analyses.

4.4.3.2 Effectiveness data

The start and end age of treatment, and the duration of treatment are shown in Table 41. For GHD, CRI, PWS and SGA there are no RCTs with a duration of more than 3 years, so we used data from the KIGS database.[152] SHOX-D was not included in the KIGS database and so we assumed that these children start treatment at the same age as those in the Blum RCT[48] and continue treatment for the same duration as for children with TS in the KIGS database. For the purposes of the model we rounded the start age and treatment duration.

Table 41 Input parameters used in the SHTAC model

	GHD	TS	PWS	CRI	SGA	SHOX-D
Source	KIGS [152]	CGHAC [86]	KIGS [152]	KIGS [152]	KIGS [152]	Blum [48]
Starting Age (years)	9	10	7	9	8	7
Age at end of treatment (years)	16	16	15	14	14	14
Treatment duration	7	6	8	5	6	7
Sex (males %)	70	0	50	71	59.6	48

For GHD, some children continue to receive rhGH treatment into adulthood. This is shown as an additional scenario for GHD where, it is assumed that 34% of GHD patients continue treatment[157] until age 25 years with a dose of 0.4 mg/day.[164] These individuals do not receive any additional benefit associated with height gain from this treatment in the model.

The clinical effect of rhGH was taken from the systematic review in Section 3. Where possible the clinical effect was taken from the best quality RCT where children had treatment for a sufficiently long time to capture HtSDS height gain, which we assumed would be at least two years. For GHD, these data were not available, as the only available RCT was for only 1 year, and so we have used observational data (KIGS database)[152] to estimate the clinical effect (Table 42). For SGA, there were no RCTs available for the licensed dose and so we used a study with one year treatment.[114] For TS, height gain was reported in terms of age specific TS HtSDS, but the mean age specific value was not reported. We assumed that the age specific TS HtSDS was that reported in the KIGS database.[152] Several studies have not reported the height gain in cm, and for these studies we converted HtSDS values to cm, using the height table from HSE 2003.[151]

Table 42 Clinical effect for rhGH used in the SHTAC model

Parameter	GHD	TS	PWS	CRI	SGA	SHOX-D
Source	KIGS [152]	CGHAC [86]	De Lind van Wijngaarden ‡[93]	Fine [108]	Philips [114]	Blum [48]
Treatment cohort						
Starting HtSDS	-2.99	-3.4*	-2.0	-2.9	-3.1	-3.3
Final HtSDS	-1.17	-1.8*	-0.5	-1.6	-2.3	-2.1
Control cohort						
Starting HtSDS	-2.99	-3.3*	-2.5	-2.9	-3.1	-3.3
Final HtSDS	-2.99	-3.0*	-2.6	-2.9	-3.0	-3
Treatment effect						
Treatment height gain (SDS)	1.82	1.3	1.6	1.3	0.7	0.9
Treatment height gain (cm)	12.8†	9.3	11.1†	9.2†	3.3	6.3†
QoL gain	0.069	0.069	0.021	0.059	0.043	0.055

SDS, standard deviation score; * estimated based on age-specific turner SDS score, converted to SDS score using KIGS database [152], † HtSDS gain converted to cm using HSE 2003[151], ‡ Results reported as median values

A review of compliance with rhGH was conducted by Merck Serono as part of the manufacturers' submissions. It found that estimates for compliance ranged from 69% to 95% for the studies identified. One study estimated concordance in 75 children by using data on GP prescriptions over 12 months.[156] Between one and two injections/week were missed by 16% of the children, and 23% missed >2 injections/week. Based on this study, we assumed a compliance of 85%.

4.4.3.3 Health related QoL

There was a lack of good quality HRQoL data expressed in terms of utility in the RCTs and other QoL studies for most of the conditions of interest (Section 4.2). Only one study was found that was appropriate to the conditions of interest and this was for GHD.[147] However, it was in an adult population and it was uncertain whether the participants had already benefited from growth hormone as children; the QoL utility gain from this study was similar to that from the Christensen and Colleagues study for GHD.[150] For the other studies the most appropriate utility measurement was from the study by Christiansen and colleagues[150] which measured QoL using the EQ-5D in a large sample of the general UK population (Health Survey for England). The utility values are not from the conditions of interest; nevertheless it does provide a common utility gain that could be compared across all the conditions of interest and that could be used with the clinical

effectiveness outcomes from the RCTs. It was assumed for children that the adult gain in utility from increased height derived from the Christensen and colleagues study would be the same as a utility gain in children.

This study assessed HRQoL estimates through the use of OLS linear regression which controlled for age, weight and gender. More details on the study are reported in section 4.2.2. We assumed that individuals in the treated and untreated cohorts would have no difference in terms of age, gender, social class, weight and long standing illness. The differences in HRQoL utility estimates between the treated and untreated cohorts are therefore derived from their differences in height. According to the regression, for those people shorter than – 2.0 HtSDS, an improvement of 1 HtSDS will result in a change in HRQoL utility of 0.061. For the subgroup between -2.0 and 0 HtSDS, a 1 HtSDS improvement increases utility by 0.01. These values were used in the SHTAC estimation of cost effectiveness.

For PWS patients, there may be an additional health benefit associated with improved body composition. Any improvements in body composition may lead to reduced risk of diabetes and cardiovascular disease. However there is considerable difficulty estimating the magnitude of this effect and extrapolating short term treatment in childhood to lifelong benefit. There was one study of poor quality in adults with PWS but this was not considered to be a robust estimate of QoL benefit (see section 1.3.5). The MS estimated a QoL benefit from reduced diabetes risk but it was not possible to verify this evidence. Due to the high uncertainty around the estimates of QoL benefit, we assumed no benefit due to body composition in the base case and then conducted sensitivity analyses using the studies mentioned above.

4.4.3.4 Estimation of costs

The costs used in the SHTAC model were based upon those used in the previous HTA report.[5] The annual cost of monitoring associated with each condition was calculated for each arm of the model using treatment pathways described in that report. Treatment costs are calculated on the basis of mean dose of rhGH. Unit costs for drugs were taken from the British National Formulary[165] and, for consultations, outpatient visits and procedures, from NHS Reference Costs.[166] The base year used for the analysis was 2008; where necessary, costs were inflated to this year. The resource use is based on those from the previous HTA report. Based on clinical opinion, the nurse visit time was assumed to be the same for all conditions and patients would

have two outpatient visits per year. Furthermore, patients would no longer have a hand x ray at the end of treatment. All children are monitored until they reach adulthood, assumed to be age 17 years old. The unit costs and resource use are shown in Table 43, and Table 44 respectively.

Table 43 Unit costs used in the SHTAC model

Costs component	Cost, £	Source
Cost per outpatient attendance first contact face to face paediatric endocrinology (HRG code 302F)	£206.28	NHS ref costs 2007/8[166]
Cost per outpatient attendance subsequent contact face to face paediatric endocrinology (HRG code 302F)	£127.97	NHS ref costs 2007/8 [166]
Specialist community nurse per patient contact (1 hour)	£73	PSRU 2008 [167]
Community nurse per patient visit (1 hour)	£64	PSRU 2008 [167]
Blood tests (for full blood count, chemical profile, thyroid and IGF)	£51	SUHT 2008 [168]
X-Ray-hand (bone age test)	£28.64†	NHS ref costs 2006/7 [166]
Pituitary function test (glucagon, insulin stress test) includes 2 hours nurse time	£207.50	SUHT 2008 [168]

† original cost of £27.71 inflated to 2008 costs

Table 44 Administration and monitoring resource use

	GHD	TS, PWS, CRI, SGA, SHOX-D
No treatment monitoring		
Outpatient visit	2	2
Blood test	1	1
Treatment 1st year		
Specialist Nurse home visit	1 hour	1 hours
Community nurse home visits	4 hour	4 hours
Outpatient visit	2	2
Blood test	1	1
Pituitary function test	0.2	0
GH treatment subsequent year		
Outpatient visit	2	2

Blood test	1	1
Hand X-ray	1	1
Pituitary function test	0.2	0
End of treatment		
Outpatient visit	1	1

The cost of the drug used in the manufacturers' models varies between £18 and £23.19 per mg. We have assumed a drug cost of £23.18 in the base case, as two drugs are this price, and vary the price in sensitivity analysis. Drug costs are calculated according to the dosage used (Table 45) and the weight of the child.[165] The weight of children at different ages was taken from a long term observational database (Appendix 13). [152]

Table 45 Drug dosage (mg/kg/day)

Condition	GHD	TS	PWS	CRI	SGA	SHOX-D
Drug dosage	0.025	0.045	0.035	0.045	0.035	0.045

4.5 Estimation of cost-effectiveness

This section reports the cost effectiveness results for a cohort of 1000 children for each of the conditions of interest who received rhGH treatment. Results for costs and quality adjusted life years (QALYs) are presented for children in the cohort for a treated and untreated cohort, with costs and benefits discounted at 3.5%. The cost effectiveness of rhGH compared to no treatment is presented as incremental cost per QALY and incremental cost per cm gained. The results are shown in Table 46 for each condition. In the base case analysis, all conditions except GHD, used the clinical benefit seen in the best quality RCT for each condition (Section 3). The cost effectiveness of rhGH versus no treatment varied from £25,483 for GHD to £148,860 for PWS per QALY gained. With the exception of PWS, all conditions have an ICER lower than £45,000 per QALY gained.

Table 46 Cost effectiveness results for the base case analysis

Condition		Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)	cm gain	ICER (£/cm)
GHD	No rhGH treatment	£2,211	16.8					
	rhGH treatment	£41,562	18.4	£39,351	1.54	£25,483	12.80	£3,074
TS	No rhGH treatment	£1,965	15.9					
	rhGH treatment	£68,829	17.4	£66,864	1.54	£43,405	9.30	£7,190
PWS	No rhGH treatment	£2,646	17.6					
	rhGH treatment	£74,317	18.1	£71,671	0.48	£148,860	11.10	£6,457
CRI	No rhGH treatment	£1,876	11.6					
	rhGH treatment	£39,289	12.4	£37,413	0.87	£43,214	9.20	£4,067
SGA	No rhGH treatment	£2,432	17.1					
	rhGH treatment	£37,636	18.1	£35,204	0.97	£36,392	3.30	£10,668
SHOX-D	No rhGH treatment	£2,646	16.8					
	rhGH treatment	£58,527	18.1	£55,881	1.25	£44,596	6.30	£8,870

Inc.= incremental

A further analysis was undertaken to see the effect of continuation of rhGH treatment into adulthood for 34% of the original cohort until the age of 25. The incremental cost per QALY was £31,026 (Table 47).

Table 47 Cost effectiveness results for continuation of rhGH treatment into adulthood for GHD patients

Condition		Costs (£)	QALYs	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)	cm gain	ICER (£/cm)
	No rhGH treatment	£2,211	16.8					
GHD continuers	rhGH treatment	£50,123	18.4	£47,912	1.54	£31,026	12.80	£3,743

Inc.=incremental

4.6 Sensitivity analyses

4.6.1 Cost effectiveness of rhGH treatment – deterministic sensitivity analysis

One-way deterministic sensitivity analyses were performed, in which model parameters were systematically and independently varied, using a realistic minimum and maximum value. The sensitivity analysis investigated the effect of uncertainty around the model structure and for variation in parameters on the cost-effectiveness results, in order to highlight the most influential parameters. The effects of uncertainty in multiple parameters were addressed using probabilistic sensitivity analysis, which is reported later in this section. Where possible, the parameters were varied according to the ranges of the confidence intervals of these parameters, based on the published estimate. Where these data were not available a alternative suitable range was chosen. The same ranges were used in the deterministic and probabilistic sensitivity analyses and these are described in Appendix 12.

Table 48 shows the results for each of the conditions using the KIGS database[152] for estimate of the clinical benefit. The KIGS database, a large observational study of children treated with rhGH, was used for the effectiveness of GHD in the base case reported above. According to these results, an ICER of rhGH versus no treatment varied from an ICER of £20,880 per QALY gained for SGA to £158,470 per QALY gained for PWS. Results are of a similar magnitude to the base case with the exception of the SGA analyses. The ICER for SGA is much lower in this analysis because the incremental clinical height gain is lower in the RCT effectiveness data compared to the KIGS effectiveness data.

Table 48 Cost effectiveness results with clinical benefit from KIGS database

		Height			Incremental	Incremental	ICER
Condition		(HtSDS)	Costs (£)	QALYs	costs (£)	QALYs	(£/QALY)
TS	No treatment	-3.18	£1,965	15.8			
	rhGH treatment	-2.24	£68,829	17.1	£66,864	1.28	£52,307
PWS	No treatment	-2.22	£2,646	17.4			
	rhGH treatment	-1.36	£74,317	17.9	£71,671	0.45	£158,473
CRI	No treatment	-2.99	£1,876	11.5			
	rhGH treatment	-2.17	£39,289	12.2	£37,413	0.74	£50,885
SGA	No treatment	-3.23	£2,432	16.8			
	rhGH treatment	-2.01	£37,636	18.4	£35,204	1.69	£20,881
SHOX-D	No treatment	-3.18	£2,646	16.6			
	rhGH treatment	-2.24	£58,527	17.9	£55,881	1.31	£42,698

The discount rates used for the analyses have a large effect on the results. This is due to the upfront costs at the beginning of the model and the health outcomes stretching over the life time of the model. Table 49 shows the results using the discount rates used in the previous HTA report, i.e. costs 6% and benefits 1.5%. Using these discount rates, rhGH treatment is more cost effective. For all conditions, except PWS, the ICER reduces to less than £30,000 per QALY.

Table 49 Cost effectiveness results with alternative discount rates

Condition	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	cm gain	ICER (£/cm)
GHD	£35,597	2.49	£14,279	12.80	£2,781
TS	£61,324	2.49	£24,592	9.30	£6,594
PWS	£63,884	0.79	£81,222	11.10	£5,755
CRI	£34,776	1.22	£28,389	9.20	£3,780
SGA	£32,297	1.57	£20,558	3.30	£9,787
SHOX-D	£50,536	2.05	£24,683	6.30	£8,022

Table 50 to Table 55 report the results of the deterministic sensitivity analyses for the conditions for the most influential parameters. Other variables were varied in sensitivity analyses but were found to only have a negligible effect on the results. The costs effectiveness results are fairly sensitive to the variation in parameters included in the deterministic sensitivity analysis. For all the conditions, the model results are most sensitive to treatment start age and length, compliance and utility gain.

The deterministic sensitivity results for GHD are shown in Table 50. The results were most sensitive to dosage and varied between £23,480 and £39,480 per QALY gained.

Table 50 Deterministic sensitivity analyses for GHD

Parameter	Baseline	Upper value	Lower value	Upper value ICER, (£/QALY)	Lower value ICER, (£/QALY)	Range
Dosage, mg/kg	0.025	0.039	0.023	£39,484	£23,482	£16,002
Utility gain per HtSDS	0.061	0.073	0.049	£21,725	£30,812	£9,087
Compliance	85%	100%	70%	£29,895	£21,070	£8,824
Treatment age, years	9 – 16	11 – 16	7 – 16	£21,180	£28,187	£7,007
Utility benefit spread over	2 years	1 year	7 years	£24,973	£28,165	£3,192
Cost of rhGH treatment £/mg	£23.18	£23.18	£22.00	£25,483	£24,210	£1,273

Standard mortality rate	1	2.4	1	£25,483	£24,371	£1,112

The deterministic sensitivity results for TS are shown in Table 51. The results were most sensitive to utility gain and varied between £36,440 and £53,660 per QALY gained.

Table 51 Deterministic sensitivity analyses for TS

Parameter	Baseline	Upper value	Lower value	Upper value ICER, (£/QALY)	Lower value ICER, (£/QALY)	Range
Utility gain per HtSDS	0.061	0.073	0.049	£36,444	£53,655	£17,211
Treatment age, years	10 – 16	12 – 16	8 – 16	£33,552	£49,616	£16,064
Compliance	85%	100%	70%	£51,018	£35,793	£15,224
Dosage, mg/kg	0.045	0.05	0.4	£48,198	£38,612	£9,586
Utility benefit spread over	2 years	1 year	6 years	£42,538	£47,027	£4,489
Standard mortality rate	1	2.4	1	£43,405	£41,038	£2,367
Cost of rhGH treatment £/mg	£23.18	£23.18	£22.00	£43,405	£41,209	£2,196

The deterministic sensitivity results for PWS are shown in Table 52. The results were most sensitive to compliance and varied between £122,720 and £175,000.

Table 52 Deterministic sensitivity analyses for PWS

Parameter	Baseline	Upper value	Lower value	Upper value ICER, (£/QALY)	Lower value ICER, (£/QALY)	Range
Compliance	85%	100%	70%	£175,002	£122,718	£52,284
Treatment age, years	7 – 15	9 – 15	5 – 15	£130,959	£158,587	£27,628
Utility benefit spread over	2 years	1 year	8 years	£145,927	£167,522	£21,595
Dosage, mg/kg	0.035	0.035	0.03	£148,860	£127,697	£21,163
Utility gain per HtSDS	0.061	0.073	0.049	£140,849	£157,836	£16,987
Cost of rhGH treatment £/mg	£23.18	£23.18	£22.00	£148,860	£141,319	£7,541
Standard mortality rate	1	2.4	1	£148,860	£142,621	£6,239

The deterministic sensitivity results for CRI are shown in Table 53. The results were most sensitive to the treatment start age and length of treatment and varied between £30,902 and £51,137 per QALY gained.

Table 53 Deterministic sensitivity analyses for CRI

Parameter	Baseline	Upper value	Lower value	Upper value ICER, (£/QALY)	Lower value ICER, (£/QALY)	Range
Treatment age, years	9 – 14	11 – 14	7 – 14	£30,902	£51,137	£20,235
Utility benefit spread over	2 years	1 year	5 years	£42,092	£59,534	£17,442
Utility gain per HtSDS	0.061	0.073	0.049	£26,798	£38,833	£12,035
Standard mortality rate	21	1	21	£31,712	£43,214	£11,502
Compliance	85%	100%	70%	£37,293	£26,131	£11,162
Dosage, mg/kg	0.045	0.05	0.04	£48,005	£38,426	£9,579
Cost of rhGH treatment £/mg	£23.18	£23.18	£22.00	£43,214	£41,020	£2,194

The deterministic sensitivity results for SGA are shown in Table 54. The deterministic sensitivity results were most sensitive to utility gain and varied between £30,410 and £45,305 per QALY gained.

Table 54 Deterministic sensitivity analyses for SGA

Parameter	Baseline	Upper value	Lower value	Upper value ICER, (£/QALY)	Lower value ICER, (£/QALY)	Range
Utility gain per HtSDS	0.061	0.073	0.049	£30,410	£45,305	£14,895
Treatment age, years	8 – 14	10 – 14	6 – 14	£28,251	£41,718	£13,467
Compliance	85%	100%	70%	£42,786	£29,999	£12,787
Dosage, mg/kg	0.035	0.04	0.035	£41,568	£36,392	£5,176
Utility benefit spread over	2 years	1 year	6 years	£35,670	£39,406	£3,736
Cost of rhGH treatment £/mg	£23.18	£23.20	£22.00	£36,392	£34,548	£1,844
Standard mortality rate	1	2.4	1	£36,392	£34,828	£1,564

The deterministic sensitivity results for SHOX-D are shown in Table 55. The deterministic sensitivity results were most sensitive to utility gain and varied between £37,265 and £55,517 per QALY gained.

Table 55 Deterministic sensitivity analyses for SHOX-D

Parameter	Baseline	Upper value	Lower value	Upper value ICER, (£/QALY)	Lower value ICER, (£/QALY)	Range
Utility gain per HtSDS	0.061	0.073	0.049	£37,265	£55,517	£18,252
Compliance	85%	100%	70%	£52,438	£36,753	£15,685
Treatment age, years	8 – 15	10 – 15	6 – 15	£37,178	£49,142	£11,964
Dosage, mg/kg	0.045	0.04	0.05	£39,658	£49,534	£9,876
Utility benefit spread over	2 years	1 year	7 years	£43,717	£49,214	£5,497

Cost of rhGH treatment £/mg	£23.18	£23.20	£22.00	£44,596	£42,333	£2,263
Standard mortality rate	1	2.4	1	£44,596	£42,716	£1,880

For PWS patients, there may be an additional health benefit associated with improved body composition which may reduce the risk of diabetes and other morbidities. This difficulty with extrapolating between childhood treatment and adult morbidity and QoL has been discussed in section 4.4.3.3. In the base case we have assumed no HRQoL benefit associated with changes in body composition. In this section we present a scenario analysis for additional changes in body composition. However, there is a difficulty linking changes in lean fat mass to changes in utility as there are no utility studies for lean fat mass. For this reason we have focused on changes in BMI.

Picot and colleagues [169] conducted a targeted search to identify published utility estimates for the BMI values relevant to an adult obese population. The search aimed to identify estimates of the change in utility scores based on the unit change in BMI values. Utility estimates were only considered where they used a validated, multi-attribute utility scale (e.g. EQ5-D) or appropriate methodology (e.g. standard gamble or time trade off techniques) and provided a clear definition of utility scores anchors 0 and 1. They suggest the values reported by Hakim and colleagues [170] represent the most methodologically sound estimates derived from subjects across a wide range of obesity levels. Hakim and colleagues [170] found that a one unit decrease in BMI, over a period of one year, was associated with a gain of 0.017, which was independent of age or gender.

RCTs for PWS, in Section 3.5.3, reported mixed results for changes in BMI with a maximum BMI difference of 1.8 between treated and untreated groups after two years treatment. Assuming this change in BMI is maintained lifelong, and therefore there is an additional utility of 0.031, the cost effectiveness of PWS would be £60,753 per QALY gained.

4.6.2 Probabilistic sensitivity analyses

In the probabilistic sensitivity analyses the main parameters were sampled probabilistically from an appropriate distribution using similar ranges as used in the deterministic sensitivity analyses. The parameters sampled were: starting age, length of treatment, dose, HtSDS at the start and end of treatment for both the rhGH and no treatment cohorts, utility increment for gains in height and all costs used in the base case excluding the cost of rhGH.

The distribution assigned to each variable included in the probabilistic sensitivity analysis and the parameters of the distribution are reported in Appendix 12. One thousand simulations were run for each condition of interest in this analysis. Table 56 reports the mean costs and outcomes from the probabilistic analysis and the ICER for rhGH compared with no treatment, based on the mean values generated in the probabilistic sensitivity analysis. Table 57 shows the 2.5% and 97.5% percentiles for the probabilistic sensitivity analyses.

Table 56 Costs and outcomes from the probabilistic sensitivity analysis.

				Incremental	Incremental	ICERs
Condition		QALYs	Costs (£)	QALYs	Costs (£)	(£/QALY)
GHD	No treatment	18.35	£40,992	1.54	£38,789	£25,151
	rhGH treatment	16.81	£1,995			
TS	No treatment	17.44	£68,097	1.55	£66,132	£42,617
	rhGH treatment	15.88	£1,964			
PWS	No treatment	18.18	£73,939	0.60	£71,296	£118,397
	rhGH treatment	17.57	£2,643			
CRI	No treatment	12.43	£38,822	0.86	£36,948	£43,129
	rhGH treatment	11.57	£1,874			
SGA	No treatment	18.06	£37,348	0.97	£34,907	£36,085
	rhGH treatment	17.09	£2,441			
SHOX-D	No treatment	18.06	£57,973	1.26	£55,332	£44,082
	rhGH treatment	16.81	£2,641			

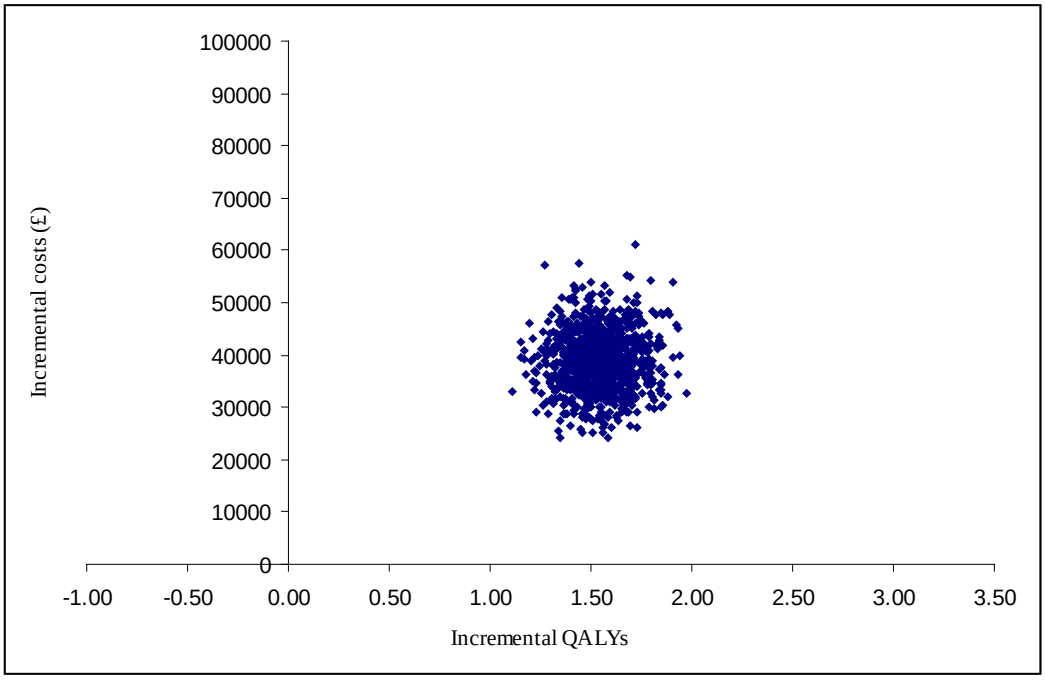
Table 57 Ranges from the probabilistic sensitivity analysis (2.5% and 97.5% percentiles)

	Incremental QALYs		Incremental	Costs (£)	ICERs	
Condition	Min	Max	Min	Max	Min	Max
GHD	1.27	1.83	£28,379	£50,383	£17,708	£34,596
TS	0.78	2.35	£47,941	£84,043	£25,412	£88,838
PWS	-0.33	1.53	£54,703	£88,079	-£872,526	£1,221,669
CRI	0.47	1.30	£24,872	£48,581	£24,023	£81,771
SGA	0.53	1.49	£25,423	£45,192	£21,700	£68,262
SHOX-D	0.53	2.10	£41,196	£67,820	£25,198	£107,108

The mean cost effectiveness ICER from the probabilistic analyses is slightly lower than the deterministic cost effectiveness for GHD, TS, CRI, SGA and SHOX-D (which was £25,483,

£43,305, £43,214, £36,392 and £44,596, respectively). The cost effectiveness from the PSA for PWS, however, is much lower at £118, 397 than the deterministic estimate. This is due to non-linearity in the PWS model due to the baseline HtSDS for the treated group being at -2.0 HtSDS where the utility gain changes. The sampling is drawing across two different utility gains for this HtSDS, therefore decreasing the ICER in the PSA.

Figure 2 Cost effectiveness plane – incremental cost and incremental QALYs for rhGH treatment and no treatment in GHD



Scatter plots are shown for the incremental cost and incremental QALYs for each of the condition in Figure 2 to Figure 7. In addition, a cost effectiveness acceptability curve was also derived, representing the proportion of simulations when GH treatment is cost effective for a range of willingness to pay thresholds, up to £100,000, see Figure 8

Figure 3 Cost effectiveness plane – incremental cost and incremental QALYs for rhGH treatment and no treatment in TS

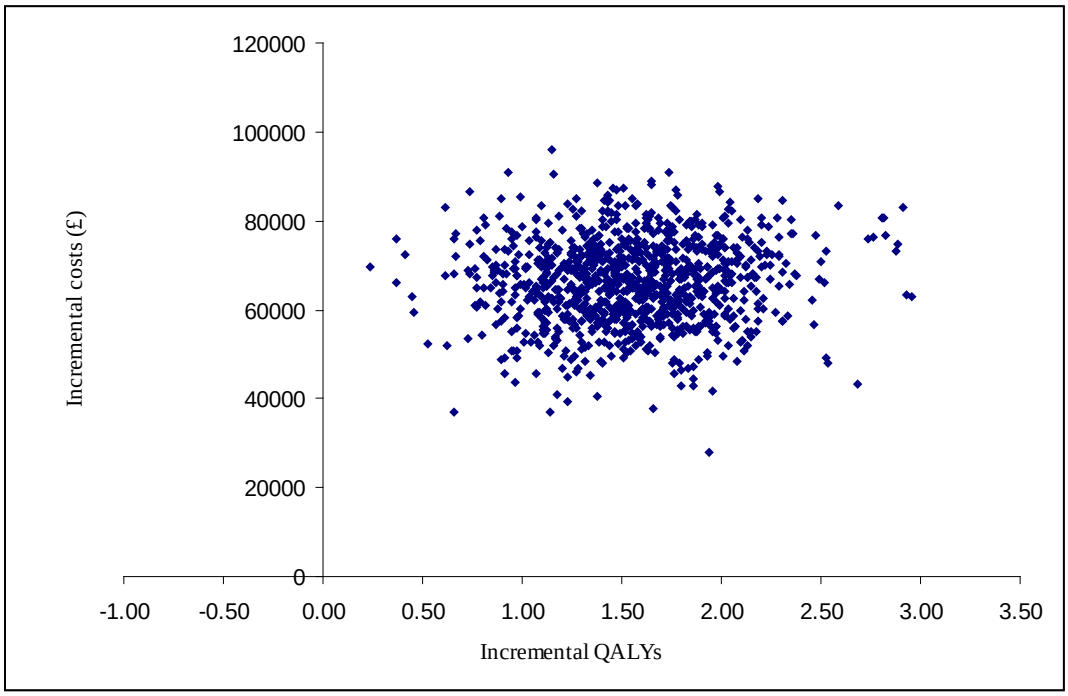


Figure 4 Cost effectiveness plane – incremental cost and incremental QALYs for rhGH treatment and no treatment in PWS

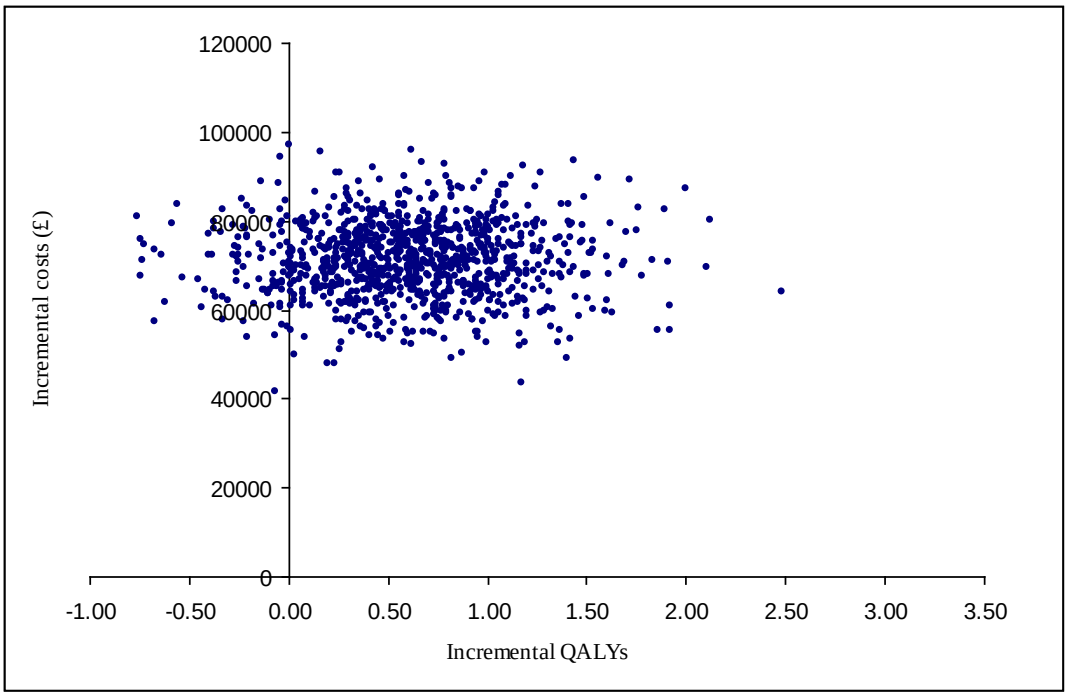


Figure 5 Cost effectiveness plane – incremental cost and incremental QALYs for rhGH treatment and no treatment in CRI

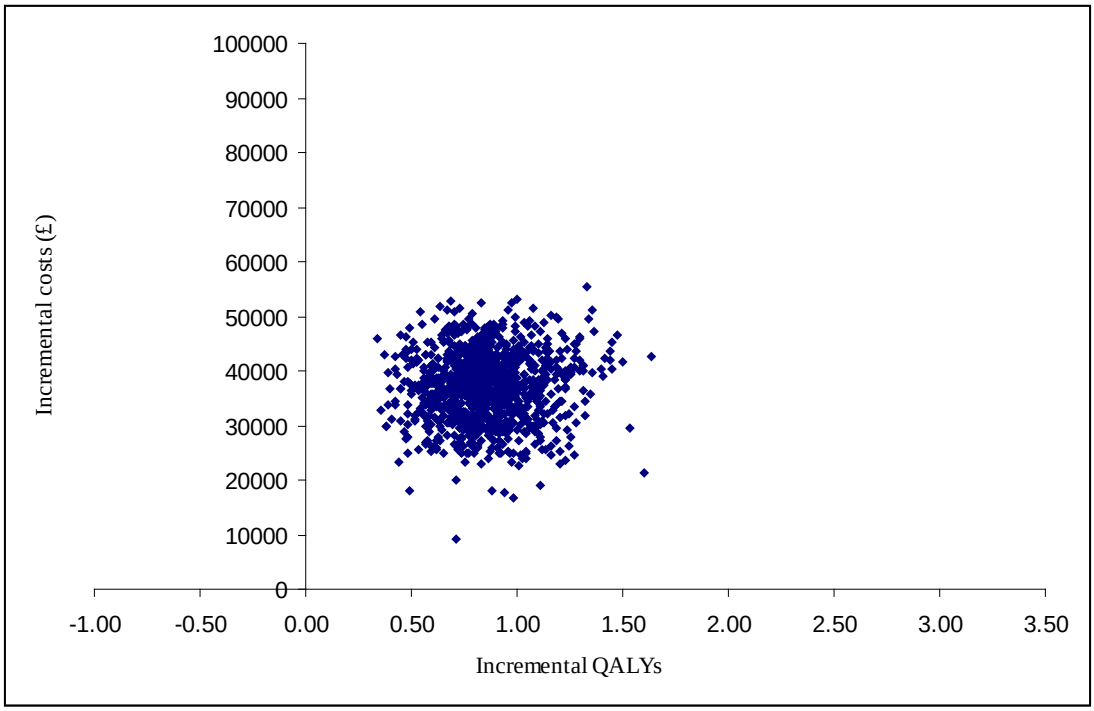


Figure 6 Cost effectiveness plane – incremental cost and incremental QALYs for rhGH treatment and no treatment in SGA

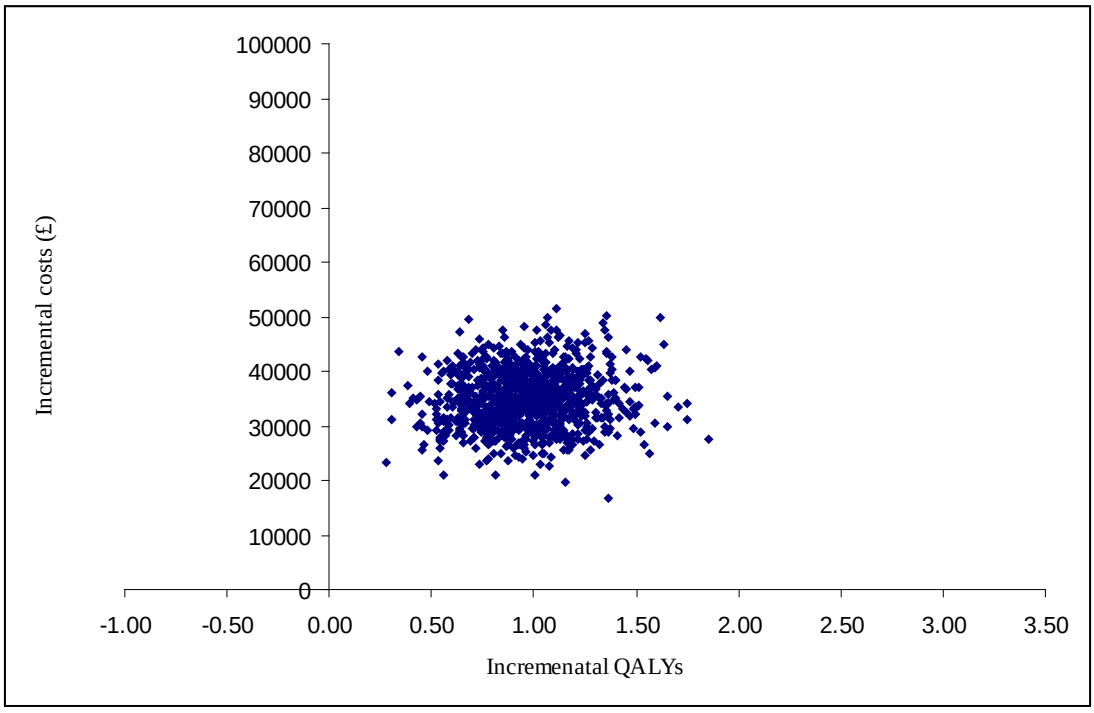


Figure 7 Cost effectiveness plane – incremental cost and incremental QALYs for rhGH treatment and no treatment in SHOX-D

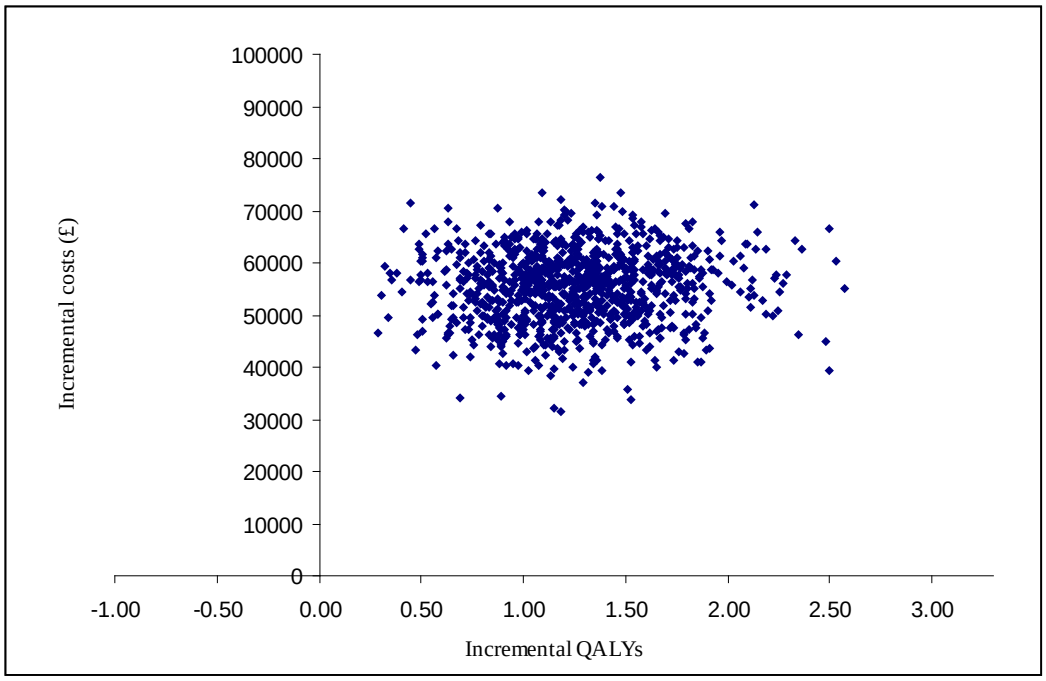
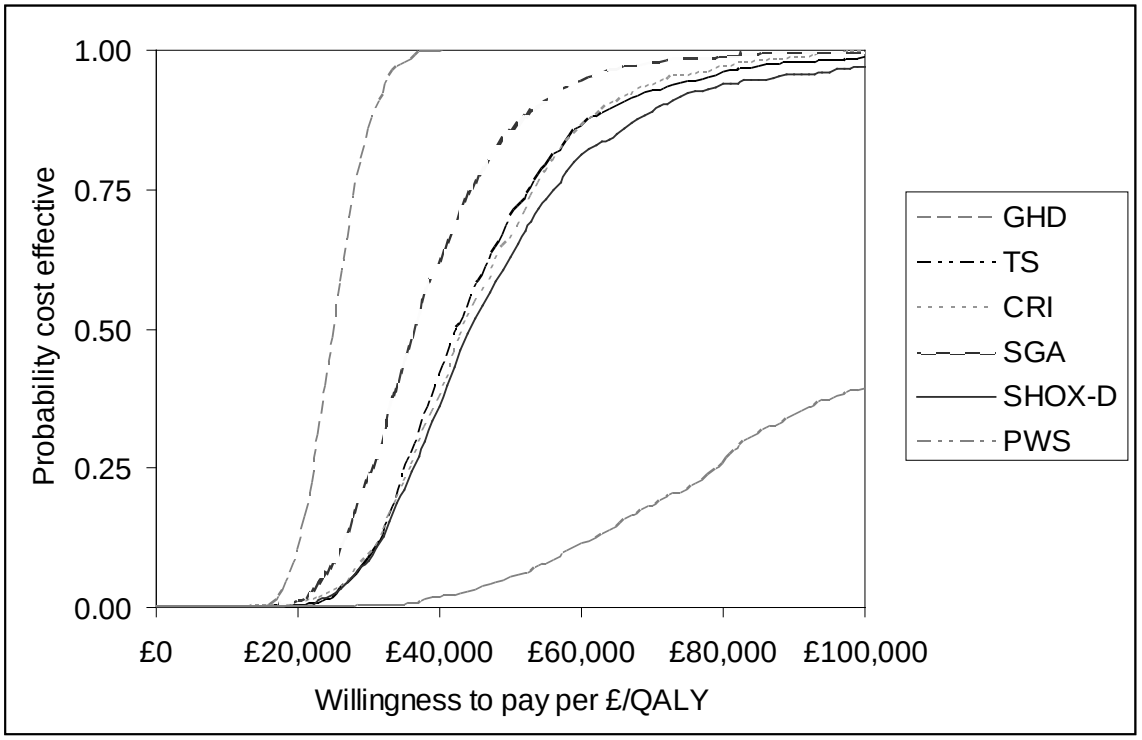


Figure 8 Cost effectiveness acceptability curve for rhGH treatment and no treatment for all the conditions.



In this analysis, rhGH treatment had the probability of being cost effective at willingness to pay thresholds of £20,000, £30,000 and £50,000 per QALY as: 15%, 88% and 100% for GHD, 0%, 12% and 68% for TS, 0%, 0% and 2% for PWS, 1%, 11% and 70% for CRI, 2%, 28% and 86% for SGA, and 3%, 15% and 63% for SHOX-D, respectively.

Summary of cost effectiveness

- A systematic search of the literature found two fully published economic evaluations of rhGH treatment for TS and GHD. The results from the studies varied due to the choice of utility estimates and assumptions on the effectiveness.
- A systematic search for published studies of QoL for patients with individuals with the conditions of interest who had rhGH identified six studies although none of these were in children. These were generally small studies of poor quality. One study was considered of reasonable quality.[147] This study estimated HRQoL for adults with GHD.
- An additional targeted search was undertaken for QoL in relation to height which identified one study [150] which provided utility estimates based on the EQ-5D for different HtSDS from the Health Survey for England.
- Six of the seven manufacturers submitted evidence to be considered for this review. One manufacturer's submission by Sandoz did not comply with the NICE template for MTA and presented a description of the product (Omnitrope) and what appears to be a cost-minimisation analysis using Genotropin as a comparator (defined as a reference product). The other five out of the six manufacturers' submissions consisted of a written report and an electronic model supporting the cost-effectiveness analyses. This model was used by the five collaborating manufacturers involved in the submission to the MTA of rhGH in the cost-effectiveness analysis of treatment in children with GHD, TS, PWS, CRI, SGA.
- Each of the collaborating manufacturers presented essentially the same model with some minor modifications. The model developed was based upon the previous HTA report[5] but has been extended to consider longer term outcomes in order to estimate cost effectiveness in terms of QALYs.
- The utility scores used in the MS model in children with GHD, TS, CRI and SGA were based upon the study by Christensen and colleagues[150] which estimate QoL associated with height for a general population survey. However they used the utility point estimates, based only on

height, instead of the regression analysis from the study which controlled for other key variables.

- In the manufacturers' base case, the cost effectiveness results for all conditions were less than £30,000 per QALY gained. They estimated ICER of: £17,552 for GHD, £29,757 for TS, £32,540 for PWS, £15,962 for CRI, and £18,167 for SGA per QALY gained.
- The authors of this report developed an independent model, based upon the previous HTA report, and extended to consider longer term outcomes in order to estimate cost effectiveness in terms of QALYs.
- From this independent model, the incremental cost per QALY estimates of rhGH compared to no treatment were: £25,483 for GHD, £43,405 for TS, £148,860 for PWS, £43,214 for CRI, £36,392 for SGA and £44,596 for SHOX-D. A further analysis was run for PWS which included a lifelong improvement of body composition of 1.8 kg/m^2 BMI and an associated additional utility of 0.031. Under these assumptions, the cost effectiveness of PWS reduced to £60,753 per QALY gained.
- The effect of a range of parameter values in the economic model were evaluated in sensitivity analyses. The model results were found to be most sensitive to the discount rate used. When the previous NICE discount rate of 6% for costs and 1.5% for benefits was used, all conditions were cost effective for a willingness to pay threshold of £30,000 per QALY. The model results are also sensitive to treatment start age and length, compliance and utility gain.
- The probabilistic sensitivity analysis estimated the probability of each of the conditions to be cost effective at £30,000 to be: 88% for GHD, 12% for TS, 0% for PWS, 11% for CRI, 28% for SGA and 15% for SHOX-D.

5 ASSESSMENT OF FACTORS RELEVANT TO THE NHS AND OTHER PARTIES

Expert opinion suggests that there has been a trend towards managing patients in tertiary centres with local paediatricians, which allows for a greater degree of centralisation, and may improve compliance. NICE guidance already recommends treatment with rhGH for children who have short stature associated with GHD, TS, PWS and CRI. Prescriptions associated with these conditions are therefore already part of PCTs' budgets, and are unlikely to increase significantly. However, expert opinion indicates that many families of children with PWS are now seeking treatment in infancy rather than in mid-childhood, and there may also be some increase in the number of prescriptions for GHD associated with oncology, since greater numbers of children are surviving childhood cancers. The newly licensed conditions SHOX-D and SGA are not yet covered by NICE guidance. Of the estimated 4758 UK patients currently receiving rhGH,[73] a breakdown by diagnosis for 3951 of them found that only 5.2% (205 patients) were receiving treatment for short stature associated with being born SGA. Clinical opinion indicates that there is unlikely to be a large increase in prescriptions for SGA children, particularly if assessment were to be undertaken in tertiary centres.

The BSPED survey did not include patients with SHOX-D, and it is not clear how many children with this condition are currently receiving treatment. Children with short stature due to unknown causes, or with other conditions such as LWS not currently covered by NICE guidance, might have an underlying SHOX-D. The availability of prescriptions to these new groups of patients could therefore have a budgetary impact. However, these conditions are very rare, so there is unlikely to be a large increase in people requiring treatment.

6 DISCUSSION

6.1 Statement of principle findings

6.1.1 GHD

The use of rhGH as replacement therapy is well-established in children who have a deficiency of the natural hormone. Therefore most clinicians would consider it unethical to withhold treatment and there is a corresponding lack of RCT evidence in the literature. Only one trial met the inclusion criteria for the review of rhGH in children with GHD, and this did not report final height. No details were reported on randomisation or allocation to treatment groups or blinding. The included patients (n=19) were part of a larger study, which was generally poorly reported. After a year's treatment, HtSDS was statistically significantly higher in treated than in untreated children, although actual height was not reported. Children who received rhGH for one year had grown at a mean velocity of 2.7 cm/year faster than untreated children, which was statistically significantly faster. The low patient numbers mean that the evidence base for GHD is weak. Thus, there is very limited evidence of a slight increase in growth for children with GHD treated with growth hormone, based on one study of mixed quality. Estimates of height gain in the previous HTA report suggested final height gains of approximately 1.3-1.6 SDS (i.e. within 2SD of the normal mean) with rhGH treatment. However, these figures were from retrospective single-cohort studies which were not included in the present review.

The cost effectiveness estimate of rhGH treatment in GHD is about £25,480 per QALY gained or £3,070 per cm gained. As there were no appropriate RCTs, the KIGS database was used for the estimate of height gain from rhGH. This estimate for height gain was higher than for the other conditions. The previous HTA report estimated a cost per cm gained of £6000 using 8 years treatment compared to the 7 years used in our analysis and a slightly lower height gain from the KIGS database. The cost effectiveness estimate fro the cohort of GHD who continue rhGH treatment into adulthood was £31,026 per QALY gained and £3,743 per cm gained.

6.1.2 TS

Six trials met the inclusion criteria for the review of growth hormone for growth disturbance in patients with TS. There is some evidence of effectiveness across all reported growth outcomes for girls with TS. However, these results are reported in studies of poor reporting and methodological

quality, and in some cases of short duration. Of the six included studies, none of the included trials employed an intention–to-treat analysis, one reported adequate randomisation to treatment groups, one study described adequate concealment of treatment allocation, and one adequately blinded the patient to treatment by administering placebo.

Children in the rhGH group in a large RCT which followed girls until final height grew an average of 9.3cm more from baseline than those in the untreated group. In a study of younger children over two years, the difference was 7.6cm. Both of these were statistically significant results. Weight and WtSDS were found to be significantly greater in the treated group in one study of younger girls with TS.

The searches for this study identified a new systematic review, conducted in Canada in 2007. The review concluded that rhGH is effective in improving growth and final height in girls with TS, but found no evidence available in the clinical trials to suggest that rhGH improves QoL. The evidence discussed in the present review reflects this, as we found some evidence for increased height but no RCT evidence for improvements in QoL.

In summary, there is some evidence of effectiveness across all reported growth outcomes for girls with growth disturbance as a result of TS. There is also evidence of improved body composition. These results are reported in studies of poor reporting and methodological quality, and in some cases short duration, issues which may affect the validity of these findings. The previous HTA report found that treated girls' final height was approximately 5cm taller than untreated controls. The full publication of the large Canadian RCT since the earlier HTA report has shown a slightly larger difference in final height of 9.3cm in final height, as reported in the present review.

The cost effectiveness estimate of rhGH treatment in TS is about £43,400 per QALY gained or £7,190 per cm gained. The estimate of cost effectiveness compares with the estimate of about £130,000 per QALY (at current exchange rates) from the Canadian Agency for drugs and technologies in health,[135] which used a lower QoL benefit for rhGH of 0.042 than used in our analysis. The previous HTA report estimated a less favourable cost per cm gained of £16,000 as they used a lower estimate for height gain of 3.9 cm (compared to 9.3 cm).

6.1.3 PWS

Eight small, rather poorly reported RCTs were included for PWS. Participants' average ages ranged from 13 months to 10 years. Only the cross-over study used a placebo injection; the parallel-group RCTs had no treatment as the comparison arm.

Treated patients grew an average of 3-5cm/year faster than untreated patients. Only one of the studies reported actual change in height, with infants treated with rhGH growing an average of 6.1cm more than untreated patients during one year. Height SDS was statistically significantly greater in treated patients than in untreated patients after one year (1-1.5 SDS higher) or two years of rhGH treatment (>2 SDS).

Four trials reported a statistically significantly lower percentage of body fat (between 1% and 10% lower) in patients treated with rhGH compared with no treatment or placebo. Three trials reported that patients treated with rhGH had statistically significantly higher lean body mass or a larger improvement in lean body mass than untreated patients. Clinical advice indicates that rhGH characteristically increases lean body mass and reduces fat mass, although weight and BMI do not always change. This is reflected in the RCTs' findings, where changes in BMI were statistically significant in two studies, there were no statistical differences in two other studies, and results were similar between groups in the other two studies.

In summary, patients treated with rhGH grew faster than untreated patients, and tended to have lower body fat percentages. Measurements in treated patients were reported to be statistically significantly better than in untreated patients in several studies, but the included studies were rather small and did not report power calculations or specify a primary outcome, so it is not clear whether they were adequately powered. These findings were comparable with growth and body composition outcomes reported in the previous HTA review. However, the previous review also reported an uncontrolled, single-cohort study of 16 children, which suggested that rhGH treatment normalised final height.

The cost effectiveness estimate of rhGH treatment in PWS is about £148,860 per QALY gained or £6,460 per cm gained. The ICER values for PWS were higher due to the majority of the height gain occurring within -2 HtSDS of average height where a lower utility gain is experienced. The

previous HTA report presented a cost per HtSDS gained of £40,815 and this compares with the current report's estimate of £44,794.

For PWS patients, there may be an additional health benefit associated with improved body composition which may reduce the risk of diabetes and other morbidities. There is considerable difficulty with extrapolating between childhood treatment and adult morbidity and QoL.

RCTs for PWS in the clinical effectiveness review, reported mixed results for changes in BMI with a maximum BMI difference of 1.8 between treated and untreated groups after two years treatment. Assuming this change in BMI is maintained lifelong, and therefore there is an additional utility of 0.031, the cost effectiveness of PWS would be £60,753 per QALY gained.

6.1.4 CRI

The evidence for rhGH in children with CRI came from six RCTs, three of which had fewer than 25 participants, and these might not have been sufficiently powered to test for a real difference between groups. Three of the studies included children who had received renal transplants, and three were for children with CRI who had not had a transplant.

One study reported that treated children grew an average of 3.6 cm more than untreated children in one year, with HtSDS being statistically significantly better in treated children than in untreated children in two studies. Growth was statistically significantly faster in treated children than in untreated children, with between-group differences in velocity ranging from 3.2cm/year to 4.2 cm/year in the parallel-group trials. Children treated with rhGH showed statistically significant improvements in weight gain or WtSDS compared with untreated children in three studies. No QoL data were reported for prepubertal children with CRI. Two rhGH-treated patients in one study experienced acute rejection episodes, but both reversed after treatment with methylprednisolone. There were no serious AE reported.

In summary, treatment with rhGH led to small but statistically significant improvements in growth in children with CRI in two trials, one of which included post-transport patients and the other which included children with CRI who had not received a transplant. The previous HTA review reported differences in HtSDS of approximately 0.8 SD and 1.3 SD for one and two years

of treatment, respectively. The present review found slightly greater differences, favouring rhGH, of approximately 1SDS for 1 year and just over 2 SDS for two years' treatment.

The cost effectiveness estimate of rhGH treatment in CRI is about £43,214 per QALY gained or £4,067 per cm gained. The previous HTA report estimated a cost per cm gained of £7,403 and this was based upon treatment for only 3 years compared with 5 years in this analysis. CRI has a lower QALY gain than the other conditions as we assumed that children with CRI would have a much shorter life expectancy than the general population due to their renal failure.

6.1.5 SGA

The licensing criteria for rhGH in children born SGA with growth disturbance state that eligible children need to have a current HtSDS ≤ -2.5, a parental adjusted HtSDS ≤ -1, a birth weight/ length SDS ≤ -2 SD, and have failed to show catch up growth, defined as GV SDS <0 during the previous year, by four years of age or later. None of the RCTs screened for this review met the inclusion criteria; these were therefore modified, retaining the current height and birth weight/ length SDS criteria. Studies' inclusion criteria were required to state that no catch up growth had taken place by three years of age, but no specific criteria were used for this. The amended inclusion criteria did not require any definition of parental height.

This could affect the generalisability of the results as it is possible that the trials included children with a genetic factor for short stature. However, such children would presumably have a shorter target height than children whose parents are closer to the population mean. So children who meet the marketing authorisation may actually have a greater possibility for increased growth than those in the clinical trials. The other difference between the marketing authorisation criteria and the adapted inclusion criteria used in this review was that the included trials had children as young as three years of age, whereas the licensed population in the UK is children over the age of four. It is possible that an early start for treatment could lead to better results than would be generalisable to the licensed population. However, in practice, the mean age of the children in the included studies was over four years of age for all the trials, so results should be generalisable to the licensed population.

Six trials met the modified inclusion criteria for this review of growth disturbance in children born short for gestational age. However, only one of the studies used the licensed dose for rhGH; the others all used two or three times the licensed dose.

One trial reported adult height, and patients who had received rhGH gained an extra 4cm of height compared with the control group. The difference between treated and untreated patients was statistically significant, as was the difference in adult HtSDS. Another study reported that patients who received 0.033mg/kg/d rhGH (the licensed dose) gained an additional 3.3 cm height compared with untreated children, and those who received 0.1 mg/kg/d gained 6.5 cm of additional height after one year's treatment. Height SDS was found to be greater in children treated with growth hormone in the four studies that reported this outcome.

WtSDS was higher in treated than in untreated groups after both one and two years of treatment in three studies reporting this outcome. Lean mass was reported in one study, being greater in the treated group.

There is very limited evidence of a slight increase in adult height gained in centimetres and SDS, and some evidence of an increase in HtSDS in children receiving growth hormone in these studies. There is also limited evidence of improved body composition outcomes, including a statistically significant mean difference in WtSDS between treated and untreated children. This evidence is from trials which did not meet the licensed inclusion criteria exactly, used higher than the licensed dose in all but one study, and were generally of poor quality with few participants in many cases.

The cost effectiveness estimate of rhGH treatment in SGA is about £36,390 per QALY gained or £10,670 per cm gained. The height gain from the clinical review indicated the gain for SGA was smaller than for the other conditions.

6.1.6 SHOX-D

Only one study reported the use of rhGH in children with SHOX-D, and this was open label and generally poorly reported. Treated children grew approximately 2cm/yr faster than their untreated counterparts after two years of treatment, with a rate of 3.5cm/year quicker than untreated children during the first year. After two years of treatment, children were approximately 6cm

taller than the control group and HtSDS was statistically significantly higher in treated than in untreated patients. Treatment with rhGH raised IGF-I and IGF-BP-3 levels to the upper normal range, but there were no serious AE reported during the study.

The ICER estimate of rhGH treatment in SHOX-D is about £44,596 per QALY gained or £8,870 per cm gained.

6.2 General discussion

This review updates a previous assessment report. [171,172] The criteria for this extended review were broadened to include children with SHOX-D or who were born SGA, as well as those with GHD, TS, PWS or CRI. In addition, we actively searched for all outcome measures including growth, body composition, biochemical markers and QoL.

In the previous HTA report, a cost effectiveness model was constructed that estimated lifetime treatment costs and benefits in terms of cost per cm gained. Those analyses are extended in the present report by including QoL factors in the economic modelling. The cost effectiveness of rhGH has been evaluated using decision analytic models using clinical trial data for the gain in height apart from GHD that used KIGS data. The analysis presented both cost per QALY outcomes together with cost per cm height gained for comparison with the previous HTA report, as shown in Table 58 to Table 60.

Table 58 Base case results for the SHTAC cost effectiveness model

Condition	GHD	TS	PWS	CRI	SGA	SHOX-D
Incremental QALYs	1.54	1.54	0.48	0.87	0.97	1.25
Incremental costs (£)	£39,351	£66,864	£71,671	£37,413	£35,204	£55,881
ICER (£/QALY)	£25,483	£43,405	£148,860	£43,214	£36,392	£44,596
Height gain (cm)	12.8	9.3	11.1	9.2	3.3	6.3
Cost per cm gain	£3,074	£7,190	£6,457	£4,067	£10,668	£8,870

Table 59 Base case results for Pfizer

	GHD	TS	PWS	CRI	SGA
Incremental QALYs	3.48	2.83	2.3	2.53	2.98
Incremental Costs (£)	£61,124	£84,078	£74,849	£40,325	£54,088

ICER (£/QALY)	£17,552	£29,757	£32,540	£15,962	£18,167
Height gain (cm)	32.24	7.95	25.59	4.48	21.92
Cost per cm gain	£1,896	£10,576	£2,925	£9,001	£2,467

Table 60 Base case results for the previous growth hormone HTA [5]

	GHD	TS	PWS†	CRI
Incremental Cost	£53,373	£61,770	£56,663	£54,009
Height gain (cm)*	8.85	3.9	1.36	7.29
Cost per cm gain	£6,029	£15,997	£40,815	£7,403

* Discounted and adjusted for drop-outs; † Height gain expressed in terms of HtSDS gained.

The cost effectiveness results from the SHTAC model for rhGH treatment vary widely between conditions, from £25,480 for GHD to £148,860 for PWS per QALY gained. The ICERs for TS, CRI and SGA and SHOX-D were between £35,000 and £45,000 per QALY gained. This indicates that rhGH is unlikely to be cost effective for TS, PWS, CRI, SGA and SHOX-D at a willingness to pay threshold of £20,000 to £30,000. However the results were sensitive to the discount rate used. All conditions, except PWS, would be cost effective at a £30,000 willingness to pay threshold using the previous NICE discount rate of 6% for costs and 1.5% for benefits. For all the conditions, the model results are most sensitive to treatment start age and length, compliance and utility gain.

The cost effectiveness results in the current report varied from those in the MS and the previous HTA report. The incremental costs reported are generally consistent between the three models, with slight variations due to different dose, cost and treatment start age and duration. In general, the results, presented in terms of cm gained, are more favourable in the current analyses compared to the previous HTA report. This is due to higher estimates in height gain and lower incremental costs in the current report. The height gains in the MS for GHD, PWS and SGA appear extremely high and inconsistent with those found in the review of clinical effectiveness. The ICERs in the MS are considerably more favourable than the current analysis, due to higher estimates of utility gain. The current analyses and the MS have chosen utility estimates from the same study[150]. However the manufacturers have not taken these values from the regression analysis from this study. Instead they have used the relationship between EQ-5D and height without controlling for other factors.

In general, the incremental costs consist primarily of the rhGH drug costs, while other costs have little effect on model results. For the cost effectiveness results, the key issue is the choice of utility values. The utility gain from rhGH is assumed to last over the patients' lifetimes and hence most of the QALY gain is in adulthood.

The results were sensitive to the length of treatment, for example by treating children from an earlier age. Current best practice is usually regarded as treating children as early as possible and this is likely to mean a longer treatment duration, which increases the cost of treatment and thus the ICER. It is unclear whether there will be an associated extra increase in height as most of the RCTs followed up children for a short time period, for less than three years. The previous HTA report suggested that height gains were greatest in the first year or two of treatment but stopping treatment before achieving final height generally leads to loss of growth gains, and so should not be advised.

The results were sensitive to the clinical effect. The treatment effect has been obtained, where possible, from the best quality RCT available. However, as indicated in Section 3.2, these trials were generally of poor quality and were not long term trials. We also used the clinical treatment effect from the KIGS observational study but the results were largely similar to those reported from the RCTs.

There are limitations to the QoL estimates used in the model. There was a lack of good QoL studies conducted in the conditions of interest. Therefore, evidence based on these studies was not used in the main analysis. The utility estimates were based upon a study which estimated utility in the general adult population according to height. The study provides a common utility gain that could be compared across all the conditions of interest. Furthermore, it also provided the possibility the outcomes from the RCTs identified in the clinical effectiveness could be used. However, this still remains a major source of uncertainty in the model.

The QoL gains were highest for individuals with lower starting height; for those with starting height of less than <-2 HtSDS the QoL gain was minimal. For example those with PWS had a starting height of -2 HtSDS, and so for this group of patients the health gain is small and therefore rhGH has high ICER values compared to no treatment. PWS patients may experience an improvement in body composition due to rhGH but this was difficult to quantify, especially in the long term, due to lack of long term data.

The current analysis assumes in the base case that all children with the conditions of interest will have reduced life expectancy. This was based upon some evidence to suggest that these children would have a lower life expectancy due to increased risk of cardiovascular disease, due to abdominal obesity and raised blood pressure. Furthermore, those children with CRI have a much reduced life expectancy. We have used the end stage renal disease mortality rates as a proxy in the absence of any available data for CRI. This may underestimate life expectancy, and overestimate ICER values, as not all CRI patients will go on to develop end stage renal disease. Bengtsson and colleagues[173] suggest that rhGH can rectify most of the cardiovascular abnormalities associated with GHD although there appear to be few long term observational studies which confirm this claim. Therefore, we assumed that rhGH will not increase life expectancy.

Apart from as a scenario analysis for PWS, the current analysis has not considered other benefits in addition to height gain within the model. The base case does not include possible benefits from changes in body composition such as reduced risk of diabetes or cardiovascular disease, which may even result in increases in life expectancy. At this stage, these health gains would be purely speculative and it is not possible to verify if they exist or quantify them. It is also possible that there may be additional psychological benefits such as improved self esteem.

Strengths and limitations of the assessment

Strengths

- The systematic review and economic evaluation were carried out independently, with no vested interest, and results are presented in a consistent and transparent manner.
- Evidence for clinical effectiveness came from RCT data, considered to be the highest level of evidence.
- The project followed established methodology and principles for conducting a systematic review. The methods used were defined *a priori* in a research protocol (Appendix 1), and this was circulated to clinical experts and agreed with NICE before the project started.
- A clinical advisory group reviewed and commented on drafts of the protocol and the final report.
- A de novo economic model was developed following recognised guidelines.

Limitations and uncertainties

- As specified in the protocol, the systematic review was restricted to RCTs, because these provide the highest level of evidence for clinical effectiveness. The majority of the studies included in this review lasted for between six months and two years, with very few continuing long term or to adult height. Many of the trials excluded patients from analyses due to incomplete follow-up data or patient withdrawal. The short duration of the RCTs means it is difficult to assess its effectiveness in the context in which it would be prescribed in real life, i.e. for many years in some cases.
- None of the RCTs included in this review reported any assessment of QoL issues, and the literature has conflicting conclusions regarding the effect of short stature on QoL. It is therefore difficult to make any judgement about the impact of rhGH on the quality of a person's daily life. Many of the children with the health conditions covered in this review will have a variety of other physical problems. Whilst rhGH treatment can help to improve growth, height and body composition to some extent, QoL issues associated with underlying health problems will continue to affect some children.
- We did not identify any RCTs which met the original inclusion criteria for children born SGA. Following discussion with NICE, we therefore amended the criteria as detailed in Section 3.1.2. The main difference was that we included studies of children who failed to show catch up growth by three years of age (rather than four), but did not specify exact criteria for this. Although this will have allowed slightly younger children to be included, the evidence presented in this report is still relevant to the UK SGA population. We also removed the reference to parental height, so it is possible that children in the included trials were naturally shorter than those in the general population. Only one of the included trials used the licensed dose, so results from the other five could over-state the effectiveness of rhGH treatment for this patient group.
- We only found one RCT of rhGH in children with GHD, so the evidence base for this condition is rather weak. However, the previous Health Technology Assessment report also included observational studies for GHD, TS, PWS and CRI. Non-randomised evidence for this condition has therefore been summarised previously in the literature and is publicly available.
- The included trials were generally poorly reported, and often had low numbers of participants. Primary outcomes were not clearly specified, and few studies reported power calculations. It is therefore possible that some trials were underpowered to detect 'real'

differences between the treatment groups, even where such differences were reported to be statistically significant.

- The included studies were heterogeneous in terms of participants, dosages and study duration. The results are therefore presented as a narrative summary, and it was not appropriate to meta-analyse the data.
- The economic model used the suggested doses given in the BNF. However, the RCTs used doses which were sometimes outside the licensed doses.

7 CONCLUSIONS

7.1 Implications for service provision

NICE guidance already recommends treatment with rhGH for children who have short stature associated with GHD, TS, PWS and CRI, so prescriptions associated with these conditions are already in place. However, possible changes to practice include a shift towards managing children in tertiary centres jointly with either local paediatricians, or sometimes GPs. Clinical opinion indicates that there may be a trend towards earlier prescribing for PWS, and many families are now seeking treatment in infancy rather than in mid-childhood. There may also be an increase in treatment associated with acquired GHD as the proportion of children surviving cancers and associated treatment increases.

The newly licensed conditions SHOX-D and SGA are not yet covered by NICE guidance. Of the estimated 4758 UK patients currently receiving rhGH, only approximately 5% were receiving treatment for short stature associated with being born SGA. Clinical opinion indicates that there is unlikely to be a large increase in prescriptions for SGA children, particularly if assessment were to be undertaken in tertiary centres.

It is not clear how many children with SHOX-D are currently receiving treatment. The availability of prescriptions to these new groups of patients could theoretically have a budgetary impact. However, the number of children with this condition is small so there is unlikely to be a large increase in prescriptions.

7.2 Suggested research priorities

- There is a lack of RCT evidence for the effects of rhGH treatment on final height, since it is impractical to run such long studies. However, longer studies beyond two years would be helpful in improving the evidence base for long term treatment, even if near-final height rather than final adult height were reported.
- None of the included RCTs reported measures of health related QoL. There is a need to develop and validate a standardised QoL assessment specifically designed for children and adults. Future RCTs should include this as an outcome measure in order to assess the impact of small increases in height on daily QoL. This would also be helpful for developing utilities for cost effectiveness analysis of rhGH treatment for these conditions.
- Good quality trials of continuation/ discontinuation of rhGH in children who have finished growing are required, that report consistent and clinically relevant outcomes, and that are standardised in terms of dose. Consensus on the most appropriate location for transition care service provision would also be helpful.
- Good quality trials are needed of GH in children born SGA, where the children included and the dose administered match the licensing criteria.
- It was difficult to establish when treatment is initiated for the different disease areas, as this depends on age at diagnosis. Further work to survey national practices or policies would be helpful in terms of providing information for future updates of this review and economic evaluation.
- Although figures for the use of renal replacement therapy are available, there is little epidemiological data available on the incidence and prevalence of CRI. Epidemiological studies would therefore be useful.
- Good quality observational studies are needed which show the long term effects of rhGH, particularly the effect of treatment on body composition, psychological benefits such as improved self esteem, and long term morbidities such as diabetes or cardiovascular disease, and life expectancy, particularly for PWS.
- Further research is also necessary to establish the QoL benefits associated with rhGH in individuals with these conditions in children and adults.
- Monitoring of AE associated with long-term rhGH treatment is required, with a central register to record the effects of long-term elevations in IGF-I levels.
- More research is needed to assess the long term effect on QoL for individuals who had rhGH as children.

8 REFERENCES

1. Kirk J, Butler G, on behalf of the British Society for Paediatric Endocrinology and Diabetes (BSPED). Treatment of children with recombinant human growth hormone (r-hGH): Shared Care Guidelines. 2006.

2. Rosenfeld, R. G. Editorial: Is growth hormone deficiency a viable diagnosis? Journal of Clinical Endocrinology & Metabolism 82[2], 349-351. 1997.

3. Juul, A., Bernasconi, S., Clayton, P. E., Kiess, W., DeMuink-Keizer Schrama, S., and Drugs and Therapeutics Committee of the European Society for Paediatric Endocrinology (ESPE). European audit of current practice in diagnosis and treatment of childhood growth hormone deficiency. Hormone Research 58[5], 233-241. 2002.

4. Warner, J. and British Society for Paediatric Endocrinology and Diabetes. Submission of Evidence for the National Institute of Clinical Excellence Health Technology Appraisal: The Use of Human Growth Hormone in Children (review). 2008.

5. Bryant J, Cave C, Mihaylova B, Chase D, McIntyre L, Gerard K *et al*. Clinical effectiveness and cost-effectiveness of growth hormone in children: a systematic review and economic evaluation. *Health Technology Assessment (Winchester, England)* 2002;6:1-168.

6. UK Child Growth Foundation. Growth Hormone Deficiency. http://www.childgrowthfoundation.org/ghd.htm *http://www.childgrowthfoundation.org/ghd.htm* (accessed 12 May 2008)

7. Thomas M, Massa G, Craen M, de ZF, Bourguignon JP, Heinrichs C *et al*. Prevalence and demographic features of childhood growth hormone deficiency in Belgium during the period 1986-2001. *European Journal of Endocrinology* 2004;151:67-72.

8. Stochholm K, Gravholt CH, Laursen T, Jorgensen JO, Laurberg P, Andersen M *et al*. Incidence of GH deficiency - A nationwide study. *European Journal of Endocrinology* 2006;155:61-71.

9. Turner H. A syndrome of infantilism, congenital webbed neck, and cubitus valgus. *Endocrinology* 1938;23:566-74.

10. Bondy CA. Clinical practice guideline - Care of girls and women with Turner syndrome: A guideline of the Turner Syndrome Study Group. *Journal of Clinical Endocrinology & Metabolism* 2007;92:10-25.

11. Quigley CA, Crowe BJ, Anglin DG, Chipman JJ. Growth hormone and low dose estrogen in Turner syndrome: results of a United States multi-center trial to near-final height. *Journal of Clinical Endocrinology & Metabolism* 2002;87:2033-41.

12. Rao E, Weiss B, Fukami M, Rump A, Niesler B, Mertz A *et al*. Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nature Genetics* 1997;16:54-63.

13. Gravholt CH. Epidemiological, endocrine and metabolic features in Turner syndrome. *European Journal of Endocrinology* 2004;151:657-87.

14. WHO Collaborating Centre for the Epidemiological Surveillance of Congenital Anomalies. Report 2004-2005. http://www.eurocat.ulster.ac.uk/pdf/EUROCAT-Annual-Report-2005-for-WHO.pdf *http://www.eurocat.ulster.ac.uk/pdf/EUROCAT-Annual-Report-2005-for-WHO.pdf* (accessed 22 April 9 A.D.)

15. Massa G, Verlinde F, De SJ, Thomas M, Bourguignon JP, Craen M *et al*. Trends in age at diagnosis of Turner syndrome. *Archives of Disease in Childhood* 2005;90:267-8.

16. Stochholm K, Juul S, Juel K, Naeraa RW, Gravholt CH. Prevalence, incidence, diagnostic delay, and mortality in Turner syndrome. *Journal of Clinical Endocrinology and Metabolism* 2006;91:3897-902.

17. Paterson WF, Donaldson MD. Growth hormone therapy in the Prader-Willi syndrome. [Review] [24 refs]. *Archives of Disease in Childhood* 2003;88:283-5.

18. Cassidy SB. Prader-Willi syndrome. Characteristics, management, and etiology. *Alabama Journal of Medical Sciences* 1987;24:169-75.

19. Chen C, Visootsak J, Dills S, Graham JM. Prader-Willi syndrome: an update and review for the primary pediatrician. *Clinical Pediatrics* 2007;46:580-91.

20. Butler M. Prader-Willi syndrome: current understanding of cause and diagnosis. *American Journal of Medical Genetics* 1990;35:319-32.

21. Carrel AL, Moerchen V, Myers SE, Bekx MT, Whitman BY, Allen DB. Growth hormone improves mobility and body composition in infants and toddlers with Prader-Willi syndrome. *Journal of Pediatrics* 2004;145:744-9.

22. Whittington JE, Holland AJ, Webb T, Butler J, Clarke D, Boer H. Population prevalence and estimated birth incidence and mortality rate for people with Prader-Willi syndrome in one UK Health Region [7]. *Journal of Medical Genetics* 2001;38:792-8.

23. Sarborio P, Hahn S, Hisano S, Latta K, Scheinman J, Chan J. Chronic renal failure: and overview from a pediatric perspective. *Nephron* 1998;80:134-48.

24. Tönshoff B, Mehls O. Growth retardation in children with chronic renal insufficiency: current aspects of pathophysiology and treatment. *Journal of Nephrology* 1995;8:133-42.

25. Schaefer F, Wingen AM, Hennicke M, Rigden S, Mehls O. Growth charts for prepubertal children with chronic renal failure due to congenital renal disorders. European Study Group for Nutritional Treatment of Chronic Renal Failure in Childhood. *Pediatr Nephrol* 1996;10:288-93.

26. Mahan JD, Warady BA, Fielder P, Gipson DS, Greenbaum L, Juarez-Congelosi MD *et al*. Assessment and treatment of short stature in pediatric patients with chronic kidney disease: A consensus statement. *Pediatric Nephrology* 2006;21:917-30.

27. National Kidney Foundation Kidney Disease Outcomes Quality Initiative. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Part 4. Definition and classification of stages of chronic kidney disease. http://www.kidney.org/PROFESSIONALS/kdoqi/guidelines_ckd/toc.htm *http://www.kidney.org/PROFESSIONALS/kdoqi/guidelines_ckd/toc.htm* (accessed 23 February 9 A.D.)

28. Kari J, Gonzalez C, Ledermann S, Shaw V, Rees L. Outcome and growth of infants with severe chronic renal failure. *Kidney International* 2000;57:1681-7.

29. Parekh R, Flynn J, Smoyer W, Milne J, Kershaw D, Bunchman T *et al*. Improved growth in young children with severe chronic renal insufficiency who use specified nutritional therapy. *J Am Soc Nephrol* 2001;12:2418-26.

30. Wingen AM, Mehls O. Nutrition in children with preterminal chronic renal failure. Myth or important therapeutic aid? *Pediatr Nephrol* 2002;17:111-20.

31. National Kidney Foundation Kidney Disease Outcomes Quality Initiative. Clinical practice guidelines for nutrition in chronic renal failure. http://www.kidney.org/PROFESSIONALS/kdoqi/guidelines_updates/doqi_nut.html *http://www.kidney.org/PROFESSIONALS/kdoqi/guidelines_updates/doqi_nut.html* (accessed 23 February 9 A.D.)

32. Furth SL. Growth and nutrition in children with chronic kidney disease. *Advances in Chronic Kidney Disease* 2005;12:366-71.

33. Lewis M, Shaw J, Reid C, Evans J, Webb N, Verrier-Jones K. Growth in children with established renal failure - a registry analysis. In: Ansell D, Feest T, Tomson C, Williams A, Warwick G, editors. UK Renal Registry Report 2006 Bristol, UK: UK Renal Registry 2006. p.243-8.

34. Groothof J, Offringa M, Van Eck-Smit B, Gruppen M, Van der Kar N, Wolff ED *et al*. Severe bone disease and low bone mineral density after juvenile renal failure. *Kidney International* 2003;63:266-75.

35. Rizzoni G, Broyer M, Brunner F, Brynger H, Challah S, Kramer P *et al*. Combined report on regular dialysis and transplantaion of children in Europe, XIII, 1983. *Proceedings of the European Dialysis and Transplant Association - European Renal Association* 1985;21:95.

36. Lewis M, Shaw J, Reid C, Evans J, Webb N, Verrier-Jones K. Demography and management of childhood established renal failure in the UK. In: Ansell BM, Feest T, Tomson C, Williams A, Warwick G, editors. UK Renal Registry Report 2006 Bristol, UK: UK Renal Registry 2006. p.227-41.

37. Royal College of Obstetricians and Gynaecologists. The investigation and management of the small-for-gestational-age fetus Guideline No. 31. 2002.

38. Lee PA, Chernausek SD, Hokken-Koelega ACS, Czernichow P. International small for gestational age advisory board consensus development conference statement:

management of short children born small for gestational age, April 24- October 1, 2001. *Pediatrics* 2003;111:1253-61.

39. Karlberg J, Albertsson-Wikland K, Kwan EY, Lam B, Low LC. The timing of early postnatal catch-up growth in normal, full-term infants born short for gestational age. *Hormone Research* 1997;48 (suppl 1):17-24.

40. Ong K, Ahmed M, Emmett P, Preece M, Dunger DB, for the Avon Longitudinal Study of Pregnancy and Childhood Study Team. Association between post-natal catch-up growth and obesity in childhood: prospective cohort study. *BMJ* 2000;320:967-71.

41. Albertsson-Wikland K, Karlberg J. Postnatal growth of children born small for gestational age. *Acta Paediatrica Suppl* 1997;423:193-5.

42. Huiming, Y, Meng, M, Chaomin, W, and Fan, Y. Growth Hormone Treatment for short stature in children with intrauterine growth retardation (protocol). Cochrane Database of Systematic Reviews [4]. 2006.

43. Karlberg J, Albertsson-Wikland K. Growth in full term small-for-gestational-age infants: from birth to final height. *Pediatr Res* 1995;38:733-9.

44. Tuvemo T, Cnattingius S, Jonsson B. Prediction of male adult stature using anthropometric data at birth: a nationwide population-based study. *Pediatr Res* 1999;46:491-5.

45. Albertsson-Wikland K, Karlberg J. Natural growth in children born small for gestational age with and without catch-up growth. *Acta Paediatrica Suppl* 1994;399:64-70.

46. Karlberg J, Albertsson-Wikland K. Growth in full term small-for-gestational-age infants: from birth to final height. *Pediatr Res* 1995;38:733-9.

47. Hokken-Koelega A, De Ridder M, Lemmen R, Den Hartog H, De Muinck Keizer-Schrama S, Drop S. Children born small for gestational age: do they catch up? *Pediatr Res* 1995;38:267-71.

48. Blum WF, Crowe BJ, Quigley CA, Jung H, Cau D, Ross JL *et al*. Growth hormone is effective in treatment of short stature associated with Short Stature Homeobox-containing gene deficiency: two year results of a randomized, controlled, multicenter trial. *The Journal of clinical endocrinology and metabolism* 2007;92:219-28.

49. Eli Lilly. Lilly and Humatrope: A lasting commitment. 2006. USA.

50. Langer L. Mesomelic dwarfism of the hypoplastic ulna, fibula, mandible type. *Radiology* 1967;89:660.

51. Zinn AR, Wei F, Zhang L, Elder FF, Scott CI, Jr., Marttila P *et al*. Complete SHOX deficiency causes Langer mesomelic dysplasia. *American Journal of Medical Genetics* 2002;110:158-63.

52. Binder G, Schwarze CP, Ranke MB. Identification of short stature caused by SHOX defects and therapeutic effect of recombinant human growth hormone. *J Clin Endocrinol Metab* 2000;85:245-9.

53. Ezquieta B, Cueva E, Oliver A, Gracia R. SHOX intragenic microsatellite analysis in patients with short stature. *J Pediatr Endocrinol Metab* 2002;15:139-48.

54. Ogata T, Muroya K, Sasaki G, Nishimura G, Kitoh H, Hattori T. SHOX nullizygosity and haploinsufficiency in a Japanese family: implication for the development of Turner skeletal features. *J Clin Endocrinol Metab* 2002;87:1390-4.

55. Rappold GA, Fukami M, Niesler B, Schiller S, Zumkeller W, Bettendorf M *et al*. Deletions of the homeobox gene SHOX (short stature homeobox) are an important cause of growth failure in children with short stature. *Journal of Clinical Endocrinology & Metabolism* 2002;87:1402-6.

56. Binder G, Ranke MB, Martin DD. Auxology is a valuable instrument for the clinical diagnosis of SHOX haploinsufficiency in school-age children with unexplained short stature. *J Clin Endocrinol Metab* 2003;88:4891-6.

57. Stuppia L, Calabrese G, Gatta V, Pintor S, Morizio E, Fantasia D *et al*. SHOX mutations detected by FISH and direct sequencing in patients with short stature. *J Med Genet* 2003;40:E11.

58. Huber C, Rosilio M, Munnich A, Cormier-Daire V. High incidence of SHOX anomalies in individuals with short stature. *Journal of Medical Genetics* 2006;43:735-9.

59. Binder G, Renz A, Martinez A, Keselman A, Hesse V, Riedl SW *et al*. SHOX haploinsufficiency and Leri-Weill dyschondrosteosis: Prevalence and growth failure in relation to mutation, sex, and degree of wrist deformity. *Journal of Clinical Endocrinology and Metabolism* 2004;89:4403-8.

60. Allen DB. Growth hormone therapy for short stature: is the benefit worth the burden? *Pediatrics* 2006;118:343-8.

61. Munns C, Glass I, Flanagan S, Hayes M, Williams B, Berry M *et al*. Familial growth and skeletal features associated with SHOX haploinsufficiency. *Journal of Pediatric Endocrinology and Metabolism* 2003;16:987-96.

62. Voss LD, Mulligan J. Bullying at school: are short pupils at risk? Questionnaire study in a cohort. *BMJ* 2000;320:612-3.

63. Tanaka, T., Cohen, P., Clayton, PE., Laron, Z., Hintz, R. L., and Sizonenko, PC. Diagnosis and management of growth hormone deficiency in childhood and adolescence - Part 2: Growth hormone treatment in growth hormone deficient children. Growth hormone & IGF research : official journal of the Growth Hormone Research Society and the International IGF Research Society 12, 323-341. 2002.

64. Rosén T, Wilhelmsen L, Bengtsson BA. Altered lipid pattern explains increased cardiovascular mortality in hypopituitary patients with growth hormone deficiency. *Clinical Endocrinology* 1998;48:525-6.

65. Elsheikh M, Dunger DB, Conway GS, Wass JAH. Turner's Syndrome in adulthood. *Endocrine Reviews* 2002;23:120-40.

66. Höybe C, Hilding A, Jacobsson H, Thorén M. Metabolic profile and body composition in adults with Prader-Willi syndrome and severe obesity. *Journal of Clinical Endocrinology & Metabolism* 2002;87:3590-7.

67. Reinehr T, Kleber M, Toschke A. Small for gestational age (SGA) status is associated with metabolic syndrome in overweight children. *European Journal of Endocrinology* 2009;160:579-84.

68. Barker D. *Mothers, babies, and disease in later life*. London: British Medical Journal Publishing Group, 1998.

69. National Institute for Health and Clinical Excellence National Collaborating Centre for Primary Care. Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children. 2006. No. Clinical Guideline 43

70. Rajaram S, Baylink D, Mohan S. Insulin-like growth factor-binding proteins in serum and other biological fluids: regulation and functions. *Endocrine Reviews* 1997;18:801-31.

71. Cohen P, Peehl D, Lamson G, Rosenfeld R. Insulin-like growth factors (IGFs), IGF receptors, and IGF-binding proteins in primary cultures of prostate epithelial cells. *Journal Clinical Endocrinology and Metabolism* 1991;73.

72. Lippman M. The development of biological therapies for breast cancer. *Science* 1993;259:631-2.

73. Kirk J, Clayton P. Specialist services and transitional care in pediatric endocrinology in the UK and Ireland. *Clinical Endocrinology* 2006;65:59-63.

74. National Institute of Health and Clinical Evidence. Guidance on the use of human growth hormone (somatropin) in children with growth failure. 2002. No. Technology Appraisal Guidance No. 42

75. National Kidney Foundation. KDOQI Clinical practice guidelines for bone metabolism and disease in children with chronic kidney disease. http://www.kidney.org/professionals/KDOQI/guidelines_pedbone/guide11.htm

76. Clayton PE, Cianfarani S, Czernichow P, Johannsson G, Rapaport R, Rogol A. Consensus statement: Management of the child born small for gestational age through to adulthood: A consensus statement of the International Societies of Pediatric Endocrinology and the Growth Hormone Research Society. *Journal of Clinical Endocrinology & Metabolism* 2007;92:804-10.

77. Pfizer. Genotropin full prescribing information. internet *http://www.accessdata.fda.gov/drugsatfda_docs/label/2008/020280s060lbl.pdf* (accessed 29 July 2009)

78. Eli Lilly. Humatrope 6mg, 12mg, or 24mg powder and solvent for solution for injection. internet *http://emc.medicines.org.uk/medicine/601/SPC/Humatrope%206mg,%2012mg,%20or%2024mg%20powder%20and%20solvent%20for%20solution%20for%20injection/* (accessed 2 June 9 A.D.)

79. Sandoz. Omintrope Summary of Product Characteristics. internet *http://www.emea.europa.eu/humandocs/PDFs/EPAR/Omnitrope/H-607-PI-en.pdf* (accessed 2 June 9 A.D.)

80. Eli Lilly and Company Ltd. Humatrope 6mg, 12mg, or 24mg powder and solvent for solution for injection. http://emc.medicines.org.uk *http://emc.medicines.org.uk* (accessed 23 December 8 A.D.)

81. GH Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH research society. *Journal of Clinical Endocrinology & Metabolism* 2000;85:3990-3.

82. Grunfeld C, Thompson M, Brown SJ, Richmond G, Lee D, Muurahainen N *et al*. Recombinant human growth hormone to treat HIV-associated adipose redistribution syndrome - 12-Week induction and 24-week maintenance therapy. *Jaids-Journal of Acquired Immune Deficiency Syndromes* 2007;45:286-97.

83. NHS Centre for Reviews and Dissemination. Undertaking Systematic Reviews of Research on Effectiveness. CRD Report 4 (2nd Edition). 2001. York, University of York.

84. Soliman AT, abdul Khadir MM. Growth parameters and predictors of growth in short children with and without growth hormone (GH) deficiency treated with human GH: a randomized controlled study. *Journal of Tropical Pediatrics* 1996;42:281-6.

85. Davenport ML, Crowe BJ, Travers SH, Rubin K, Ross JL, Fechner PY *et al*. Growth hormone treatment of early growth failure in toddlers with Turner syndrome: a randomized, controlled, multicenter trial. *Journal of Clinical Endocrinology & Metabolism* 2007;92:3406-16.

86. Stephure DK, Canadian Growth Hormone Advisory Committee. Impact of growth hormone supplementation on adult height in turner syndrome: results of the Canadian randomized controlled trial. *Journal of Clinical Endocrinology & Metabolism* 2005;90:3360-6.

87. Rovet J, Holland J. Psychological aspects of the Canadian randomized controlled trial of human growth hormone and low-dose ethinyl oestradiol in children with Turner syndrome. The Canadian Growth Hormone Advisory Group. *Hormone Research* 1993;39 Suppl 2:60-4.

88. Gravholt CH, Naeraa RW, Brixen K, Kastrup KW, Mosekilde L, Jorgensen JO *et al*. Short-term growth hormone treatment in girls with Turner syndrome decreases fat mass and insulin sensitivity: a randomized, double-blind, placebo-controlled, crossover study. *Pediatrics* 2002;110:889-96.

89. Gravholt CH, Hjerrild BE, Naeraa RW, Engbaek F, Mosekilde L, Christiansen JS. Effect of growth hormone and 17beta-oestradiol treatment on metabolism and body composition in girls with Turner syndrome. *Clinical Endocrinology* 2005;62:616-22.

90. Johnston DI, Betts P, Dunger D, Barnes N, Swift PG, Buckler JM *et al*. A multicentre trial of recombinant growth hormone and low dose oestrogen in Turner syndrome: near final height analysis. *Archives of Disease in Childhood* 2001;84:76-81.

91. Festen DA, van TA, Duivenvoorden HJ, Hokken-Koelega AC. Adiponectin levels in prepubertal children with Prader-Willi syndrome before and during growth hormone therapy. *Journal of Clinical Endocrinology & Metabolism* 2007;92:1549-54.

92. Festen DAM, Wevers M, Lindgren AC, Bohm B, Otten BJ, Wit JM *et al*. Mental and motor development before and during growth hormone treatment in infants and toddlers with Prader-Willi syndrome. *Clinical Endocrinology* 2008;68:919-25.

93. de Lind van Wijngaarden Roderick FA, de Klerk L, Festen D, Duivenvoorden HJ, Otten B, Hokken-Koelega ACS. Randomized controlled trial to investigate the effects of growth hormone treatment on scoliosis in children with Prader-Willi syndrome. *J Clin Endocrinol Metab* 2009;94:1274-80.

94. Festen DAM, de Lind van Wijngaarden R, van Eekelen M, Otten BJ, Wit JM, Duivenvoorden HJ *et al*. Randomized controlled GH trial: effects on anthropometry, body composition and body proportions in a large group of children with Prader-Willi syndrome. *Clinical Endocrinology* 2008;69:443-51.

95. Carrel AL, Myers SE, Whitman BY, Allen DB. Growth hormone improves body composition, fat utilization, physical strength and agility, and growth in Prader-Willi syndrome: A controlled study. *Journal of Pediatrics* 1999;134:215-21.

96. Myers SE, Carrel AL, Whitman BY, Allen DB. Physical effects of growth hormone treatment in children with Prader-Willi syndrome. *Acta Paediatrica Supplement* 1999;88:112-4.

97. Myers SE, Whitman BY, Carrel AL, Moerchen V, Bekx MT, Allen DB. Two years of growth hormone therapy in young children with Prader-Willi syndrome: Physical and neurodevelopmental benefits. *American Journal of Medical Genetics, Part A* 2007;143:443-8.

98. Whitman B, Carrel A, Bekx T, Weber C, Allen D, Myers S. Growth hormone improves body composition and motor development in infants with Prader-Willi syndrome after six months. *Journal of Pediatric Endocrinology* 2004;17:591-600.

99. Hauffa BP. One-year results of growth hormone treatment of short stature in Prader-Willi syndrome. *Acta Paediatrica Supplement* 1997;423:63-5.

100. Lindgren AC, Hagenas L, Muller J, Blichfeldt S, Rosenborg M, Brismar T *et al*. Effects of growth hormone treatment on growth and body composition in Prader-Willi syndrome: a preliminary report. The Swedish National Growth Hormone Advisory Group. *Acta Paediatrica Supplement* 1997;423:60-2.

101. Lindgren AC, Hagenas L, Muller J, Blichfeldt S, Rosenborg M, Brismar T *et al*. Growth hormone treatment of children with Prader-Willi syndrome affects linear growth and body composition favourably. *Acta Paediatrica* 1998;87:28-31.

102. Haqq AM, Stadler DD, Jackson RH, Rosenfeld RG, Purnell JQ, LaFranchi SH. Effects of growth hormone on pulmonary function, sleep quality, behavior, cognition, growth velocity, body composition, and resting energy expenditure in Prader-Willi syndrome. *Journal of Clinical Endocrinology & Metabolism* 2003;88:2206-12.

103. Sanchez CP, Kuizon BD, Goodman WG, Gales B, Ettenger RB, Boechat MI *et al*. Growth hormone and the skeleton in pediatric renal allograft recipients. *Pediatric Nephrology* 2002;17:322-8.

104. Hokken-Koelega ACS, Stijnen T, De Muinck Keizer-Schrama SMPF, Wit JM, Wolff ED, De Jong MCJW *et al*. Placebo-controlled, double-blind, cross-over trial of growth hormone treatment in prepubertal children with chronic renal failure. *Lancet* 1991;338:585-90.

105. Hokken-Koelega AC, Stijnen T, de Jong RC, Donckerwolcke RA, Groothoff JW, Wolff ED *et al*. A placebo-controlled, double-blind trial of growth hormone treatment in prepubertal children after renal transplant. *Kidney International - Supplement* 1996;53:S128-S134.

106. Powell DR, Liu F, Baker BK, Hintz RL, Lee PD, Durham SK *et al*. Modulation of growth factors by growth hormone in children with chronic renal failure. The Southwest Pediatric Nephrology Study Group. *Kidney International* 1997;51:1970-9.

107. Broyer M. Results and side-effects of treating children with growth hormone after kidney transplantation - a preliminary report. Pharmacia & Upjohn Study Group. *Acta Paediatrica Supplement* 1996;417:76-9.

108. Fine RN, Kohaut MEC, Brown MD, Perlman AJ. Growth after recombinant human growth hormone treatment in children with chronic renal failure: Report of a multicenter randomized double-blind placebo-controlled study. *Journal of Pediatrics* 1994;124:374-82.

109. De SJ, Thomas M, Beckers D, Craen M, Maes M, de ZF. Growth hormone treatment and fat redistribution in children born small for gestational age. *Journal of Pediatrics* 2008;152:327-30.

110. Lagrou K, Froidecoeur C, Thomas M, Massa G, Beckers D, Craen M *et al*. Concerns, expectations and perception regarding stature, physical appearance and psychosocial functioning before and during high-dose growth hormone treatment of short pre-pubertal children born small for gestational age. *Hormone Research* 2008;69:334-42.

111. Carel JC, Chatelain P, Rochiccioli P, Chaussain JL. Improvement in adult height after growth hormone treatment in adolescents with short stature born small for gestational age: results of a randomized controlled study. *Journal of Clinical Endocrinology & Metabolism* 2003;88:1587-93.

112. de Zegher F., Maes M, Gargosky SE, Heinrichs C, Du Caju MVL, Thiry G *et al*. High-dose growth hormone treatment of short children born small for gestational age. *Journal of Clinical Endocrinology and Metabolism* 1996;81:1887-92.

113. de Zegher F., Ong K, van HM, Mohn A, Woods K, Dunger D. High-dose growth hormone (GH) treatment in non-GH-deficient children born small for gestational age induces growth responses related to pretreatment GH secretion and associated with a reversible decrease in insulin sensitivity. *Journal of Clinical Endocrinology & Metabolism* 2002;87:148-51.

114. Phillip M, Lebenthal Y, Lebl J, Zuckerman-Levin N, Korpal-Szczyrska M, Marques JS *et al*. European multicentre study in children born small for gestational age with persistent short stature: comparison of continuous and discontinuous growth hormone treatment regimens. *Hormone Research* 2009;71:52-9.

115. Cutfield W, Lindberg A, Chatelain P, Ranke M, Wilton P. Final height in idiopathic growth hormone deficiency: the KIGS experience. *Acta Paediatrica Suppl* 1999;88:72-5.

116. August G, Julius J, Blethen S. Adult height in children with growth hormone deficiency. *Paediatrics* 1998;102:16.

117. The growth pattern and final height of girls with Turner syndrome with and without human growth hormone treatment. *Paediatrics* 1998;101:663-8.

118. Hochberg Z, Zadik Z. Final height in young women with Turner syndrome after GH therapy: an open controlled study.[see comment]. *European Journal of Endocrinology* 1999;141:218-24.

119. Pasquino AM, Passeri F, Municchi G, Segni M, Pucarelli I, Larizza D *et al*. Final height in Turner syndrome patients treated with growth hormone. *Horm Res* 1996;46:72.

120. Taback SP, Collu R, Deal CL, Guyda HJ, Salisbury S, Dean HJ *et al*. Does growth-hormone supplementation affect adult height in Turner's syndrome? *Lancet* 1996;348:25-7.

121. Haffner D, Schaefer F, Nissel R, Wuhl E, Tonshoff B, Mehls O. Effect of growth hormone treatment on the adult height of children with chronic renal failure. German Study Group for Growth Hormone Treatment in Chronic Renal Failure. *New England Journal of Medicine* 2000;343:923-30.

122. Janssen F, Van Damme L, Van Dyck M, Hall M, Schurmans T, Herman J *et al*. Impact of growth hormone treatment on a Belgian population of short children with renal allografts. *Pediatr Transplant* 1997;1:6.

123. Angulo M, Castro-Magana M, Canas J, Arguello R, Lamerson M, Tapiador C *et al*. Final height in Prader-Willi syndrome (PWS) children treated with recombinant growth hormone (rhGH). *Pediatr Res* 2000;47.

124. Rosenfeld RG. Acceleration of growth in Turner syndrome patients treated with growth hormone: summary of three-year results. *Journal of Endocrinological Investigation* 1989;12:49-51.

125. Rosenfeld RG. Non-conventional growth hormone therapy in Turner syndrome: the United States experience. *Hormone Research* 1990;33:137-43.

126. Ross JL, Feuillan P, Kushner H, Roeltgen D, Cutler GB, Jr. Absence of growth hormone effects on cognitive function in girls with Turner syndrome.[see comment]. *Journal of Clinical Endocrinology & Metabolism* 1997;82:1814-7.

127. Whitman BY, Myers S, Carrel A, Allen D. A treatment/control group study of growth hormone treatment: Impact on behavior - A preliminary look. *Endocrinologist* 2000;10:31S-7S.

128. Conway GS, Szarras-Czapnik M, Racz K, Keller A, Chanson P, Tauber M *et al*. Treatment for 24 months with recombinant human GH has a beneficial effect on bone mineral density in young adults with childhood-onset GH deficiency. *European Journal of Endocrinology* 2009;160:899-907.

129. Mauras N, Pescovitz OH, Allada V, Messig M, Wajnrajch MP, Lippe B *et al*. Limited efficacy of growth hormone (GH) during transition of GH-deficient patients from adolescence to adulthood: a phase III multicenter, double-blind, randomized two-year trial. *Journal of Clinical Endocrinology & Metabolism* 2005;90:3946-55.

130. Attanasio AF, Shavrikova E, Blum WF, Cromer M, Child CJ, Paskova M *et al*. Continued growth hormone (GH) treatment after final height is necessary to complete somatic development in childhood-onset GH-deficient patients. *Journal of Clinical Endocrinology & Metabolism* 2004;89:4857-62.

131. Norrelund H, Vahl N, Juul A, Moller N, Alberti KG, Skakkebaek NE *et al*. Continuation of growth hormone (GH) therapy in GH-deficient patients during transition from childhood to adulthood: impact on insulin sensitivity and substrate metabolism. *Journal of Clinical Endocrinology & Metabolism* 2000;85:1912-7.

132. Underwood LE, Attie KM, Baptista J, Genentech Collaborative Study Group. Growth hormone (GH) dose-response in young adults with childhood-onset GH deficiency: a two-year, multicenter, multiple-dose, placebo-controlled study. *Journal of Clinical Endocrinology & Metabolism* 2003;88:5273-80.

133. Drake WM, Carroll PV, Maher KT, Metcalfe KA, Camacho-Hubner C, Shaw NJ *et al*. The effect of cessation of growth hormone (GH) therapy on bone mineral accretion in GH-deficient adolescents at the completion of linear growth. *Journal of Clinical Endocrinology & Metabolism* 2003;88:1658-63.

134. Baxter L, Bryant J, Cave CB, Milne R. Recombinant growth hormone for children and adolescents with Turner syndrome. *Cochrane Database of Systematic Reviews* 2007.

135. Li H, Banerjee S, Dunfield L, Kirby J, Jones M, Hamilton J et al. Recombinant human growth hormone for treatment of Turner Syndrome: systematic review and economic evaluation. Canadian Agency for Drugs and Technologies in Health (CADTH); 2007. No. 56

136. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313:275-83.

137. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C *et al*. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value Health* 2003;6:9-17.

138. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. 2008.

139. Joshi AV, Munro V, Russell MW. Cost-utility of somatropin (rDNA origin) in the treatment of growth hormone deficiency in children. *Current Medical Research & Opinion* 2006;22:351-7.

140. Anthony D, Stevens A. Growth hormone in children (for growth hormone deficiency, Turner's syndrome, chronic renal failure and idiopathic short stature) (CRD Record). 1996.

141. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes (3rd edition). *Oxford University Press* 2005.

142. Wit J-M, Boersma B, De Muinck Keizer-Schrama SMPF, Nienhuis HE, Oostdijk W, Otten BJ *et al*. Long-term results of growth hormone therapy in children with short stature, subnormal growth rate and normal growth hormone response to secretagogues. *Clinical Endocrinology* 1995;42:365-72.

143. Bannink EM, Raat H, Mulder PG, Keizer-Schrama SM. Quality of life after growth hormone therapy and induced puberty in women with Turner syndrome. *Journal of Pediatrics* 2006;148:95-101.

144. Bertella L, Mori I, Grugni G, Pignatti R, Ceriani F, Molinari E *et al*. Quality of life and psychological well-being in GH-treated, adult PWS patients: a longitudinal study. *Journal of Intellectual Disability Research* 2007;51:302-11.

145. Busschbach JJ, Rikken B, Grobbee DE, De Charro FT, Wit JM. Quality of life in short adults. *Horm Res* 1998;49:32-8.

146. Carel JC, Ecosse E, Bastie-Sigeac I, Cabrol S, Tauber M, Leger J *et al*. Quality of life determinants in young women with turner's syndrome after growth hormone treatment: results of the StaTur population-based cohort study. *J Clin Endocrinol Metab* 2005;90:1992-7.

147. Koltowska-Haggstrom MK. Growth hormone (GH) replacement in hypopituitary adults with GH deficiency evaluated by a utility-weighted quality of life index: A precursor to cost-utility analysis. *Clinical Endocrinology* 2008;68:Jan.

148. 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 comment]. *Journal of Clinical Endocrinology & Metabolism* 1998;83:1134-42.

149. Ara R, Brazier J. Deriving an Algorithm to Convert the Eight Mean SF-36 Dimension Scores into a Mean EQ-5D Preference-Based Score from Published Studies (Where Patient Level Data Are Not Available). *Value Health* 2008.

150. Christensen TL, Djurhuus CB, Clayton P, Christiansen JS. An evaluation of the relationship between adult height and health-related quality of life in the general UK population. *Clin Endocrinol (Oxf)* 2007;67:407-12.

151. Sproston K, Primatesta P. Health Survey for England 2003: Cardiovascular Disease. Volume 1: Cardiovascular disease. 2004.

152. Pfizer UK. KIGS (Pfizer Inernational growth study) database for Western Europe. in: The clinical and cost-effectiveness of growth hormone (Genotropin®) in children: Pfizer submission to NICE.; 2008.

153. Fryback DG, Lawrence WF, Martin PA, Klein R, Klein BE. Predicting Quality of Well-being scores from the SF-36: results from the Beaver Dam Health Outcomes Study. *Med Decis Making* 1997;17:1-9.

154. Shmueli A. The SF-36 profile and health-related quality of life: an interpretative analysis. *Qual Life Res* 1998;7:187-95.

155. Cooper K, Brailsford SC, Davies R, Raftery J. Modelling health care interventions. *J Oper Res Soc* 2007;58:168-76.

156. Kapoor RR, Burke SA, Sparrow SE, Hughes IA, Dunger DB, Ong KK *et al*. Monitoring of concordance in growth hormone therapy. *Arch Dis Child* 2008;93:147-8.

157. Bonfig W, Bechtold S, Bachmann S, Putzker S, Fuchs O, Pagel P *et al*. Reassessment of the optimal growth hormone cut-off level in insulin tolerance testing for growth hormone secretion in patients with childhood-onset growth hormone deficiency during transition to adulthood. *J Pediatr Endocrinol Metab* 2008;21:1049-56.

158. Briggs A, Sculpher M, Claxton K. *Decision modelling for health economic evaluation*. Oxford University Press, 2006.

159. Nielsen EH, Lindholm J, Laurberg P. Excess mortality in women with pituitary disease: a meta-analysis. *Clin Endocrinol (Oxf)* 2007;67:693-7.

160. Schoemaker MJ, Swerdlow AJ, Higgins CD, Wright AF, Jacobs PA. Mortality in women with turner syndrome in Great Britain: a national cohort study. *J Clin Endocrinol Metab* 2008;93:4735-42.

161. Lee PDK. Growth hormone and mortality in prader-willi syndrome. *Growth, genetics and hormones* 2006;22:17-23.

162. Groothoff JW. Long-term outcomes of children with end-stage renal disease. *Pediatric Nephrology* 2005;20:849-53.

163. Kajantie E, Osmond C, Barker DJ, Forsen T, Phillips DI, Eriksson JG. Size at birth as a predictor of mortality in adulthood: a follow-up of 350 000 person-years. *Int J Epidemiol* 2005;34:655-63.

164. Bryant J, Loveman E, Chase D, Mihaylova B, Cave C, Gerard K *et al*. Clinical effectiveness and cost-effectiveness of growth hormone in adults in relation to impact on quality of life: a systematic review and economic evaluation. *Health Technol Assess* 2002;6:1-106.

165. BMA, Royal Pharmaceutical Society of Great Britain. *British National Formularly 57*. London: BMJ Publishing Group, 2009.

166. Department of Health. NHS Reference Costs. UK: 2006.

167. Curtis L. Unit costs of health and social care. University of Kent, Canterbury: Personal Social Services Research Unit; 2008.

168. Contracting Unit SUHT. Unit costs. (unpublished database); 2008.

169. Picot J, Jones J, Colquitt J, Gospodaresvskaya E, Loveman E, Baxter L *et al*. The clinical-effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. *Health Technol Assess (In press)* 2009.

170. Hakim Z, Wolf A, and Garrison L.P. Estimating the effect of changes in body mass index on health state preferences. Pharmacoeconomics 20[6], 393-404. 2008.

171. Colquitt J, Clegg A, Sidhu M, Royle P. Surgery for morbid obesity. *The Cochrane Database of Systematic Reviews* 2003;Issue 2. Art. No.: CD003641. DOI: 10.1002/14651858.CD003641.

172. Colquitt J, Clegg A, Loveman E, Royle P, Sidhu MK. Surgery for morbid obesity. *The Cochrane Database of Systematic Reviews* 2005;Issue 4. Art. No.: CD003641. DOI: 10.1002/14651858.CD003641.pub2.

173. Bengtsson BA. Untreated growth hormone deficiency explains premature mortality in patients with hypopituitarism. *Growth Hormone & Igf Research* 1998;8 Suppl A:77-80.

174. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *International Journal of Technology Assessment in Health Care* 2005;21:240-5.

175. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R *et al*. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment* 2004;8:1-158.

Appendix 1 Protocol methods

A review of the evidence for the clinical and cost effectiveness of somatropin will be undertaken systematically following standard guidelines from the NHS Centre for Reviews and Dissemination (CRD).[83] An expert advisory group of clinical experts and service users where appropriate will support the review team at key stages of the project.

Search strategy

- A search strategy will be developed and tested by an experienced information scientist. The strategy will be designed to identify studies reporting clinical-effectiveness, cost-effectiveness, health-related QoL, resource use and costs, epidemiology and natural history.
- The draft clinical effectiveness search strategy for Medline is shown in Appendix 2. This will be adapted for other databases.
- A number of electronic databases will be searched including: The Cochrane Database of Systematic Reviews (CDSR); The Cochrane Central Register of Controlled Trials; NHS CRD (University of York) Database of Abstracts of Reviews of Effectiveness (DARE) and the NHS Economic Evaluation Database (NHS EED); Medline (Ovid); Embase (Ovid); National Research Register; Current Controlled Trials; ISI Proceedings; Web of Science; and BIOSIS. Bibliographies of related papers will be assessed for relevant studies where possible.
- The manufacturers' submissions to NICE will be assessed for any additional studies which meet the inclusion criteria.
- Experts will be contacted to identify additional published and unpublished references.
- Searches will be carried out from the inception date of the database. Although this will involve duplication of searches carried out for the previous review, it will be necessary to identify trials reporting body composition as an outcome measure, as these may not have been identified for all conditions in the previous review. For databases of abstracts and conference presentations searches will only be carried out for the past two years to capture any research that has not yet been fully published. All searches will be limited to the English language, and will be updated around February 2009.

Inclusion and exclusion criteria

Patients

Children with growth disturbance, as per licensed indication for each preparation available.

Interventions

Recombinant human growth hormone (somatropin)

Comparators

Management strategies without somatropin

Outcomes

The following outcomes will be included, where data are available:

- Final height gained
- Height standard deviation score
- Growth velocity
- Growth velocity standard deviation score
- Body composition, and biochemical/metabolic markers as appropriate
- Adverse effects of treatment
- Health-related QoL

Types of studies

- Fully published randomised controlled trials (RCTs) or systematic reviews of RCTs will be included. Indicators of a systematic review include: explicit search strategy, inclusion criteria, data extraction and assessment of quality. Where we judge it necessary and appropriate, we will consider the inclusion of evidence from other non-randomised studies. Full economic evaluations (cost-effectiveness studies, cost-utility studies, cost-benefit studies) and reviews of economic evaluations will be included in the review of cost effectiveness.
- Studies published only as abstracts or conference presentations will only be included in the primary analysis of clinical and cost-effectiveness if sufficient details are presented to allow an appraisal of the methodology and assessment of results.
- Non-English language studies will be excluded.

Inclusion and data extraction process

- Two reviewers will assess the titles and abstracts of studies identified by the search strategy for potential eligibility.
- The full text of relevant papers will be requested for further assessment, and these will be screened independently by two reviewers.
- Data will be extracted by one reviewer using a standard data extraction form and checked by a second reviewer.
- At each stage, any discrepancy will be resolved by discussion, with involvement of a third reviewer where necessary.

Quality assessment

- The quality of included clinical effectiveness studies will be assessed using NHS CRD (University of York) criteria.[83] The methodological quality of the economic evaluations will be assessed using accepted frameworks such as the international consensus-developed list of criteria developed by Evers and colleagues,[174] and Drummond and colleagues.[141] For any studies based on decision models we will also make use of the checklist for assessing good practice in decision analytic modelling (Philips and colleagues).[175,175]
- Quality criteria will be applied by one reviewer and checked by a second reviewer, with differences in opinion resolved by discussion and involvement of a third reviewer where necessary.

Methods of analysis/synthesis

- Clinical- and cost-effectiveness studies will be synthesised through a narrative review with tabulation of results of included studies.
- Where data are of sufficient quality and homogeneity, a meta-analysis of the clinical-effectiveness studies will be performed, using appropriate software.
- Quality of life studies will be synthesised using the same methods as above, i.e. narrative review and meta-analysis as appropriate.

Appendix 2 Literature search strategies

Search strategies for Medline are shown below. Strategies for other databases are available from the authors.

HGH Clinical Effectiveness
Medline all years 1950-2008
Search date: 23/06/09

1 growth disorders/
2 growth failure.ti,ab.
3 growth deficien*.ti,ab.
4 Prader-Willi Syndrome/
5 prader-willi.ti,ab.
6 turner syndrome/
7 (Turner*2 adj syndrome).ti,ab.
8 growth hormone deficien*.ti,ab.
9 GH deficien*.ti,ab.
10 GHD.ti,ab.
11 exp renal insufficiency chronic/
12 (chronic adj2 (renal or kidney*) adj2 (failure or insufficien*)).ti,ab.
13 (CRI or CRF).ti,ab.
14 "small for gestational age".ti,ab.
15 "short for gestational age".ti,ab.
16 infant small for gestational age/
17 "short stature homeobox-containing gene".ti,ab.
18 "short stature homeobox".ti,ab.
19 SGA.ti,ab.
20 SHOX.ti,ab.
21 PHOG.ti,ab.
22 "Pseudoautosomal homeobox-containing osteogenic gene".ti,ab.
23 or/1-22
24 human growth hormone/ or growth hormone/
25 (somatropin* or somatotropin* or somatotrophin* or genotropin* or saizen* or zomacton* or nutropin* or norditropin* or omnitrope* or humatrope*).ti,ab.
26 24 or 25
27 exp child/ or exp adolescent/ or exp infant/
28 child preschool/
29 (child* or infant* or adolescen* or girl* or boy* or prepubert* or pre-pubert*).ti,ab.
30 or/27-29
31 23 and 26 and 30
32 randomized controlled trial.pt.
33 controlled clinical trial.pt.
34 exp Randomized Controlled Trial/
35 exp Randomized Controlled Trials as Topic/
36 exp random allocation/
37 Double-Blind Method/
38 Single-Blind Method/
39 ((singl* or doubl* or trebl*) adj9 (blind* or mask*)).ti,ab.

40 placebo*.ti,ab,sh.
41 random*.ti,ab.
42 (medline or medlars or embase or scisearch or cinahl).ti,ab,sh.
43 (systematic* adj5 review*).mp.
44 (systematic adj5 overview*).mp.
45 (methodolog* adj5 review).mp.
46 (methodolog* adj5 overview).mp.
47 (methodolog* adj5 research*).mp.
48 meta analysis.pt.
49 meta-analysis.sh.
50 (meta-analys* or meta analys* or metaanalys*).mp.
51 ((hand adj5 search*) or (manual* adj5 search)).mp.
52 (electronic* database* or bibliographic* database* or computer* database* or online database*).mp.
53 (Health Technology Assessment* or Medical Technology Assessment*).ti,ab,in.
54 or/32-53
55 31 and 54
56 limit 55 to (english language and humans)
57 kidney transplantation/
58 (renal or kidney*).ti,ab.
59 57 or 58
60 26 and 30 and 54 and 59
61 60 not 56
62 growth hormone/ or human growth hormone/
63 30 and 54 and 59 and 62
64 63 not 56
65 61 or 63
66 limit 65 to (english language and humans)
67 55 or 66
68 (editorial or letter or comment).pt.
69 67 not 68
70 from 69 keep 1-13,21-22

Cost Effectiveness

Medline all years1950 to current: search date 24/06/09

1 exp economics/
2 exp economics hospital/
3 exp economics pharmaceutical/
4 exp economics nursing/
5 exp economics medical/
6 exp "Costs and Cost Analysis"/
7 Cost Benefit Analysis/
8 value of life/
9 exp models economic/
10 exp fees/ and charges/
11 exp budgets/
12 (value adj2 (money or monetary)).tw.
13 (economic adj2 burden).tw.
14 (expenditure* not energy).tw.
15 budget*.tw.

16 (economic* or price* or pricing or financ* or "fee" or "fees" or pharmacoeconomic* or pharma economic* or pharmaco-economic*).tw.
17 (decision adj1 (tree* or analys* or model*)).tw.
18 Resource Allocation/
19 (unit cost or unit-cost or unit-costs or unit costs or drug cost or drug costs or hospital costs or health-care costs or health care cost or medical cost or medical costs).tw.
20 ((value or values or valuation) adj2 (money or monetary or life or lives or costs or cost)).tw.
21 (cost adj2 (util* or effective* or efficac* or benefit* or consequence* or analys* or minimi* or saving* or breakdown* or lowering or estimate* or variable* or allocation* or control* or illness* or affordable* or instrument* or technolog* or fee* or charge* or charges)).tw.
22 Markov Chains/
23 Monte Carlo Method/
24 exp Decision Support Techniques/
25 (resource adj2 (use* or utili* or allocat*)).tw.
26 or/1-25
27 growth disorders/
28 growth failure.ti,ab.
29 growth deficien*.ti,ab.
30 Prader-Willi Syndrome/
31 prader-willi.ti,ab.
32 turner syndrome/
33 (Turner*2 adj syndrome).ti,ab.
34 growth hormone deficien*.ti,ab.
35 GH deficien*.ti,ab.
36 GHD.ti,ab.
37 exp renal insufficiency chronic/
38 (chronic adj2 (renal or kidney*) adj2 (failure or insufficien*)).ti,ab.
39 (CRI or CRF).ti,ab.
40 "small for gestational age".ti,ab.
41 "short for gestational age".ti,ab.
42 infant small for gestational age/
43 "short stature homeobox-containing gene".ti,ab.
44 "short stature homeobox".ti,ab.
45 SGA.ti,ab.
46 (SHOX or PHOG).ti,ab.
47 "idiopathic short stature".ti,ab.
48 "Pseudoautosomal homeobox-containing osteogenic gene".ti,ab.
49 or/27-48
50 human growth hormone/
51 (somatropin* or somatotropin* or somatotrophin* or genotropin* or saizen* or zomacton* or nutropin* or norditropin* or omnitrope* or humatrope*).ti,ab.
52 or/50-51
53 26 and 49 and 52
54 growth disorders/ec or growth hormone/ec
55 53 or 54
56 limit 55 to (human and english language)
57 (editorial or letter).pt.
58 56 not 57
59 "growth hormone".ti,ab.

60 26 and 49 and 59
61 58 or 60
62 limit 61 to (english language and humans)

Quality of life searches
Searched 30/09/08
1 "Quality of Life"/
2 (hql or hqol or "h qol" or hrqol or "hr qol").ti,ab.
3 ("hye" or "hyes").ti,ab.
4 (euroqol or "euro qol" or "eq5d" or "eq 5d").ti,ab.
5 Quality-Adjusted Life Year/
6 "quality adjusted life".ti,ab.
7 (qaly$ or qald$ or qale$ or qtime$).ti,ab.
8 "disability adjusted life".ti,ab.
9 "quality of wellbeing".ti,ab.
10 "quality of well being".ti,ab.
11 daly$.ti,ab.
12 (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab.
13 health$ year$ equivalent$.tw.
14 disutil*.ti,ab.
15 "Value of Life"/
16 rosser.ti,ab.
17 willingness to pay.tw.
18 standard gamble$.tw.
19 time trade off.tw.
20 time tradeoff.tw.
21 health utilit*.ab.
22 exp Health Status/
23 exp Health Status Indicators/
24 "Activities of Daily Living"/
25 "Patient Acceptance of Health Care"/
26 "health related quality of living".ti,ab.
27 "health related quality of life".ti,ab.
28 (patient* adj2 (preference* or satisfaction or acceptance)).ti,ab.
29 (health adj ("state" or "status" or "states")).ti,ab.
30 or/1-29
31 growth disorders/
32 growth failure.ti,ab.
33 growth deficien*.ti,ab.
34 Prader-Willi Syndrome/
35 prader-willi.ti,ab.
36 turner syndrome/
37 (Turner*2 adj syndrome).ti,ab.
38 growth hormone deficien*.ti,ab.
39 GH deficien*.ti,ab.
40 GHD.ti,ab.
41 exp renal insufficiency chronic/
42 (chronic adj2 (renal or kidney*) adj2 (failure or insufficien*)).ti,ab.
43 (CRI or CRF).ti,ab.

44 "small for gestational age".ti,ab.
45 "short for gestational age".ti,ab.
46 infant small for gestational age/
47 "short stature homeobox-containing gene".ti,ab.
48 "short stature homeobox".ti,ab.
49 SGA.ti,ab.
50 SHOX.ti,ab.
51 PHOG.ti,ab.
52 "Pseudoautosomal homeobox-containing osteogenic gene".ti,ab.
53 or/31-52
54 human growth hormone/
55 (somatropin* or somatotropin* or somatotrophin* or genotropin* or saizen* or zomacton* or nutropin* or norditropin* or omnitrope* or humatrope*).ti,ab.
56 54 or 55
57 30 and 53 and 56
58 limit 57 to (english language and humans)
59 (edtorial or letter or comment).pt.
60 58 not 59
61 HIV.ti,ab.
62 60 not 61

Appendix 3 Quality assessment

Criteria	Judgement
1. Was the assignment to the treatment groups really random?	Adequate/partial/ inadequate/unknown
2. Was the treatment allocation concealed?	Adequate/inadequate/ unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported/unknown
4. Were the eligibility criteria specified?	Adequate/partial/ inadequate/unknown
5. Were outcome assessors blinded to the treatment allocation?	Adequate/inadequate/ unknown
6. Was the care provider blinded?	Adequate/partial/ inadequate/unknown
7. Was the patient blinded?	Adequate/partial/ inadequate/unknown
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate/partial/ inadequate/unknown
9. Did the analyses include an ITT analysis?	Adequate/inadequate

Appendix 4 Data extraction tables

GHD Data extraction forms

Reviewers: LB, AT		Date: 22/12	Version: checked
Reference and Design	Intervention	Participants	Outcome measures
Soliman et al., 1996 [84] Country: Egypt Study design: RCT Number of centres: Not stated Funding: Not reported	(Group 1 not data extracted as dose-response arm) 1a GH 30 U/m2/week as a daily s.c. dose 1b GH 15 U/m2/week as a daily s.c. dose 2 a GH 15 U/m2/week as a daily s.c. dose 2 b No treatment (Group 3 not data extracted as not GHD) 3 a GH 15 U/m2/week as a daily s.c. dose 3 b No treatment Duration of treatment:1 year	Target population: Pre-pubertal children with growth hormone deficiency Number of Participants: Total: 77 (19 in Group 2) 1. Group I : 34 children with peak GH response to provocation <7µg (not data extracted as dose response arm) 2. Group II: 19 children with peak GH response to provocation between 7 and 10 µg/l 2a: 9 2b: 10 3. Group III: 24 children with normal peak GH response (not data extracted as not GHD) Sample attrition/dropout: None reported for group 2 Inclusion/exclusion criteria for study entry: Inclusion/ exclusion criteria not clearly stated. Subjects were prepubertal, and bone age was <10 years at initiation of therapy, and <3rd percentile height for chronological age. None of the children had hemoglobinopathy, hepatic or renal impairment. No child had a reduced weight relative to height, other systemic disease, history of head trauma or cranial irradiation, malnutrition, psychosocial dwarfism or hypothyroidism.	Primary outcomes: Not stated Secondary outcomes: GV, HtSDS, bone age delay, IGF-I, glucose, FT4, TSH, GH Method of assessing outcomes: Height measured on a stadiometer, normal population data were according to Tanner, skeletal age examined yearly according to Greulich and Pyle, height determined at 3 month intervals, height GV calculated from height at beginning and end of therapy

Characteristics of participants: Growth parameters and hormonal data			
	GH 15 U/m2/week (n=9)	No treatment (n=10)	Overall
Age, years	7.1 ± 1.9	6.6 ± 1.6	6.8 ± 2.1
Growth velocity (cm/yrs)	3.65 ± 1.1	4.3 ± 1	3.9 ± 1.1
HtSDS (-)	3.4 ± 0.8	3.1 ± 0.6	2.8 ± 1
Bone age delay	2.1 ± 0.8	1.8 ± 0.65	1.9 ± 1
GH peak after clonidine (µg/l)			8.4 ± 1.3
GH peak after insulin (µg/l)			8.1 ± 1.6
IGF-I (ng/ml)	58.5 ± 42.5	52.4 ± 21.3	59 ± 33
Glucose (mmol/l) 0-min	3.6 ± 0.6	4.1 ± 0.5	
Glucose (mmol/l) 120-min	5.4 ± 0.5	4.9 ± 0.45	
FT4 (pmol/l)	16.5 ± 2.1	14.6 ± 1.4	
TSH (uIU/ ml)	1.4 ± 0.4	1.6 ± 0.3	
Results			
Outcomes	GH 15 U/m2/week (n=9)	No treatment (n=10)	P Value
Growth velocity (cm/yrs)	8.4 ± 1.4*†	5.7 ± 1.8	
HtSDS (-)	2.3 ± 0.45*†	2.8 ± 0.45	
Bone age delay	2.25 ± 0.8	1.93 ± 0.75	
GH peak after clonidine (µg/l)	8.6 ± 1.1	8.2 ± 1	
GH peak after insulin (µg/l)	8.5 ± 1.4	8.3 ± 1.2	
IGF-I (ng/ml)	91.2 ± 30.4*†	49.4 ± 19	
Glucose (mmol/l) 0-min	4.3 ± 0.6	4.5 ± 0.8	
Glucose (mmol/l) 120-min	5.1 ± 0.4	4.4 ± 0.6	
FT4 (pmol/l)	17.4 ± 2.2	15.6 ± 1.4	
TSH (uIU/ml)	2.4 ± 0.5	2.2 ± 0.5	
Comments*p <0.05 before v. after 1 year † p<0.05 a vs. b subgroup			
Methodological comments Allocation to treatment groups: Three groups of children were identified and recruited according to their peak GH response to provocation, then subsequently allocated 'at random' to 2 subgroups within that group. No further details on randomisation were provided. Blinding: Blinding is not reported Comparability of treatment groups: Treatment groups appear comparable, but no p value is reported. Method of data analysis: Data presented as mean ± SD Sample size/power calculation: None reported Attrition/drop-out: None reported for group 2, although n=4 excluded from group 1b due to lack of compliance			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Inadequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Adequate
10. Were withdrawals and dropouts completely described?	Adequate

TS Data extraction forms

Reviewers: LB, AT		Date: 08/09/08	Version: checked
Reference and Design	Intervention	Participants	Outcome measures
Quigley et al, 2002[11] U.S RCT, dose response Number of centres: 50 Funding: Author/group appears to be employed by Eli Lilly	1. Growth hormone (Humatrope) (GH) 0.27 mg.kg·wk, with oral placebo (GH 0.27/Pla) 2. GH 0.27 mg/kg·wk with low dose estrogen (GH 0.27/LDE)(Not data extracted) 3. GH 0.36 mg/kg·wk with oral placebo (GH 0.36/Pla) 4. GH 0.36 mg/kg·wk with low dose estrogen (GH 0.36/LDE)(Not data extracted) 5. Placebo injection with oral placebo (Pla/Pla) GH/ Placebo injections: sc in	Target population: Pre-pubertal girls with Turner syndrome (first 18 months of the study data extracted, as placebo group joined group 3 after this time) Number of Participants: Total: 232, stratified by age and randomised. 224 completed 180 days active therapy and have baseline data reported 1.45 2. 47 3. 49 4. 42 5. 41 Sample attrition/dropout: No further details on withdrawals are given (n=8) Inclusion/exclusion criteria for study entry: Inclusion criteria: Karyotypically proven TS ≥ 5 yrs old Bone age ≤ 12 years	Primary outcomes: Near final height (cm)(no placebo group) Changes in HtSDS from baseline to end point (no placebo group) Secondary outcomes: Changes in: bone age, height (cm) Impact of GH dose Effect of low dose oestrogen Method of assessing outcomes: Subjects were assessed every 3 mos for first 6 yr, then 6 mos until study completion: Height using stadiometer, weight and pubertal status. Blood chemistry and thyroid function tests at every visit. Glucose and insulin every 6 mos. IGF-I every 3 mos for first 18 mos, at 24 mos, then annually. X-ray of the left wrist and hand

	equally divided doses, initially three times per wk. Oral placebo given daily Duration of treatment: Placebo group for first 18 months of the study; subjects completed the full study when HV was less than 2 cm/yr and BA ≥ 15 yr Other interventions used: Ethinyl E2 daily 25 ng/kg·d – 200 ng/kg·d depending on age	Prepubertal < 10th percentile for height on NCHS standard HV < 6 cm/year Exclusion criteria: Presence of any Y chromosomal component in karyotype Concurrent treatment with agent that might influence growth Clinically significant systemic illness	for BA performed every 6 mos for 24 mos, then annually. Length of follow-up: 18 months for placebo controlled study

Characteristics of participants:

Baseline (mean ± SD)	GH 0.27/Pla (n=45)	GH 0.36/Pla (n=49)	Pla/Pla (n=41)
Age, years	9.7 ± 2.7	9.8 ± 2.9	9.4 ± 2.7
Bone age (yr)	7.9 ± 2.3	7.9 ± 2.3	7.9 ± 2.4
Height (cm)	119.2 ± 13.6	118.6 ± 12.5	117.6 ± 13.6
Height SDS (NCHS)	-2.7 ± 0.9	-2.9 ± 0.9	-2.9 ± 0.9
Height SDS (NCHS)	0.3 ± 1.0	0.2 ± 0.8	0.2 ± 0.9
Midparental height (cm)	164.6 ± 6.1	162.9 ± 5.9	162.4 ± 5.0
Midparental height SD score	0.27 ± 0.93	0.00 ± 0.91	-0.08 ± 0.77
Pre-study GV	4.1 ± 1.2	4.0 ± 1.2	4.1 ± 1.2

Results

Outcomes	GH 0.27/Pla (n=45)	GH 0.36/Pla (n=49)	Pla/Pla (n=41)	P Value
Height velocity 0-18 months (cm/yr)	6.6 ± 1.1	6.8 ± 1.1	4.2 ± 1.1	<0.001[a]

Comments: [a]Compared with placebo. The 6 monthly GV results are presented on a difficult to read graph – could not data extract. Authors state that HV declined slightly in all GH groups after the initial peak, but was significantly greater than that in the placebo group.

Adverse Effects	Growth Hormone	Placebo	P value
Otitis Media (occurrence/ worsening)	54/186 (29%)	6/46 (13%)	0.037

Comments: Ear pain and ear disorder were not different in frequency between groups. Otitis media was reported in 41% of subjects overall, ear pain in 27% and hypothyroidism in 16%, edema in 3%. There were no disorders that occurred significantly more frequently in subjects receiving the higher

dose. Serious AE (defined as death, life-threatening cancer, hospitalisation, permanently disabling, drug overdose or resulting in congenital anomaly in an offspring) were reported for 47 of 232 subjects. 31/47 of these were hospitalised for surgical procedures, either for elective management of conditions associated with TS or related to accidental injury. 11 were hospitalised for other reasons: infectious illness/dehydration n=5, psychosis n=1, abnormal liver function tests n=1, vaginal bleeding n=1, hematuria n=1, cardiac failure n=1, hypertension n=1. The remaining 5 were reported to have accidentally overdosed on the study drug. Adverse events that were considered unexpected and possibly related to the study drug were reported for 5/232 subjects (2%): hypertension n=2 (in 1 subject this had been present for 11 yrs), surgical procedures n=2, scoliosis n=1. There were no reports of deaths, cancer or neoplasia.
Methodological comments Allocation to treatment groups: Authors state that subjects were randomised in a double blind fashion, but no further details are given. Blinding: States double blind. Placebo is given by injection. BA X-rays were read by a single observer who was blinded to treatment status. Comparability of treatment groups: Treatment groups appear similar at baseline Method of data analysis: Data obtained during the initial 18 month placebo controlled phase are reported for each of the five original randomisation groups. ITT performed for all subjects who received 180 d of active treatment Sample size/power calculation: Not reported Attrition/drop-out: Withdrawals not discussed. 8 patients were randomised but did not complete treatment.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Partial
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Inadequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Inadequate

Reviewers: LB, AT		Date: 10/09/08	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Stephure and The Canadian Growth Hormone Advisory Committee[86] Rovet et al, 1993 [87]	Intervention: (GH group) recombinant human GH (Humatrope, Eli Lilly Canada) by daily sc injection six times weekly (0.30mg/kg·wk, maximum weekly dose 15mg)	Target population: Pre-pubertal girls, aged 7-13 years, with a diagnosis of Turner syndrome documented by peripheral blood karyotype Number of Participants: 154 (95 in Rovet) prepubertal girls Intervention: 76 (51 in Rovet)	Primary outcomes: Bone age (yr) Height (cm) Height SDS (age specific/ adult turner) Change in height(cm) Change in HtSDS (age specific turner) Secondary outcomes:

(No extractable data, so no further information extracted here) Year: 2005 Country: Canada Study design: RCT Number of centres: Multicentre Funding: Eli Lilly Canada, Inc	Control: no GH treatment Other interventions used: Girls with primary ovarian failure received standardised sex steroid replacement: ethinyl estradiol 2.5 µg/d at age 13, 5.0µg/d at age 14, and 2.0µg on d 1-24 with medroxyprogesterone acetate 10 mg on d 15-24 of each month at age 15 and thereafter.	Control: 78 (44 in Rovet) Sample attrition/dropout: Overall, 15 withdrew from GH; 35 from control: -addendum follow up 8 from GH; 9 from control -1997 follow up only, 5 from GH; 13 from control - core protocol data only, 2 from GH; 13 from control Sample crossovers: N/A Inclusion criteria for study entry: height less than the 10th percentile for chronological age on the growth charts of the National Center for Health Statistics of the USA An annualised GV less than 6.0cm/yr during a 6 month prerandomisation period Diagnosis of Turner syndrome documented by peripheral blood karyotype. Phenotypic females with identifiable Y chromosome eligible to participate if had undergone prior gonadectomy Exclusion criteria for study entry: Clinically significant chronic systemic illness, prior treatment with GH, anabolic steroids, estrogens, craniospinal radiation or inadequate thyroxine replacement for hypothyroidism were excluded A spontaneous or stimulated serum GH level was 8.0 µg/litre or greater in all subjects	Method of assessing outcomes: Routine haematology, biochemistry and thyroid function studies were monitored every 3 months (every 6 in control after first year), bone age interpreted by central reader using Greulich and Pyle annually. Length of follow-up: Subjects returned for follow up every 3 months until study completion, protocol completion criteria required annualized GV less than 2cm/yr and bone age 14yr or greater. Addendum follow up = height and safety follow-up at least one year following latest core protocol visit
Characteristics of participants:			

Baseline characteristics Mean ± SD	Growth Hormone n=61	No treatment n=43	P Value
Age	10.3 ± 1.8	10.9 ± 1.7	
Baseline bone age (yr)	8.8 ± 1.4	8.9 ± 1.3	
Baseline height (cm)	119.1 ± 8.5	122.0 ± 7.8	
Baseline HtSDS (age specific Turner)	-0.2 ± 0.9	-0.1 ± 0.8	
Adjusted midparental height (cm)[a]	160.7 ± 6.2	159.3 ± 5.8	
45, X karyotype (%)	62.3	58.1	

Comments: Baseline results for patients who completed the protocol. Baseline data for patients who also had follow up are very similar. No baseline characteristics differed at $p<0.05$ [a] adjusted mid-parental height = [(father height – 13cm) + mother height]/2

Results: Protocol completion characteristics (mean ± SD)

Primary Outcomes Mean ± SD	Growth Hormone n= 61	No treatment n= 43	GH effect [b] mean (95% CI)	P Value
Age (yr)	16.0 ± 0.8	16.5 ± 0.9	-[c]	0.002
Time since randomisation (yrs)	5.7 ± 1.6	5.7 ± 1.6		
Bone age (yr)	14.4 ± 0.8	14.5 ± 0.9	-0.1 (0.5,0.3)	NS
Height (cm)	147.5±6.1	141.0 ± 5.4	7.2 (6.0, 8.4)	p<0.001
Height SDS (age-specific turner)	1.4 ± 1.0	0.2 ± 0.9	1.2 (1.0, 1.5)	p<0.001
Height SDS (adult Turner)	0.7 ± 0.9	-0.3 ± 0.8	1.1 (0.8, 1.3)	p<0.001
Change in height (cm)	28.3 ± 8.9	19.0 ± 6.1	7.2 (6.0, 8.3)	p<0.001
Change in HtSDS (age-specific Turner)	1.6 ± 0.6	0.3 ± 0.4	1.3 (1.1, 1.5)	p<0.001

Comments: [b]ANCOVA model with treatment, baseline HtSDS, baseline HtSDS by treatment interaction, baseline age, and baseline age by treatment interaction. Explanatory variables were removed from the model when not significant. GH effect is estimated by differences of least-squares means for treatment [c] Age at protocol completion was significantly different between control and GH, this reflects the similar numerical difference at baseline and completion, and the lower SD at completion due to the narrower age range
Protocol completion criteria required annualized GV less than 2cm/yr and bone age 14yr or greater.

Results: Addendum follow-up characteristics (mean ± SD)

Primary Outcomes	Growth hormone n= 40	No treatment n= 19	GH effectb mean (95% CI)	P value
Age (year)	20.7 ± 2.5	21.2 ± 2.0		
Time since randomisation (yrs)	10.6 ± 1.7	10.7 ± 1.4		
Bone age (yr)	15.1 ± 1.0	15.2 ± 1.0	0.0 (-0.6, 0.6)	NS
Height (cm)	149.0 ± 6.4	142.2 ± 6.6	73. (5.4, 9.2)	p<0.001
Height SDS (age-	0.9 ± 0.9	-0.1 ± 1.0	1.1 (0.8, 1.4)	p<0.001

specific Turner)				
Height SDS (adult Turner)	0.9 ± 0.9	-0.1 ± 1.0	1.1 (0.8, 1.4)	p<0.001
Change in height (cm)	30.3 ± 8.3	21.6 ± 6.2	7.3 (5.4, 9.1)	p<0.001
Change in HtSDS (age-specific Turner)	1.1 ± 0.5	0.0 ± 0.5	1.1 (0.8, 1.4)	p<0.001
Comments: As for completion characteristics				

Adverse event	Growth Hormone(n=74)	No treatment (n=64)	P value
Surgical procedures	37	17	0.005
Otitis media	35	17	0.014
Ear disorder	15	4	0.024
Joint disorder	10	2	0.036
Respiratory disorder	8	1	0.037
Sinusitis	14	4	0.041
Goiter	0	4	0.004
Death (ruptured aortic aneurysm)	0	1	Not reported
Elevated transamine levels[d]	1	0	Not reported
Intracranial hypertensiond	1	0	Not reported

[d] Leading to withdrawal from study

After protocol completion there was no significant difference in auditory acuity (conductive or neurosensory) between groups (data not shown)

There were no significant between group differences in change from baseline to end point in fasting blood glucose, haemoglobin A1c, serum T4, or TSH (data not shown)

Methodological comments

Allocation to treatment groups: Eligible subjects were stratified for height relative to chronological age at entry and randomly assigned

Blinding: Unblinded – control received no treatment. No mention of blinding of assessors

Comparability of treatment groups: No statistically significant differences between groups at baseline (stated, p values not given)

Method of data analysis: Data are reported as mean ± 1 SD unless stated otherwise. Differences between groups at baseline and end-point for characteristics such as age and duration of therapy were assessed by one-way ANOVA or Fisher's exact test, as appropriate. Age-specific and adult height SD scores (SDS; height SD score) and the change in height SD scores at protocol completion and follow-up relative to baseline were calculated according to published standards for girls with Turner syndrome. No intention-to-treat analysis

Sample size/power calculation: Not calculated

Attrition/drop-out: Drop out is discussed. 15 withdrew from the GH group; 35 from the control. -addendum follow up 8 from GH; 9 from control -1997 follow up only, 5 from GH; 13 from control-core protocol data only, 2 from GH; 13 from control

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate (no treatment)
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: LB, AT		Date: 03/09/08	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
Davenport et al, 2007[85] U.S RCT, open label Number of centres: 11 Setting: US pediatric endocrine centres Funding: Supported by Eli Lilly (EL) and Company, along with grants from universities. Four of the authors are employed by EL, most of the authors have received grant support from EL as	1. Recombinant growth hormone (Humatrope) daily sc injections of 50 µg/kg·d 2. No treatment Duration of treatment:2 years Other interventions used: None	Target population: Girls with Turner syndrome (TS), aged 9 months - 4 years Number of Participants: Total: 89 (The efficacy data exclude one subject who was found after study entry to have a 46, XX karyotype) 1.45 2. 44 Sample attrition/dropout: Overall drop outs: 10 GH group: 4 No treatment: 6 Reasons for discontinuation: Control: Parents' decision n=2 Scheduling problems n=1 Request for GH n=2 Lost to follow up n=1 GH: Relocation n=1 Lost to follow up n=3 Compliance rated as excellent by authors: 95% of subjects received 80% of scheduled injections	Primary outcomes: Change in SDS for length or height (depending on age) from baseline to 2 years A height gain of at least 0.5 was considered clinically significant Secondary outcomes: Serum IGF –I, IFGBP-3 Bone tumour markers Identify factors associated with treatment response determine whether outcome could be predicted by regression model using these factors assess safety of GH treatment in young cohort Method of assessing outcomes: Age – appropriate measures were obtained at each visit for length using infant measuring box (children <2 yr or older children for whom accurate standing measurements could not be obtained) Standard wall-mounted stadiometer (children older

well as consulting and lectureship fees from EL and other pharmaceutical companies in the past		Inclusion/exclusion criteria for study entry: Inclusion criteria: Aged 9 months – 4 years Karyotype proven TS Normal urinalysis, haemoglobin and TSH Adequate thyroid hormone replacement for at least 6 months in those with hypothyroidism Written informed consent from legal guardians Exclusion criteria: Presence of Y chromosomal component in the karyotype in subjects with gonads in situ Autosomal abnormality Concurrent treatment that might influence growth Clinically relevant systemic illness No specific eligibility criteria based on height or GV	than 2 yr) Both length and height measured for girls between 2 and 3 years old. Length measurements in these cases were used for the analyses Length/HtSDS were calculated on the basis of data for aged matched girls from the US Centers for Disease Control Mid-parental height (MPH) calculated as follows: (father's height – 13cm + mother's height)/2 and converted to SDS using normative height data for women at 20 yr of age Serum IGF-I, IGF-binding protein 3 (IGFBP-3) and bone turnover markers were measured at baseline, 4 months, 1 yr and 2 yr. SDS were calculated using Esoterix's data for healthy controls. Bone age x rays obtained at baseline, 1 yr, and 2 yr and read by blinded independent assessors Safety was assessed on each visit based on reported AE, detailed history and physical examinations Length of follow-up: 4 monthly intervals for the 2 years of treatment

Baseline characteristics of participants:

Variable Mean ± SD	Growth hormone (n= 45)	No Treatment (n=43)	P Value
Chronological age, years	1.98 ± 1.01	1.97 ± 1.01	NR
Bone age (yr)[a]	1.95 ± 0.89	1.88 ± 0.96	NR
Bone age-	-0.06 ± 0.56	-0.14 ± 0.42	NR

chronological age			
Length/height (cm)	78.9 ± 8.6	77.6 ± 8.7	NR
Length/HtSDS	-1.42 ± 1.00	-1.76 ± 1.07	NR
MPH (cm)[b]	164.4 ± 5.0	164.4 ± 4.7	NR
MPH SDS[b]	0.17 ± 0.77	0.16 ± 0.73	NR
Weight (kg)	10.35 ± 2.28	9.92 ± 2.47	NR
WtSDS	-1.31 ± 1.18	-1.77 ± 1.46	NR
BMI (kg/m2)	16.48 ± 1.37	16.24 ± 1.29	NR
Head circumference (cm)c	47.2 ± 2.4	46.7 ± 2.1	NR
Head circumference SDS[c]	0.09 ± 1.05	-0.14 ± 1.19	NR
Karyoptype distribution: 45, X	27/45 (60%)	29/43 (67%)	
Karyoptype distribution: 45, X/ 46, XX	7/45 (16%)	7/43 (16%)	
Karyoptype distribution: Other	11/45 (24%)	7/43 (16%)	
IGF-I SDS[d]	-0.25 ± 0.85	-0.39 ± 0.95	NR
IGFBP-3[d]	-0.66 ± 1.08	-0.83 ± 1.05	NR
[a] Baseline bone age missing for 2 subjects in each group [b]Father's height missing for one GH subject at both baseline and endpoint [c]Baseline data missing for one subject in each group; one control subject had an erroneous value at baseline, so the value was not used; endpoint data missing for 2 control subjects [d]baseline data missing for eight control subjects and three GH-treated subjects; endpoint data missing for four control subjects and seven GH subjects			
Results			
Outcomes Mean ± SD	Growth hormone (n= 41)	No Treatment (n=37)	P Value
Chronological age, years	4.03 ± 1.05	4.03 ± 1.03	0.9944
Bone age (yr)[a]	4.24 ± 1.35	3.38 ± 1.11	0.0033
Bone age-chronological age	-0.64 ± 0.80	0.21 ± 0.96	<0.0001
Length/height (cm)	99.5 ± 7.6	91.9 ± 7.2	<0.0001
Length/HtSDS	-0.34 ± 1.10	-2.16 ± 1.22	<0.0001
MPH (cm)[b]	164.7 ± 4.9	164.1 ± 4.9	0.5608
MPH SDS[b]	0.22 ± 0.76	0.12 ± 0.76	0.5607
Weight (kg)	16.62 ± 2.86	13.81 ± 2.50	<0.0001
WtSDS	0.20 ± 1.06	-1.37 ± 1.36	<0.0001
BMI (kg/m2)	16.72 ± 1.70	16.24 ± 1.29	0.1724
Head circumference (cm)[c]	51.1 ± 1.5	49.9 ± 1.4	0.0004
Head circumference SDS[c]	1.17 ± 1.03	0.30 ± 0.99	0.0004
IGF-I SDS[d]	1.26 ± 0.72	-0.69 ± 0.84	<0.0001
IGFBP-3[d]	0.97 ± 0.94	-1.12 ± 1.13	<0.0001

Comments: see notes above
The between group difference for change in HtSDS after 2 yrs was 1.6 ± 0.6 p<0.001 – this analysis was performed on data from the 78 subjects with karyotype proven TS who completed the 2 yr study
The between group difference was significant by 4 months and increased progressively
Total 2 yr height gain was 13.6 ± 3.5 cm for the control group, vs. 20.4 ± 3.3cm for the GH group (p<0.001) Data are reported as mean ± SD unless noted otherwise.

	No Treatment (n=37)	Growth hormone (n= 41)	P Value
First year GV[e] (cm/yr)	8.0 ± 2.4	11.7 ± 2.4	<0.0001
Second year GV (cm/yr)	5.5 ± 1.8	8.4 ± 1.6	<0.0001
First year GV SDS	-0.83 ± 0.95	1.75 ± 1.25	<0.0001
Second year GV SDS	-1.63 ± 1.29	0.70 ± 1.11	<0.001

Comments:[e] Numbers in groups not known for first year results, data are reported as mean ± SD unless noted otherwise.

At the 2 year time point, (when heights of both groups were compared with U.S. standards) only 7% of GH treated subjects remained below -2.0 SDS (~2.3rd percentile); in contrast, 57% of the controls were below -2.0 SDS at 2 yr (p<0.0001).

	Growth hormone (n= 41)	No Treatment (n=37)	P Value
Baseline to 2 yr change: IGF –I SDS	1.53 ± 0.93	-0.09 ± 0.87	Not reported

Adverse Effects	Growth hormone (n= 45)	No Treatment (n=44)	
Serious AE, n (%)[f]	4 (9)	4 (9)	
Treatment emergent AE[g]	42 (93)	43 (98)	

Comments [f]Control group: one subject each was hospitalised for surgical repair of an atrial septal defect, croup/bronchiolitis, gastroenteritis, and dehydration. GH: one subject each was hospitalised for gastroenteritis/dehydration, bacterial pneumonia, persistent bleeding after tonsillectomy and hypoxemia after adenoidectomy [g] events or conditions that began or worsened after study entry: many of these events were related to ear disorders. There was no detrimental effect of GH treatment on frequency of episodes of otitis media, rates of ear tube insertion, middle ear function, or hearing. Most other events reported with a high frequency were typical childhood illnesses considered unlikely to have been related to GH treatment. There were no significant changes or between-group differences in serum TSH.

Adverse events have been reported for the full group numbers.

Methodological comments
Allocation to treatment groups: Children stratified by age (9 months to 2.5 years and >2.5 yr to 4 yr) and then randomised using a blinded phone in process, in a 1:1 ratio
Blinding: Assessors of bone age x-rays were blinded, it is not reported if assessors of other outcomes were, control group did not receive placebo injections
Comparability of treatment groups: The two groups appear broadly similar at baseline. Bone age-chronological age, length/HtSDS, IGF-I SDS and IGFBP-3 SDS were slightly lower in the GH group at baseline. Weight measures were slightly higher in this group. No p value, so unknown if these differences are minimal.

Method of data analysis: The primary efficacy analysis was conducted on the baseline-2 yr change in HtSDS for all subjects who had measurements at both time points (not ITT) using an ANOVA model with treatment group and baseline age group as explanatory variables. For analyses of changes in HtSDS, one-sided tests were used with the significance level set at 0.05. All other analyses of efficacy variables were conducted using two-sided tests with the significance level set at 0.05. Serious AE, treatment-emergent AE and laboratory data were summarised for all subjects who entered the study. Data are reported as mean ± SD unless noted otherwise. Sample size/power calculation: No calculation Attrition/drop-out: Overall drop outs: 10, GH group: 4, no treatment: 6

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Adequate
2. Was the treatment allocation concealed?	Adequate
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: LB, AT		Date: 29/10	Version: Checked
Reference and Design	Intervention	Participants	Outcome measures
Gravholt et al., 2005 [88] Country: Denmark Study design: Randomised, placebo controlled cross over study Number of centres: NR Funding: Govt grant to Novo Nordisk Centre for Research in Growth and	1. GH 0.1 IU/kg/d s.c 2. Placebo Age matched control group studied once (not data extracted) Duration of treatment: 2 months in each arm No washout period between the two study periods Other interventions used: At least 6 months before inclusion in the study all girls had received GH (0.1	Target population: Girls with Turner syndrome Number of Participants: Total: 12 Numbers allocated to each group not given Sample attrition/dropout: Not reported Inclusion/exclusion criteria for study entry: Not stated	Primary outcomes: Not stated Secondary outcomes: Body composition, insulin sensitivity, other biochemical/ metabolic markers, markers of ovarian function (not data extracted) Method of assessing outcomes: Participants studied at the end of every 2 month period, IGF-I, IGFBP-3 and IGFBP-1 and other biochemical markers tested at the end of every study period. Body composition measured by whole body DEXA

Regeneration. One author recipient of honoraria from Pharmacia and Novo Nordisk, and a second is recipient of research grant from Eli Lilly, Novo Nordisk and Roche	IU/kg/d).		Length of follow-up: 4 months

Characteristics of participants: 12 TS girls aged 9.5-14.8 years, (median 12.9) –not reported

Outcomes	GH 0.1 IU/kg/d s.c	Placebo	P Value
FM arms (g/ kg total body weight)	32.9 ± 8.2	36.0 ± 8.6	0.12
FM legs (g/ kg total body weight)	98.7 ± 18.7	104.9± 17.8	0.340
FM trunk (g/ kg total body weight)	80.7 ± 27.4	88.1 ± 35.4	0.1
FM head (g/ kg total body weight)	18.7 ± 3.3	18.7 ± 3.1	0.5
FM total (g/ kg total body weight)	231.0 ± 49.5	247.8 ± 58.1	0.04
BMC arms (g/ kg total body weight)	3.6 ± 0.8	3.5 ± 0.7	0.6
BMC legs (g/ kg total body weight)	10.5 ± 1.7	10.6 ± 1.8	0.3
BMC trunk (g/ kg total body weight)	7.9 ± 1.5	8.0 ± 1.4	0.4
BMC head (g/ kg total body weight)	7.9 ± 1.1	8.0 ± 1.2	0.9
BMC total (g/ kg total body weight)	29.6 ± 3.6	30.1 ± 3.6	0.1
LBM arms (g/ kg total body weight)	62.9 ± 6.4	60.5 ± 6.6	0.1
LBM legs (g/ kg total body weight)	205.7 ± 23.7	202.0 ± 25.9	0.2
LBM trunk (g/ kg total body weight)	378.8 ± 17.4	369.3 ± 29.6	0.046
LBM head (g/ kg total body weight)	78.0 ± 15.2	78.8 ± 13.6	0.5
LBM total (g/ kg total body weight)	725.4 ± 44.8	710.5 ± 54.6	0.05

IGF-I (µg/l)	380.5 ± 116.3	179.8 ± 79.4	<0.0005
IGFBP-1 (µg/l)	3.1 ± 2.4	7.3 ± 4.7	0.002
IGFBP-3 (µg/l)	5982 ± 1557	4344 ± 787	0.002
IGF-I/IGFBP-3 ratio	0.065 ± 0.014	0.041 ± 0.013	<0.0005
Fasting glucose (mmol/l)	4.28 ± 0.59	4.02 ± 0.44	0.046
Fasting insulin (pmol/l)	17.17 ± 8.30	8.58 ± 4.27	0.007a
Fasting glucagon (ng/l)	97.8 ± 43.4	79.2 ± 23.3	0.08
ISIcomp	10.3 ± 9.8	20.9 ± 16.0	0.003
RHOMA	3.34 ± 1.70	1.56 ± 0.87	0.001
AUC insulin (pmol/l/24h)	61 344 ± 28 547	40 868 ± 16 112	0.006
AUC glucose	6922 ± 570	6707 ± 464	0.3
AUC lactate (mmol/l/540 min)	5255 ± 1224	4589 ± 1165	0.2
AUC alanine (µmol/l/540 min)	2230 ± 548	2081 ± 368	0.4
AUC glycerol (µmol/l/540 min)	648 ± 208	527 ± 104	0.1
AUC BOH (µmol/l/540 min)	1215 ± 1486	589 ± 385	0.2
AUC lactateOGTT (mmol/l/120 min)	11569 ± 2438	10239 ± 1674	0.09
AUC alanineOGTT (µmol /l/120 min)	2848 ± 730	2665 ± 459	0.3
AUC glycerolOGTT (µmol /l/120 min)	444 ± 83	408 ± 96	0.2
AUC BOHOGTT (µmol /l/120 min)	564 ± 812	319 ± 268	0.3
AUC FFAOGTT (µmol /l/120 min)	2.43 ± 0.77	2.06 ± 0.91	0.1
Comments: Numbers entered into each group unclear a Wilcoxon 2 tailed test			
Adverse Effects	Not reported		
Methodological comments Allocation to treatment groups: States randomised, but no other details. No details of numbers allocated to groups. Blinding: States placebo used, no other details given Comparability of treatment groups: Appear comparable, but unclear if the details are from baseline Method of data analysis: Groups were compared using Student 2 tailed paired t test, independent t-test, Mann-Whitney U test or Wilcoxon test as appropriate. States that all data were tested for period as well as carryover effects: authors state this did not affect significance. Results expressed as mean ± SD. Statistical significance was assumed for p<5% Sample size/power calculation: Not reported Attrition/drop-out: Not reported/ discussed, no numbers allocated to groups specified			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown

2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Unknown
4. Were the eligibility criteria specified?	Inadequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Unknown
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Inadequate

Reviewers: LB, AT		Date: 27/10/2008	Version: Checked
Reference and Design	Intervention	Participants	Outcome measures
Gravholt et al, 2005[89] Country: Denmark Study design: Randomised, placebo controlled, crossover trial, Number of centres: Not reported Funding: Govt grant to Novo Nordisk Centre for Research in Growth and Regeneration -	All girls were treated with placebo + placebo, GH + placebo or GH + 17β oestradiol (this latter group's results are not data extracted) for a two month period each completed by a 24 h blood sampling period. The treatment regimen was given sequentially and in random order Doses: 1. GH [1.3 ± 0.3 (0.7-1.8)] mg/day [mean ± SD (range)] 2. 17β oestradiol [0.39 ± 0.16 (0.25-0.6) mg/day A pubertal stage matched healthy control group (n=10) was studied once (not data extracted) Duration of treatment: 6 months	Target population: Girls with Turner syndrome Number of Participants: Total:9 No numbers given for treatment groups Sample attrition/dropout: One girl was excluded for non-compliance with study protocol Inclusion/exclusion criteria for study entry: All TS previously verified by chromosomal karyotyping. No other criteria stated.	Primary outcomes: Not stated Secondary outcomes: Insulin sensitivity, glucose tolerance, body composition Method of assessing outcomes: Participants were studied at the end of every 2 month period. IGF-1, IGFBP-3 and IGFBP-1 tested at each study visit. Body composition measured by DEXA Length of follow-up: 8 months (including initial observation period of 2 months)

	Other interventions used: At least 5 months before inclusion in the study all TS girls received GH [1.3 ± 0.3 (0.7-1.8)] mg/day [mean ± SD (range)] and 17β oestradiol [0.39 ± 0.16 (0.25-0.6) mg/day		
Characteristics of participants: Baseline data given for Turner participants as one group, did not extract data for healthy controls			
	Turner syndrome		P Value
Age, years	15.9 ± 1.8		
Weight (kg)	49.1 ± 11.0		
Height (cm)	148.3 ± 4.0		
BMI (kg/m2)	22.2 ± 4.0		
Results			
Outcomes	Growth Hormone	Placebo	P Value
FM arms	41.2 ± 10.2	46.3 ± 12.9	Unclear which groups the p values in the paper are referring to: not data extracted here
FM legs	122.4 ± 22.2	135.1 ± 30.2	
FM trunk	96.2 ± 27.9	116.6 ± 38.7	
FM head	14.7 ± 2.1	14.8 ± 2.5	
FM total	274.5 ± 55.5	312.9 ± 74.7	
BMC arms	4.5 ± 0.4	4.2 ± 0.3	
BMC legs	11.7 ± 0.8	11.9 ± 0.9	
BMC trunk	9.0 ± 1.1	8.9 ± 0.7	
BMC head	7.3 ± 1.2	7.2 ± 1.2	
BMC total	32.5 ± 2.6	32.1 ± 2.0	
LBM arms	61.2 ± 6.5	56.5 ± 10.4	
LBM legs	213.2 ± 24.1	197.2 ± 29.0	
LBM trunk	356.8 ± 20.9	339.9 ± 30.4	
LBM head	61.6 ± 10.7	61.3 ± 10.4	
LBM total	692.8 ± 55.5	655.2 ± 73.7	
IGF-I (µg/l)	661 ± 192	288 ± 69	
IGFBP-1 (µg/l)	1.8 ± 1.2	4.2 ± 2.8	
IGFBP-3 (µg/l)	5157 ± 741	4146 ± 573	
Fasting glucose (mmol/l)	4.46 ± 0.40	4.04 ± 0.47	
Fasting insulin (pmol/l)	147.1 ± 54.0	86.1 ± 41.0	
Fasting glucagon	37.4 ± 12.6	43.0 ± 26.1	

(ng/l)			
ISIcomp	7.0 ± 3.7	14.7 ± 8.7	
RHOMA	4.12 ± 1.60	2.24 ± 1.31	
AUC insulin (pmol/l/24h)	8710 ± 4728	5848 ± 4312	
AUC glucose	119 ± 10	111 ± 13	
AUC lactate (nmol/l/480 min)	4853 ± 1520	5532 ± 2120	
AUC alanine (µmol/l/480 min)	1864 ± 627	2230 ± 543	
AUC glycerol (µmol/l/480 min)	516 ± 245	491 ± 220	
AUC BOH (µmol/l/480 min)	947 ± 1372	338 ± 437	
AUC lactateOGTT (mmol/l/120 min)	3614 ± 976	3718 ± 948	
AUC alanineOGTT (µmol /l/120 min)	855 ± 190	840 ± 159	
AUC glycerolOGTT (µmol /l/120 min)	117 ± 56	99 ± 42	
AUC BOHOGTT (µmol /l/120 min)	96 ± 96	57 ± 68	
AUC FFAOGTT (µmol /l/120 min)	0.83 ± 0.18	0.75 ± 0.27	
Comments FM: fat mass, BMC: bone mineral content, LBM: lean body mass, AUC: area under the curve OGTT: oral glucose tolerance test, BOH: 3-hydroxybutyrate, FFA: free fatty acids			
Adverse Effects	Not reported/ discussed		

Methodological comments
Allocation to treatment groups: Unclear whether allocation to treatment groups has taken place, or whether participants all took the same combination of drugs in the same time period
Blinding: No details given, although is stated that placebo + placebo given and GH+ placebo in those groups
Comparability of treatment groups: Not reported – baseline information given for TS participants as a whole
Method of data analysis: Groups were compared using Student's two tailed paired t-test and an independent t-test when normally distributed, Mann-Whitney and Wilcoxon used for non-parametric data. Results expressed as mean ± SD. Statistical significance was assumed for $p<5\%$
Sample size/power calculation: Not reported
Attrition/drop-out: One patient excluded for non-compliance with study protocol. No further details given.

No washout period. Unclear on whether is randomised or treatment simply given 'in a random order' (p617)

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Not reported
4. Were the eligibility criteria specified?	Inadequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Adequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: LB, AT		Date: 04/11/2008	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
Johnston et al., 2001[90] Country: UK Study design: RCT Number of centres: 6 Funding: Pharmacia Upjohn	1. GH 28-30 IU/m2 surface area/wk daily subcutaneous injection 2.Low dose oestrogen: ethinyloestradiol 1.0 µg/day for <10 y.o and 2.0 µg/day for >10 y.o (approx 50-75 ng/kg body weight daily) 3. Combined ethinyloestradiol and GH (not data extracted) Duration of treatment: 1 year in these groups (group 2 changed to group 3 after the first year, not data extracted, and treatment continued until height increases had fallen below 1cm/year) Other interventions	Target population: Girls with Turner syndrome Number of Participants: Total: 58 1. 22 2. 13 3.23 Sample attrition/dropout: 7 withdrawals, 5 girls reallocated from oestrogen to GH: it is unclear at what point this occurred Inclusion/exclusion criteria for study entry: Inclusion criteria: not stated Exclusion criteria: other growth limiting disorders, prior hormone therapy	Primary outcomes: Height gain at adult height Secondary outcomes: growth enhancing effect of low dose oestrogen (not data extracted), change in HSDS Method of assessing outcomes: Standing height, sitting height, and weight were measured at 3 month intervals; Height standard deviation scores were derived from published Turner height standards, bone age (BA) was initially determined at yearly intervals and calculated using the Tanner-Whitehouse RUS method applicable to normal female population. Various biochemical measures performed at study entry and annually, including triglycerides, cholesterol and TSH Length of follow-up: 1 year

	used: Not stated for year 1		

Characteristics of participants:

	GH 28-30 IU/m2 surface area/wk (n=22)	Low dose oestrogen: ethinyloestradiol (n=13)	P Value*
Age, years	9.0 (5.2 - 15.4)	9.1 (6.0 – 13.7)	
Bone age (y)	8.0 (3.3 -13.5)	7.9 (3.0 – 13.7)	
Height (cm)	113.2 (93.2 – 135.1)	114.0 (94.6 – 140)	
HSDS for CA	-0.3 (-2.1 – 1.2)	-0.1 (-1.5 – 1.8)	
HSDS for BA	0.6 (-0.8 -3.3)	1.0 (-0.6 – 2.4)	
Mid parental HSDS	-0.2 (0.8)	-0.3 (1.1)	

* Not extracted as unclear which groups of the three this refers to. CA: chronological age,
Results are expressed as mean (range) or (SD)
Results

Outcomes	GH 28-30 IU/m2 surface area/wk (n=unclear)	Low dose oestrogen: ethinyloestradiol (n=unclear)	P Value
Change in HSDS in first year	+0.7 (0.7)	+0.4 (0.9)	<0.05

Adverse Effects: Three of 58 girls ceased growth hormone early because of serious health events not directly related to GH or low dose oestrogen: one each with hypertension, ulcerative colitis, and brain tumour. One patient in group 3 died from aortic dissection shortly after treatment cessation. Compliance problems led to the withdrawal of four patients. Seven others developed coincidental disorders but these were not considered sufficient to invalidate continued participation in the study. Five girls from group 2 were allocated to low dose oestrogen were re-allocated to GH due to concerns over early breast development at age range 6.2-8.9 y.o.

Methodological comments
Allocation to treatment groups: States randomised, no other details given. 5 girls reallocated from oestrogen to GH: it is unclear at what point this occurred
Blinding: Unknown, no details given
Comparability of treatment groups: Authors state that the groups were similar for the main monitoring parameters
Method of data analysis: Within group results were compared using the paired Student's t test. Between group results were compared using analysis of variance.
Sample size/power calculation: Not reported
Attrition/drop-out: 7 withdrawals: Three of 58 girls ceased growth hormone early because of serious health events not directly related to GH or low dose oestrogen, Compliance problems led to the withdrawal of four patients. Treatment centres had the option of stopping ethinyloestradiol therapy if girls showed unacceptable premature breast development or excessive bone maturation: this occurred in 5 cases. Group numbers for final height data are lower, for the 1 year data they are unclear

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Inadequate
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Inadequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown

6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Unknown
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Inadequate

PWS Data extraction forms

Reviewers: AT, LB		Date: 20/10/2008	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Lindgren et al., 1998 [101] Lindgren et al., 1997 [100] Country: Sweden and Denmark Study design: RCT Number of centres: multicentre Funding: Pharmacia & Upjohn AB	1. 0.1 IU/kg/day GH by s.c. injection 2. no treatment Duration of treatment: 2 years (only year 1 data extracted as no control arm in year 2) Other interventions used: special dietary instructions more than 1 year before start of treatment and throughout the study period to ensure constant energy intake per kg body weight	Target population: prepubertal children aged 3-12 with PWS Number of Participants: Total: n=29 1. n=15 2. n=14 An additional group of non-PWS obese children was also studied, but data from this group were not data extracted. Sample attrition/dropout: 2 control group patients excluded from analysis Inclusion criteria for study entry: fulfilled diagnostic criteria for PWS and had either a paternal deletion or maternal disomy of chromosome region 15q11-13; projected final height <165cm (boys) and 154 cm (girls).	Primary outcomes: Secondary outcomes: HtSDS; GV SDS, BMI SDS, lean mass, % body fat Method of assessing outcomes: height and WtSDS calculated with reference to the standard for healthy Swedish children; bone age was assessed according to Tanner-Whitehouse 2/RUS; % body fat estimated by dual-energy X-ray absorptiometry QoL questionnaires completed (but no extractable data reported) Length of follow-up: 1 year

Characteristics of participants:

Mean (range)	0.1 IU/kg/day GH (n=15)	No treatment (n=12)	P Value
Age (years)	6.8 (3.6 – 11.9)	6.4 (3.3 – 11.7)	
Bone Age (years)	6.6 (3.3 – 13.0)	5.4 (3.3 – 10.2)	
Sex	7 female, 8 male	5 female, 7 male	
Target HtSDS	0.4 (-1.3 – 1.8)	-0.1 (-1.5 – 1.0)	
HtSDS	–1.6 (-4.0 – 0.5)	-1.7 (-5.3 – 0.4)	
BMI (SDS)	3.0 (-0.7 – 7.6)	2.1 (-1.3 – 5.1)	
Height velocity (SDS) mean ± SD (range)	–1.9 ± 2.0 (-6.4 – 0.9)	–0.1 (-1.7 – 2.71)	

IGF-I (SDS)	-1.6 (-3.0 to -0.6)	-1.4 (-2.4 to -0.1)	
Mean (± SD)			
Fat-free mass (kg) By DEXA By BIA	 14.9 ± 4.1 14.6 ± 3.9	 14.1 ±3.0 13.6 ± 3.3	
Body fat (%) By DEXA By BIA	 40.0 ± 10.5 44.6 ± 9.2	 34.8 ± 7.9 41.3 ± 10.7	
Comments DEXA=dual-energy X-ray absorptiometry; BIA=bioelectrical impedance analyser Height velocity SDS was during 12 months before treatment commenced			
Results			
Mean (range)	0.1 IU/kg/day GH (n=15)	No treatment (n=12)	P Value
Bone Age (years)	8.0 (5.5 – 13.9)*	6.9 (3.9 – 11.4)	
Bone Age (years) change from baseline	1.4 (0.0-2.8)	1.5 (0.4 – 2.6)	
HtSDS	-0.4 (-2.7 -1.9)*	-1.8 (-5.1 -0.2)	
BMI (SDS)	2.0 (-2.4 -6.7)*	2.5 (0.1-6.1)	
Height velocity (SDS) mean ± SD (range)	6.0 ± 3.2 (1.4-11.9)*	-1.4 (-3.2 -0.3)	
IGF-I (SDS)	1.8 (-0.1 -4.1)*	-1.4 (-2.9 to -0.3)	
Mean (± SD)			
Fat-free mass (kg) By DEXA By BIA	 19.8± 5.2** 21.7± 8.9**	 15.2± 2.9 14.8± 3.5	
Body fat (%) By DEXA By BIA	 30.9± 11.4** 30.3± 10.5**	 38.2± 9.1 43.3 ±12.9	
Comments * change from baseline $p<0.05$ ** change from baseline $p<0.001$			
Adverse Effects i.v. glucose-tolerance test was normal and unchanged in all children. Basal fasting insulin levels were significantly increased throughout the group in the GH group (from 10.4 mU/I±2.7 SD to 19.2 mU/I±10.5 SD, $p<0.001$). No severe progression of scoliosis (angle $\geq 20^{o}$) in either group. Bone mineral density did not differ between groups. One child developed low levels of thyroxine without any change in TSH levels. He received substitution with L-thyroxine during the GH treatment. The increased levels of fasting insulin during the treatment may be regarded as laboratory AE. However, both fasting glucose and HbA1C were unchanged and, although increased compared to pre-treatment, insulin levels were still within the normal range.			
Methodological comments Allocation to treatment groups: states children were randomized, but no further details given Blinding: open label Comparability of treatment groups: baseline age, height, BMI and height velocities stated to be similar in both PWS groups. Method of data analysis: Student's 2-tailed paired and unpaired t-tests were used for normally distributed values, and non-parametric tests were used otherwise. Single regression analysis used for			

statistical comparisons. Not ITT. Data were analysed as change from baseline rather than between-group differences. Sample size/power calculation: not reported Attrition/drop-out: 1 patient excluded at baseline evaluation because she had a severe scoliosis that required surgical intervention; one patient was excluded after 6 months in the control arm because she developed central precocious puberty

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 23/10/2008	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Carrel et al., 2004 [21] and Myers et al. 2007 [97] Whitman et al. 2004 [98] Country: USA Study design RCT Number of centres 2 Funding: supported by Pharmacia Inc. (Pfizer)	1. 1mg/m2/day GH 2. no treatment Duration of treatment: 1 year Other interventions used: 0.1g/kg of deuterium-labelled water was given on day 1 and 0.15g/kg of oxygen-18 water.	Target population: infants and toddlers with PWS Number of Participants: Total: n=32 (Whitman et al); n=29 (Carrel); n=25 (Myers) 1. n=15 2. n=14 In Whitman paper – 30 patients completed first 6 months: n=18 n=12 Sample attrition/dropout: none in Difference in n between Whitman paper and others suggests 7 patients dropped out Inclusion criteria for study entry: confirmed diagnosis of PWS; age 4-37 months;	Primary outcomes: not stated Secondary outcomes: % body fat, lean body mass, bone mineral density, GV SDS, change in height, IGF-I; mobility (not data extracted as not per protocol) Method of assessing outcomes: Harpenden stadiometer used for length/height for children >2, otherwise an infantometer was used; body composition measured by dual-energy x-ray absorptiometry; Length of follow-up: 1 year

Characteristics of participants:			
Mean ± SD	1mg/m2/day GH (n= 15)	No treatment (n=14)	P Value
Age, months	13 ± 8	15 ± 0	ns
% female	50	42	ns
Length/HtSDS*	-1.6 ± 1.2	-1.3 ± 1.1	
Growth velocity SDS	1.4 ± 1.8	1.2 ± 1.4	
Body fat, %*	28 ± 7	29 ± 12	
Lean mass, kg*	5.8 ± 1.9	6.9 ± 2.0	
BMD, g/cm2*	0.60 ± 0.08	0.64 ± 0.09	
Total cholesterol mg/dL	163 ± 34	170 ± 30	
IGF-I (ng/dL)*	34 ± 21	Not reported	
Fasting insulin μIu/mL	4.8 ± 3.7		
Comments * from Myers paper, which had unclear patient numbers Baseline data are also given by Whitman et al. These have not been data extracted as they differ slightly from the group presented here. Whitman's results were for 6 months, so it is assumed that the Carrel data supersede these.			
Results			
Mean ± SD	1mg/m2/day GH (n= 15)	No treatment (n=14)	P Value
Mean % body fat	23.2 ± 8.9	32.7 ± 8.8	0.03
Change in body fat	-4.8% ± 5.7%	+4.1%4.6%	P=0.001
Change in lean body mass (kg)	3.6 ± 0.5	1.8±0.7	P<0.001
Change in height (cm)	+15.4 ± 2.3	9.2 ± 3.2	P<0.001
Growth velocity SDS	5.0 ± 1.8	1.2 ± 1.4	
IGF-I ng/mL	231 ± 98	51 ± 28	P<0.001
Fasting insulin μIu/mL	5.6 ± 7.1	5.7± 7.1	ns
Bone mineral density (%)	14.1 ± 10.4	9.0 ± 6.9	ns
Total cholesterol mg/dL	159 ± 40	183 ± 43	
Length/HtSDS*	-0.2 ± 1.5	-1.5± 0.7	
Comments GVSDS in GH patients p<0.001 compared with baseline. * from Myers paper, which had unclear patient numbers Length/HtSDS change from baseline in GH group, p<0.005			
Adverse Effects No changes in the prevalence of scoliosis were seen between the treatment and control groups (Carrel) although Myers et al. comment on progression of scoliosis in one patient. No other adverse effects were noted during this study, and no subject required thyroid hormone therapy. After the first 6 months, 2 children showed a 3.5 SD increase in head circumference. This was monitored, but the later papers do not mention it.			
Methodological comments			

Allocation to treatment groups: randomisation following stratification by age (4-18 months and 19-37 months) and sex. No further details given. Myers and Whitman papers state a 60:40 ratio was used, but this doesn't reflect numbers in Carrel suggesting that attrition bias may have affected the results. Blinding: none Comparability of treatment groups: similar at baseline Method of data analysis: t-test for between group comparisons. Doesn't appear to be ITT. Data reported by Whitman et al was for 25 patients who completed the first 6 months. All three papers appear to report data for a slightly different version of the patient group. Sample size/power calculation: not reported Attrition/drop-out: Difference in n between Whitman paper and others suggests 7 patients dropped out

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	inadequate

Reviewers: AT, LB		Date: 29/10	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Carrel et al., 1999 [95] and Myers et al. 1999 [96] Country USA Study design open RCT Number of centres – not reported Funding: Genentech foundation for growth	1.GH 1 mg/m2/d 2. no treatment Duration of treatment: 1 year Other interventions used: standardised caloric intake	Target population: children with PWS without prior GH therapy Number of Participants: Total: n=54 1. n=35 2. n=19 Sample attrition/dropout: none Inclusion criteria for study entry: Genetically confirmed PWS Pts were aged 4-16, with skeletal maturation <13 for girls and <15 for boys Exclusion criteria: prior GH	Primary outcomes: not clearly stated Secondary outcomes: HtSDS; GV; GVSDS; Body fat; Lean mass; BM; IGF-1; IGFBP-3; insulin; cholesterol; HDL-C; strength and agility (not data extracted as not per protocol). Method of assessing outcomes: height measured by Harpenden stadiometer; Greulich and Pyle method of determining bone age;

and development		therapy	body composition assessed using dual-energy x-ray absorptiometry Length of follow-up: 1 year

Characteristics of participants:			
Mean ± SD	GH 1 mg/m2/d (n=35)	No treatment (n=19)	P Value
Sex (% female)	42	58	
Mean age (y)	9.8	10.0	
Prepubertal (n)	34 (97%)	17 (90%)	
Height SDS	–1.1 ± 1.3	-1.5 ± 0.8	
Mean GV (cm/y)	4.72 ± 2.2	5.18 ± 1.5	
Mean GV SDS	-1.0 ± 2.5	-0.9 ± 1.7	
Bone age	9.1 ± 3.6	8.4 ± 3.1	
Body fat (%)	46.3± 8.4	42.6 ± 8.1	
Lean mass (kg)	20.5 ± 6.3	20.5 ± 5.0	
BMI (kg/m2)	25.0 ±6.7	24.2 ± 6.5	
IGF-1 (ng/mL)	127 ± 67	139 ± 64	
IGFBP-3 (ng/mL)	1.73 ± 0.49	1.84 ± 0.64	
Insulin-0 hour (mIU/L)	11.2 ± 9.9	9.3 ± 6.2	
Insulin-2 hour (mIU/L)	49.5 ± 40.7	41.6 ± 42.5	
Total cholesterol (mg/dL)	184 ± 36	190 ± 36	
HDL-C (mg/dL)	42 ± 8	44 ± 9	
Femoral neck BMD (g/cm3)	0.656 ± 0.19	0.636 ± 0.9	
Spine BMD (g/cm3)	0.744 ± 0.14	0.753 ± 0.12	
Scoliosis (°)	9.1 ± 6.0	14.7 ± 11.0	
Free fatty acids (mmol/I)	0.6 ± 0.4	0.6 ± 0.3	
Triglycerides (mg/dl)	91.6 ± 57.9	84.3 ± 39.6	
Results			
Mean ± SD	GH 1mg/m2/d (n=35)	No treatment (n=19)	P Value
Height SDS	-0.6 ± 1.2	-1.6 ± 1.2	p < 0.01
Mean GV (cm/y)	10.1 ± 2.5	5.0 ± 1.8	p < 0.01
Mean GV SDS	4.6 ± 2.9	-0.7 ± 1.9	p < 0.01
Bone age	10.6± 3.5	9.8 ± 3.0	n/s
Body fat (%)	38.4 ± 10.7	45.8 ± 8.8	p < 0.01
Lean mass (kg)	25.6 ± 4.3	21.7 ± 5.0	p < 0.01
BMI (kg/m2)	23.7 ± 6.3	25.2 ± 8.9	n/s
IGF-1 (ng/mL)	522 ± 127	121 ± 52	p < 0.01
IGFBP-3 (ng/mL)	3.5 ± 0.73	2.07 ± 0.45	p < 0.01
Insulin-0 hour	18.6 ± 14.6	8.8 ± 5.4	

(mIU/L)			
Insulin-2 hour (mIU/L)	70.2 ± 44.2	47.1 ± 34.1	
Total cholesterol (mg/dL)	166 ± 34	193 ± 34	p < 0.01
HDL-C (mg/dL)	50 ± 10	44 ± 8	p < 0.01
Femoral neck BMD (g/cm3)	0.797± 0.09	0.707 ± 0.09	P<0.05
Spine BMD (g/cm3)	0.834 ± 0.15	0.793± 0.13	
Scoliosis (°)	12.1 ± 7.0	16.6 ± 10.0	
Free fatty acids (mmol/I)	0.72 ± 0.40	0.64 ± 0.30	P<0.01
Triglycerides (mg/dl)	86.0 ± 62.0	94.2 ± 49.0	
Comments P values are for paired t-test before and after GH therapy, compared with either baseline values of treated patients or 12-month values of non-treated patients.			
Adverse Effects Headaches in 2 patients treated with GH within first 3 weeks. Symptoms resolved with temporary cessation and gradual re-institution of GH. No pseudotumor cerebri			
Methodological comments Allocation to treatment groups: reported as randomised 60:40. Method not stated Blinding: none Comparability of treatment groups: similar at baseline Method of data analysis: ITT. Data were analysed using a Student's t-test for paired samples or two related samples. Sample size/power calculation: not reported Attrition/drop-out: none			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	unknown
2. Was the treatment allocation concealed?	unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	adequate
9. Did the analyses include an ITT analysis?	Adequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 10-03-09	Version: Initial
Reference and Design	Intervention	Participants	Outcome measures
de Lind van Wijngaarden	1. 1mg/m2 s.c. daily	Target population: infants and prepubertal children with PWS,	Primary outcomes: not stated

2009 et al. [93] Festen et al., 2008 [94] Country: the Netherlands Study design: RCT Number of centres: 18 Funding: not stated	2. no treatment Duration of treatment: 1 year for infants and 2 years for children After 1st year infants were all offered a second year of GH treatment. Not discussed here as no control group. Other interventions used: caloric intake and activity level standardised and monitored	who were not severely overweight, naïve to GH treatment Number of Participants: Total: n=104 enrolled, n=91 were available for follow-up: 42 infants (<3.5 years) and 49 children over 3.5 years. Randomised groups not clear. The following are the groups analysed at year 1: Infants: 1 n=19 2 n=19 Children: 1 n= 23 2 n=21 Sample attrition/dropout: 4 infants and 5 children excluded from analysis Inclusion criteria: genetically confirmed diagnosis of PWS; age 6 mths – 12 yrs (girls) or 14 yrs (boys); bone age <14 (girls) or 16 (boys); prepubertal - Tanner breast stage ≤ 2 for girls and testicular volume <4ml for boys Exclusion criteria: non-cooperative behaviour; on medication to reduce fat	Secondary outcomes: HtSDS; BMI; BMI SDS; head circumference SDS; IGF-I; IGF-I SDS; IGFBP-3; IGFBP-3 SDS; IGF-I/BP3 (SDS); LBM and scoliosis Method of assessing outcomes: Harpenden stadiometer used to measure height, using a mean of 3 values. Anthropometric measurements taken at baseline and every 3 months; dual energy x-ray absorptiometry used for fat measurements Length of follow-up: 1 year (infants), 2 years (children)

Characteristics of participants from Festen et al., 2008 [94] (other than scoliosis and Trunk LBM:BSA) as this is the most complete:

Baseline characteristics of infants (6 months – 3 years)

Median (IQR)	1mg/m2 s.c. daily rhGH (n=20)	No treatment (n=22)	P Value
Sex (m/f)	12/8	16/6	
Age, years	2.0 (1.6 to 3.1)	1.3 (1.0 to 2.8)	
HtSDS	-2.3 (-2.8 to -0.7)	-2.1 (-3.2 to -1.0)	
BMI (kg/m2)	16.4 (15.1 to 18.6)	16.1 (14.7 to 18.2)	
BMI (SDS)	0.5 (-0.9 to 1.9)	-0.8 (-1.7 to 1.6)	
Head circumference (SDS)	-0.8 (-1.6 to -0.3)	-1.1 (-1.8 to -0.5)	
IGF-I (ng/ml)	27.0 (22.0 to 35.0) (n=11)	47.0 (17.0 to 52.0)	
IGF-I (SDS)	-1.9 (-2.8 to -1.3) (n=11)	-1.6 (-2.6 to -0.4) (n=11)	
IGFBP-3 (ng/ml)	0.8 (0.7 to 1.1) (n=11)	1.1 (0.8 to 1.3) (n=11)	
IGFBP-3 (SDS)	-2.6 (-3.3 to -2.0) (n=11)	-1.5 (-2.6 to -0.7) (n=11)	
IGF-I/BP3 (SDS)	-0.9 (-2.0 to -0.4) (n=11)	-0.3 (-1.7 to 0.6) (n=11)	

Scoliosis (%)	7 (37) (n=19)	4 (21) (n=19)	
TrunkLBM:BSA	7.4 (6.9 to 8.0) (n=19)	7.3 (7.0 to 7.7)(n=19)	
Baseline characteristics of children (3-14 years)			
Median (IQR)	1mg/m2 s.c. daily rhGH (n=25)	No treatment (n=22)	P Value
Sex (m/f)	13/12	8/14	
Age, years	6.8 (5.4 to 8.8)	5.9 (4.7 to 7.4)	
HtSDS	-2.0 (-3.1 to -1.7)	-2.5 (-3.3 to -1.9)	
BMI (kg/m2)	17.7 (16.0 to 22.3)	18.1 (17.2 to 19.9)	
BMI (SDS)	1.2 (0.1 to 2.2)	1.3 (1.1 to 1.6)	
Head circumference (SDS)	-0.8 (-1.5 to -0.2)	-0.6 (-1.2 to -0.1)	
IGF-I (ng/ml)	60.0 (46.5 to 96.5) (n=21)	56.0 (42.0 to 88.0) (n=18)	
IGF-I (SDS)	-1.7 (-2.3 to -1.2) (n=21)	-1.9 (-2.6 to -1.2) (n=18)	
IGFBP-3 (ng/ml)	1.3 (0.9 to 1.5) (n=21)	1.2 (0.9 to 1.5) (n=18)	
IGFBP-3 (SDS)	-1.9 (-2.8 to -1.2) (n=21)	-2.2 (-3.1 to -1.4) (n=18)	
IGF-I/BP3 (SDS)	-0.5 (-1.0 to 0.5) (n=21)	-0.6 (-1.6 to 0.3) (n=18)	
Fat % (SDS)	2.1 (1.7 to 2.7) (n=?)	2.3 (1.9 to 2.6) (n=?)	
Fat (SDS)	1.2 (0.8 to 2.0) (n=?)	1.2 (0.7 to 1.6) (n=?)	
LBMage (SDS)	-1.7 (-3.0 to -1.0) (n=?)	-1.9 (-3.4 to -1.2) (n=?)	
LBMHtSDS	-1.7 (-3.8 to -0.6) (n=?)	-1.4 (-2.9 to 0.9) (n=?)	
Trunk fat (%)	36.0 (24.8 to 46.2) (n=?)	36.0 (29.2 to 41.2) (n=?)	
Scoliosis (%)	7 (30) (n=23)	9 (43) (n=21)	
TrunkLBM:BSA	8.0 (7.5 to 8.4)(n=23)	7.6 (7.1 to 8.1) (n=21)	
Comments N is unclear for body composition measures, as these were only available for children over the age of 4 at the start of the study. P vals are for change in GH group vs. control group			
Results infants (6 months – 3 years) – mostly from de Lind van Wijngaarden 2009 et al. [93] as this is the most complete data			
Median (IQR)	1mg/m2 s.c. daily rhGH 1 year (n=19)	No treatment (n=19)	P Value
HtSDS	-0.9 (-1.6 to -0.1)	-1.8 (-3.5 to -1.4)	0.003
ΔHtSDS	1.2 (1.0 to 1.6)	-0.2 (-0.6 to 0.3)	<0.0001
BMI (kg/m2)	16.3 (15.7 to 18.2)	16.4 (15.4 to 19.8) (n=15)	
BMI (SDS)	0.3 (-0.1 to 1.6)	0.3 (-0.6 to 1.6)	0.72
Δtrunk LBM	1.7 (1.3 to 2.1)	0.7 (0.4 to 0.9)	<0.0001
Δtrunk LBM:BSA	1.2 (0.7 to 1.8)	0.3 (-0.3 to 0.6)	0.002
Head circumference (SDS)	0.0 (-0.9 to 0.7) (n=16)	-0.8 (-1.6 to -0.3) (n=15)	P<0.001
IGF-I (ng/ml)	179.0 (119.5 to 241.0) (n=12)	33.0 (22.5 to 47.8) (n=15)	
IGF-I (SDS)	2.5 (1.4 to 2.9)	-2.6 (-4.1 to -0.7)	<0.0001
IGFBP-3 (ng/ml)	2.2 (1.6 to 2.4) (n=12)	0.9 (0.7 to 1.3) (n=12)	
IGFBP-3 (SDS)	0.5 (0.0 to 1.2) (n=12)	-2.4 (-3.5 to -1.2) (n=12)	
IGF-I/BP3 (SDS)	2.3 (1.7 to 3.4) (n=12)	-1.1 (-2.1 to 0.0) (n=12)	P<0.001
Onset scoliosis (%)	4 (21) (n=19)	2 (11) (n=19)	P=0.71
Progression of scoliosis	-6.0 (-12.5 to 12.8) (n=19)	-7.5 (-7.5 to-5.0) (n=19)	P=0.48
Results for children (3-14 years) mostly from de Lind van Wijngaarden 2009 et al. [93] as this is the most complete data			
Median (IQR)	1mg/m2 s.c. daily rhGH	No treatment	P Value
Year 1 results	N=23	N=21	

HtSDS	-1.0 (-1.5 to -0.3)	-2.5 (-3.4 to -2.3)	<0.0001
ΔHtSDS	0.9 (0.7 to 1.3)	-0.1 (-0.2 to 0.1)	<0.0001
BMI (kg/m2)	17.5 (15.3 to 19.8) (n=21)	18.6 (17.6 to 19.7) (n=21)	
BMI (SDS)	0.8 (-0.1 to 2.1)	1.4 (1.0 to 1.6)	0.05
Δtrunk LBM	1.8 (1.4 to 2.3)	0.7 (0.1 to 0.8)	<0.0001
Δtrunk LBM:BSA	1.3 (0.7 to 1.7)	0.0 (-0.4 to 0.3)	<0.0001
Head circumference (SDS)	-0.2 (-1.2 to 0.2) (n=21)	-0.6 (-0.9 to 0.3) (n=21)	
IGF-I (ng/ml)	337.0 (274.3 to 474.3) (n=21)	55.0 (42.5 to 94.8) (n=12)	
IGF-I (SDS)	2.3 (1.5 to 2.8)	-2.5 (-3.1 to -1.5)	<0.0001
IGFBP-3 (ng/ml)	2.5 (2.2 to 2.9) (n=21)	1.3 (0.8 to 1.5) (n=12)	
IGFBP-3 (SDS)	0.4 (-0.1 to 0.8) (n=21)	-2.4 (-3.5 to -1.8) (n=12)	P<0.001
IGF-I/BP3 (SDS)	2.5 (2.0 to 3.0) (n=21)	-0.8 (-1.4 to -0.2) (n=12)	P<0.001
Fat % (SDS)	1.5 (0.7 to 2.1) (n=?)	2.3 (2.0 to 2.6) (n=?)	P<0.001
Fat (SDS)	0.9 (0.2 to 1.4) (n=?)	1.3 (0.7 to 1.9) (n=?)	P<0.001
LBMage (SDS)	-0.5 (-1.3 to 0.7) (n=?)	-2.1 (-4.1 to -1.3) (n=?)	P<0.001
LBMHtSDS	-1.5 (-2.3 to -0.7) (n=?)	-1.9 (-2.9 to 0.0) (n=?)	P<0.05
Trunk fat (%)	28.0 (16.9 to 36.7) (n=?)	37.2 (32.0 to 42.5) (n=?)	P<0.001
Onset scoliosis (%)	5 (22) (n=23)	6 (29) (n=21)	P=0.52
Progression of scoliosis	-3.5 (-7.3 to 1.8) (n=23)	0.0 (-1.0 to 1.0) (n=21)	P=0.60
Year 2 results	N=23	N=21	
HtSDS	-0.5 (-0.8 to 0.0)	-2.6 (-3.4 to -2.3)	<0.0001
ΔHeight SDS	1.4 (1.3 to 1.8)	-0.1 (-0.4 to 0.1)	<0.0001
BMI (kg/m2)	17.5 (16.1 to 21.1) (n=20)	19.1 (17.8 to 20.8) (n=20)	
BMI (SDS)	1.1 (-0.2 to 1.7)	1.4 (1.1 to 1.6)	0.19
Δtrunk LBM	2.8 (2.6 to 3.5)	0.8 (0.4 to 1.0)	<0.0001
Δtrunk LBM:BSA	1.4 (0.5 to 1.7)	-0.2 (-0.5 to -0.1)	<0.0001
Head circumference (SDS)	-0.1 (-1.1 to 0.5) (n=20)	-0.6 (-1.1 to 0.3) (n=20)	P<0.05
IGF-I (ng/ml)	424.0 (313.0 to 570.0) (n=20)	92.0 (61.8 to 130.0) (n=16)	
IGF-I (SDS)	2.4 (2.1 to 2.8)	-1.6 (-2.5 to -1.0)	<0.0001
IGFBP-3 (ng/ml)	2.8 (2.6 to 3.2) (n=20)	1.5 (1.2 to 1.8) (n=16)	
IGFBP-3 (SDS)	0.6 (0.3 to 1.1) (n=20)	-1.7 (-2.3 to -1.2) (n=16)	P<0.001
IGF-I/BP3 (SDS)	2.5 (1.8 to 2.9) (n=20)	-0.6 (-1.2 to -0.1) (n=16)	P<0.001
Fat % (SDS)	1.9 (0.7 to 2.3) (n=?)	2.4 (2.1 to 2.7) (n=?)	P<0.001
Fat (SDS)	1.1 (0.6 to 2.0) (n=?)	4.5 (0.9 to 2.0) (n=?)	P<0.01
LBMage (SDS)	-0.1 (-1.3 to 0.6) (n=?)	-2.5 (-3.8 to -1.4) (n=?)	P<0.001
LBMHtSDS	-1.9 (-2.4 to -1.4) (n=?)	-2.3 (-2.7 to -1.3) (n=?)	P<0.05
Trunk fat (%)	33.3 (17.3 to 40.9) (n=?)	37.9 (35.0 to 45.7) (n=?)	P<0.001
Onset scoliosis (%)	5 (22) (n=23)	7 (33) (n=21)	P=0.14
Progression of scoliosis	3.3 (-4.3 to 11.9) (n=23)	-5.0 (-9.0 to -2.0) (n=21)	P=0.27
Comments N is unclear for body composition measures, as these were only available for children over the age of 4 at the start of the study. P vals are for change in GH group vs. control group Progression of scoliosis is change in Cobb angle during study			
Adverse Effects Not reported – the reader is referred to 3 other papers by the same author, but 2 of these appear to be other, smaller studies.			
Methodological comments			

Allocation to treatment groups: Prior to randomisation, infants were stratified for age and children (>3.5 years) for BMI. All participants were randomized to GH-treatment or no GH-treatment.
Blinding: A double blind placebo controlled study was considered unethical.
Comparability of treatment groups: Anthropometric parameters were similar in the two groups, although no p vals are presented.
Method of data analysis: Fat mass, Fat % and LBM were transformed into SDS adjusting for age and sex. LBM is related to height, so LBMHtSDS were computed by comparing LBM of PWS with LBM of health children with the same height and sex. Reference data for the dual energy X-ray absorptiometry were not available for children under the age of 4, so only those >4 years were included in the analysis. IGF-I and IGFBP-3 were transformed to SDS using sex- and age- matched Dutch references. Data were expressed as median (IQR) as most were not Gaussian distributed. Differences from baseline between groups were calculated using Mann Whitney U-tests. P vals are for change in GH group vs. control group
Sample size/power calculation: deLind van W reports that the power calculation estimated a total number of 40 patients (infants and prepubertal children) to yield a power of 0.80.
Attrition/drop-out: 2 excluded before treatment (one had a dose reduction due to high IGF-I levels, another had spinal surgery for scoliosis and two other medical problems). In total 4 infants and 5 children excluded from analysis – presumably due to incomplete study period for the other patients. Infants with repeated measures were older (p=0.025), possibly reflecting early diagnosis of PWS during recent years.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 3/10/08	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Festen et al., 2007 [91] Country Netherlands Study design RCT Number of centres not stated Funding: supported by Pfizer	1. 1mg/m2/day somatropin by sc injection (restricted to 0.5mg/m2/day in the 1st 4 weeks to avoid fluid retention). 2. no treatment Duration of treatment:2 years Other interventions used: caloric intake and activity levels standardized 3 months before study	Target population: prepubertal, generally not overweight children with PWS Number of Participants: Total: n=20 1. n=10 2. n=10 Sample attrition/dropout: none Inclusion criteria for study entry: Genetically confirmed diagnosis of PWS; age 4-9 yrs; prepubertal.	Primary outcomes: adiponectin levels, body composition, carbohydrate metabolism and triglyceride levels Secondary outcomes: associations between adiponectin and body composition, carbohydrate metabolism and triglyceride levels; effect of GH on these parameters Method of assessing outcomes: anthropometric measurements at baseline, year 1 and year 2 (standing height, weight, BMI); body composition assessed using dual energy x-ray absorptiometry; biochemical marker assays performed in the same laboratory Length of follow-up: 2 years

Characteristics of participants:

Median, IQR	1mg/m2/day GH (n=10)	No treatment (n=10)	P Value
N (male/female)	10 (5/5)	10 (3/7)	
Age (yr)	6.2 (5.1-7.1)	5.8 (4.9-7.8)	
Height SDS	-2.2 (-3.1 to -1.8)	-2.8 (-3.4 to -2.0)	
BMI (kg/m2)	16.9 (15.8 – 17.7)	17.3 (16.4-19.3)	
BMI SDS	0.8 (0.1 – 1.2)	1.1 (0.6 – 1.5)	
Adiponectin (mg/litre)	15.9 (13.3-23.9)	17.1 (13.1-23.1)	
Glucose (mmol/litre)	4.8 (4.6-5.0)	4.4 (4.3-4.7)	
Insulin (mU/litre)	6.0 (3.8-10.0)	5.5 (4.8-7.3)	
Insulin glucose ratio	1.3 (0.8-2.1)	1.3 (1.0-1.6)	
HOMA index	0.8 (0.5-1.3)	0.7 (0.6-0.9)	
Triglycerides (mmol/litre)	0.9 (0.7-1.7)	0.7 (0.6-1.0)	
IGF-I SDS	-1.7 (-2.2 to -1.2)	-1.7 (-2.9 to -1.0)	
IGFBP-3 SDS	-2.0 (-3.0 to -1.3)	-2.5 (-3.2 to -1.5)	
LBM SDS	-2.2 (-2.7 to -2.0)	-2.3 (-2.8 to -1.8)	

Fat mass SDS	0.8 (0.6 to 1.0)		0.8 (0.6 to 1.2)		
Percent fat SDS	1.7 (1.6 to 2.0)		1.8 (1.5 to 2.4)		
Trunk fat/total fat	0.44 (0.34 to 0.47)		0.4 (0.35 to 0.46)		
Comments Adiponectin levels were compared with healthy matched controls					
Results					
Outcomes Median, IQR	1mg/m2/day GH (n=10)		No treatment (n=10)		P Value change from baseline grp 1 vs. grp 2
	Year 1	Year 2	Year 1	Year 2	
Height SDS	-1.3 [a] (-1.7 to -0.8)	-0.6 [a] (-0.9 to -0.3)	-2.8 (-3.5 to -2.0)	-3.0 (-3.5 to -1.8)	[b] P<0.01
BMI (kg/m2)	16.1[c] (15.2-17.6)	16.3 (15.8 – 19.0)	18.5 (17.6 – 19.3)	18.5 (17.5-20.6)	[c] P<0.05
BMI SDS	0.2 [c] (-0.2 to 0.8)	0.4 (-0.3 to 1.1)	1.3 (1.0 – 1.6)	1.2 (0.9-1.5)	[c] P<0.05
Comments [a] p<0.05 compared with baseline p values for between-group tests corrected for multiple testing					
Adiponectin (mg/litre)	24.7 (15.0-25.9) a, b	24.6 (15.4-28.2) a, b	13.4 (11.6-21.4)	15.8 (12.5-19.2)	b P<0.05
Glucose (mmol/litre)	4.4 (4.2-5.0)	4.6 (4.2-5.0)	4.6 (4.3-4.8)	4.7 (4.3-4.9)	
Insulin (mU/litre)	9.0 (6.5-13.5) [a]	7.5 (6.0-11.5)	6.0 (3.3-8.3)	11.0 (6.0-24.0) [a]	
Insulin glucose ratio	2.1 (1.5-2.6) a	1.6 (1.5-2.2)	1.3 (0.8-1.9)	2.3 (1.4-2.2) [a]	
HOMA index	1.2 (0.8-1.8)	1.0 (0.7-1.5)	0.8 (0.4-1.0)	1.4 (0.8-3.0) [a]	
Triglycerides (mmol/litre)	0.8 (0.6-1.3)	0.7 (0.6-0.8)	0.6 (0.5-1.0)	1.0 (0.6-1.0)	
IGF-I SDS	2.3 (1.6-3.0) [a, c]	2.3 (2.1-2.9) [a, c]	-2.5 (-3.2 to -0.8)	-2.0 (-2.7 to 1.0)	[c] P<0.001
IGFBP-3 SDS	0.5 (-0.1 to 1.0) [a, c]	0.6 (0.4-1.1) [a, c]	-2.4 (-3.8 to -1.9)	-1.8 (-2.7 to -1.5)	[c] P<0.001
LBM SDS	-1.6 (-1.9 to -1.4) [a]	-1.2 (-1.7 to -1.1) [a]	-2.5 (-3.0 to -1.8)	-2.8 (-3. to 1.9) [a]	
Fat mass SDS	0.5 (0.2 to 1.0)	0.9 (0.4 to 1.4)	1.1 (0.9 to 1.2) [a]	1.2 (0.9 to 1.4) [a]	
Percent fat SDS	1.4 (0.9 to 1.7)[a]	1.7 (0.9 to 1.9)[a]	2.1 (1.8 to 2.2)	2.1 (1.9 to 2.4) [a]	
Trunk fat/total fat	0.4 (0.33 to 0.42)	0.41 (0.34 to 0.46)	0.41 (0.40 to 0.44)	0.41 (0.38 to 0.45)	
Comments Adiponectin levels were compared with healthy matched controls [a] p<0.05 compared with baseline p values corrected for multiple testing					
Adverse Effects – not reported					

Methodological comments Allocation to treatment groups: stratified by age and BMI prior to randomisation. No further details given. Blinding: open label trial Comparability of treatment groups: similar at baseline. Nb adeponectin levels were compared against healthy controls, not the untreated PWS group Method of data analysis: HtSDS and BMI SDS calculated from Dutch reference data. Most data not Gaussian distributed, so data expressed as median (interquartile range) and non parametric tests were used. Mann-Whitney U tests used for differences between groups. Adiponectin levels of PWS children were compared with reference data of healthy sex- and age-matched controls (n=40) with Wilcoxon signed rank test. Sample size/power calculation: not reported Attrition/drop-out: none

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	adequate
9. Did the analyses include an ITT analysis?	adequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 14/10/2008	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Festen et al., 2007 [92] Country The Netherlands and Sweden Study design RCT Number of centres multicentre Funding: Pfizer	1. GH 1mg/m2/day 2. no treatment Duration of treatment: 12 months Other interventions used: Dietary advice given and compliance evaluated every 3 months	Target population: PWS infants and toddlers Number of Participants: Total: 43 evaluated at baseline, then 29 entered treatment 1. n=15 2. n=14 Sample attrition/dropout: 14 were excluded from the study, and this appears to have taken place post-randomisation Inclusion criteria for study entry: Genetically confirmed	Primary outcomes: psychomotor development (BSID-II) (not data extracted as not per protocol) Secondary outcomes: Body composition; IGF-I and IGFBP-3 Method of assessing outcomes: Height measured with a Harpenden stadiometer; Dutch references used to calculate SDS for median height, BMI and head

		diagnosis of PWS; aged 6 months-3 years at start of protocol; Exclusion criteria – severe scoliosis (>20°); extremely low dietary intake	circumference; body composition in Dutch participants measured using dual energy X-ray absorptiometry; IGF in Dutch children measured using an immunometric technique, and in Swedish infants using a semi-illuminiscent technique Length of follow-up:12 months

Characteristics of participants:

Median (IQR)	GH 1mg/m2/day (n=15)	No treatment (n=14)	P Value
Gender (M/F)	7/8	8/6	
Age (years)	2.3 (1.7-3.0)	1.5 (1.2-2.7)	
Height SDS	-2.6 (-3.3 to -1.8)	-2.3 (-3.3 to -1.1)	
BMI (kg/m2)	16.3 (14.5 – 17.8)	15.9 (14.7 – 16.8)	
BMI SDS	-0.3 (-1.1 – 1.3)	-0.9 (-1.8 to -0.8)	
Head circumference SDS	-1.0 (-1.7 to -0.3)	-1.1 (-1.8 to -0.9)	
Body fat (%)	26.2 (22.2-28.9)	25.8 (23.1 – 27.7)	
LBM (%)	72.1 (69.8-75.7)	73.3 (70.9 – 75.2)	
IGF-SDS	-2.1 (-2.7 to -1.7)	-2.0 (-2.6 to -0.3)	
IGFBP-3SDS	-2.8 (-3.5 to -2.4)	-1.8 (-3.4 to -0.9)	

Results

Median (IQR)	GH 1mg/m2/day (n=15)	No treatment (n=14)	P Value
Age (years)	3.3 (2.7-4.0)	2.6 (2.3 – 3.8)	
Height SDS	-1.6† (-2.1 to -0.8)	-2.3 (-3.9 to -1.5)	
BMI (kg/m2)	16.4 (15.2 – 18.5)	15.5 (14.9-17.6)	
BMI SDS	0.3 (-0.9 – 1.8)	-0.4* (-0.8-1.3)	
Head circumference SDS	-0.2†‡ (-1.2 - 0.6)	-1.1‡ (-1.6 to -0.6)	
Body fat (%)	22.5 (11.3 – 33.2)	22.8 (19.5-32.9)	
LBM (%)	74.8 (63.7 – 82.3)	73.6 (61.6-75.9)	
IGF-SDS	1.7†¶ (0.1 – 2.5)	-2.6¶ (-4.1 to -0.4)	
IGFBP-3SDS	0.4*‡ (-0.3 -1.1)	-3.1‡ (-4.0 to -2.2)	

Comments
*p<0.05; †p<0.005: 12 vs. 0 months
‡P<0.05; ¶ p<0.001: GH vs. control

Adverse Effects
No results presented. Paper states that compared to randomized controls, GH did not induce disadvantageous effects on carbohydrate metabolism, sleep-related breathing disorders, and thyroid hormone levels.

Comments

Methodological comments
Allocation to treatment groups: children were stratified for age before randomisation. No further details given.

Blinding: open label Comparability of treatment groups: similar at baseline, although GH group had slightly older median age. Method of data analysis: For repeated measurement analysis, only children with 2 Bayley Scales of Infant Development II (BSID II) scores were included. BSID-II can only be used if developmental age is maximally 3-5 years. Non-parametric statistics used as data not Gaussian distributed. Mann-Whitney U tests used for 2-tail differences at baseline, one-tailed ANCOVA used for data analysis. Sample size/power calculation: not reported Attrition/drop-out: 14 of the original 43 were excluded from repeated BSID-II analysis, and therefore do not appear to have been randomised. However, the paper later states that results of 14 patients were excluded from analysis – not clear if this is the same 14, but assumed to be so, i.e. they were excluded post-randomisation. Reasons for exclusion: 5 children had not reached 1 year of study, 1 infant was excluded due to thyroid hormone deficiency, 8 had already passed the upper limit of BSID-II after 1 year of follow up (divided equally between the GH group and the control group).

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	unknown
2. Was the treatment allocation concealed?	unknown
3. Were the groups similar at baseline in terms of prognostic factors?	reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Inadequate

Reviewers: AT, LB		Date: 31/10/2008	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Haqq et al., 2003 [102] Country: USA Study design: Double blind placebo-controlled cross-over Number of centres: 1 Funding:	1. GH 0.043 mg/kg.d plus inactive ingredients, by daily s.c. injection 2. placebo injection of inactive ingredients, by daily s.c. injection Duration of treatment: 6 months in each treatment arm, 12 months overall.	Target population: children with PWS Number of Participants: Total: n=14 randomised, but data only given for n=12. 1. n=6 2. n=6 Sample attrition/dropout: n=2 Inclusion criteria for study entry: PWS; naïve to GH treatment Exclusion criteria: other	Primary outcomes: not stated Secondary outcomes: linear growth velocity, body composition pulmonary function, sleep, behaviour, cognition, resting energy expenditure (last 5 not DX as not per protocol) Method of assessing outcomes: assessed at 0.6 and 12 months; anthropometric

grants from the General Clinical Research Center and Pharmacia Corp.	Other interventions used: none	chronic illnesses; taking medications that impact on long-term bone mineralisation or body composition	measurements, side effects and compliance measured at 3 and 9 months; bone age determined at 0 and 12 months using Greulich and Pyle analysis of wrist x-rays; height measured at 0.6 and 12 months using wall-mounted stadiometer; body composition measured using dual-energy x-ray absorptiometry. Length of follow-up: 6 months for outcomes, 12 months overall

Characteristics of participants:

Mean ± SD	All patients (n=12)		P Value
Age, years	9.7 ± 3.3		
Sex	6m, 6f		
Bone age, years	10.0 ± 4.2		
BMI SDS	2.5 ± 0.7		
IGF-I ng/ml	169.3 ± 155.7		
IGF-I SDS	-1.10 ± 1.15		
IGFBP-3 ng/ml	2169 ± 1010		
IGFBP-3 SDS	-1.67 ± 1.10		
Mean height (cm)	128.9 ± 19.7		
BMI (kg/m2)	30.8 ± 8.3		
BMI (SDS)	2.5 ± 0.7		
HtSDS	-1.3 ± 1.2		
Growth velocity (cm/yr)	4.2 ± 2.3		
Body fat (%)	54 ± 5.3		
Fat mass (kg)	29.6 ± 16.7		
Lean mass (kg)	22.5 ± 10.9		
Lumbar spine BMD (SDS)	-0.51 ± 0.30		
Total BMC (g)	1263 ± 451		

Comments
Mean bone age also reported as 10.2 ± 4.1 yr later in the paper.

Results

Outcomes	GH 0.043 mg/kg.d (n=12)	Placebo (n=12)	P Value
BMI (kg/m2)	31.2 ± 8.9	32.8 ± 9.7	P<0.05
BMI (SDS)	2.4 ± 0.5	2.5 ± 0.6	
HtSDS	-1.2 ± 1.1	-1.3 ± 1.3	
Growth velocity (cm/yr)	7.5 ± 3.5	4.5 ± 2.7	P<0.05
Body fat (%)	49.7 ± 5.8	54.1 ± 5.6	P<0.05
Fat mass (kg)	26.1 ± 12.8	29.1 ± 14.1	P<0.05
Lean mass (kg)	24.1 ± 8.8	22.4 ± 8.5	P<0.05
Lumbar spine BMD (SDS)	-0.33 ± 1.4	-0.4 ±1.4	

Total BMC (g)	1337 ± 453	1342 ± 453	
IGF-I (ng/ml)	720 ± 379	232 ± 182	P<0.001
IGFBP-3 (ng/ml)	6029 ±1311	4247 ± 1209	P<0.01
Leptin (ng/ml)	49.7 ± 39.3	54.3 ± 46.2	P=0.06
Ghrelin (pmol/liter)	272 ± 204	361 ± 309	P=0.11
FT4 (pmol/liter)	12.9 ± 1.5	14.8 ± 1.4	P<0.05
TSH (mU/liter)	1.81 ± 0.79	2.04 ± 1.13	
Insulin (pmol/liter)	64.2 ± 42.6	64.2 ± 39	
Glucose (mmol/liter)	5.0 ± 0.7	4.8 ± 0.5	
Osteocalcin (nmol/liter)	10.5 ± 5.7	7.8 ± 5.9	P=0.06
Triglycerides (mmol/liter)	0.80 ± 0.52	0.92 ± 0.42	
Total cholesterol (mmol/liter)	4.7 ± 0.9	4.5 ± 1.7	
Comments Mean bone age (in all patients) increased to 11.3 ± 3.7 by the end of 12 months, compared with a chronological age of 9.7 ± 3.3 years. Mean height increased to 134.6± 19.3 cm. Only one patient required thyroid hormone replacement while receiving GH treatment.			
Adverse Effects No patient developed a significant degree of scoliosis (>20°). No evidence of impaired fasting glucose concentrations. GH treatment resulted in supranormal IGH-I and normal IGFBP-3 concentrations, but the consequences of this are unknown.			
Methodological comments Allocation to treatment groups: reported to be randomised, but no further details given Blinding: Both GH and placebo injections were given using a Genotropin pen. Comparability of treatment groups: data only presented for whole group - cross-over study design. Method of data analysis: Not ITT. Differences between groups calculated using paired t-tests. For data not distributed Normally, Wilcoxin sign-rank tests were used. P<0.05 considered statistically significant. Sample size/power calculation: not reported Attrition/drop-out: 2 patients withdrew – one due to relocation, one due to non-compliance with daily injections. Not clear which group they belonged to.			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Not reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Adequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	inadequate

Reviewers: AT, LB		Date: 20/10/2008	Version: checked
Reference and Design	Intervention	Participants	Outcome measures
Hauffa, 1997 [99] Country Germany Study design: Open RCT Number of centres: 1 Funding: Pharmacia and Upjohn, Germany	1. GH: 0.075 IU/kg/day for first month, then continued at dose of 0.15 IU/kg/day to a maximum of 8 IU/day 2. no treatment Duration of treatment: 2 year study with control arm during 1st year. Other interventions used: not stated	Target population: children aged 3-12 with PWS Number of Participants: Total: n=19 randomised, n=17 included in study, n=16 analysed 1. n=8 2. n=9 Sample attrition/dropout: 2 not entered following randomisation, 1 excluded from analysis due to AE-related dose reduction Inclusion/exclusion criteria for study entry: Prepubertal 3 to 12 years old Prader-Willi syndrome (confirmed by molecular genetics) projected final Ht < 3rd centile for German population	Primary outcomes: not stated Secondary outcomes: changes in HtSDS; growth velocity SDS; IGF-I; IGFBP-3 Method of assessing outcomes: not reported Length of follow-up:1 year

Characteristics of participants:			
Mean ± SD	GH 0.15 IU/kg/day (n=7)	No treatment (n=9)	P Value
Age, years	8.25 ± 2.4	7.56 ± 2.0	
Sex – female/male	3/4	4/5	
Bone age (years)	7.91 ± 4.3	6.76 ± 2.4	
Height (cm)	120.9 ± 16.3	120.5 ± 11.2	
Weight (kg)	35.9 ± 18.2	32.5 ± 8.7	
Hip circumference (cm)	78.8 ± 19.6	77.6 ± 11.5	
Target height (cm)	172.9 ± 8.5	174.8 ± 8.2	
Results			
Outcomes	GH 0.15 IU/kg/day (n=7)	No treatment (n=9)	P Value
HV SDS	5.5	-2.3	P=0.0012
Height SDS	+1.07	-0.25	
IGF-I	Increased significantly (P<0.008), sometimes to above the upper limit of the reference range	'at or slightly below lower limit of reference range'	
IGFBP-3	Increased significantly (P<0.008), mostly to above the upper limit of the reference	'within normal range'	

	range		
Comments Height gain (+1.02 SD) remained unchanged when analysed in relation to bone age. No significant within- or between-group changes were detected for sitting height, BMI, skinfold thickness, waist or hip circumference or serum lipids.			
Adverse Effects 1 patient in GH group developed pseudotumour cerebri after increasing the starting dose to the final dose. Symptoms resolved on discontinuation. No abnormalities of glucose regulation observed in either group.			
Methodological comments Allocation to treatment groups: randomised (method not stated) Blinding: open label Comparability of treatment groups: similar at baseline Method of data analysis: no details given Sample size/power calculation: not reported Attrition/drop-out: 19 randomised, 2 not entered (reasons not stated), 1 not included in analysis (discontinued after an AE then resumed at half the dose)			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	unknown
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	inadequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	Inadequate

CRI Data extraction forms

Reviewers: AT, LB		Date: 19/09/08	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Sanchez et al, 2002 [103] Country USA Study design RCT Number of centres 1 Funding: partly funded by Genetech Foundation for Growth and Development, and the Casey Lee Ball Foundation.	1. 0.05 mg/kg rhGH, daily subcutaneous injection 2. no treatment Duration of treatment: 12 months Other interventions used: All patients received either monoclonal or polyclonal anti-T cell therapy and were maintained on a 3-drug immunosuppressive regimen. None were given vitamin D sterols, oral calcium supplements, or anti-convulsant medications	Target population: Prepubertal paediatric kidney allograft recipients Number of Participants: Total: 23 1. 12 2. 11 Sample attrition/dropout: gp1: 1; gp2: 1 Inclusion criteria for study entry: Pre-pubertal children Stable renal function for at least 1 year post op Normal bone formation rates Pts with adynamic lesions who had not previously been treated with rhGH were also included Exclusion criteria Secondary hyperparathyroidism	Primary outcomes: Appears to be skeletal changes, but not stated clearly Secondary outcomes: ht SDS; wt SDS; growth velocity Method of assessing outcomes: Ht and Wt measured at 3-month intervals; ht measured using fixed wall-mounted stadiometer; bone biopsy and histomorphometry bone mass measured by dual-energy X-ray absorptiometry; blood samples every 3 months; bone age determined by Greulich and Pyle method from X-rays of left hand and wrist Length of follow-up: 12 months

Characteristics of participants:			
	0.05 mg/kg rhGH (n=12)	No treatment (n=11)	P Value
Mean age ± SD, years	9.7 ± 4.5	11 ± 1.8	n/s
Sex	18 boys, 5 girls		
Mean interval since transplantation (yrs)	3.4 ± 2.5		
SDS for height	-2.0 ± 1.1	Not given, but 'did not differ' stated	
Mean SDS for height 12 mths before study	-2.2 ±0.8	-2.6 ±1.0	n/s
Annual growth velocity 12 mths before study cm/yr	5 ±2.0	4 ±2.0	n/s
Bone age (yrs)	7.1 ± 3.6	8.8 ± 2.4	n/s
Tanner score	1.9 ± 0.8	2.1 ± 1.1	n/s
Glomerular filtration rate (ml/min)	58 ± 15	58 ± 14	
Results			

Outcomes (mean ± SE)	0.05 mg/kg rhGH (n=12)		No treatment (n=11)		P Value
SDS for height at end of study	-1.1 ± 1.0 (p<0.02 compared with baseline)		No change from baseline		
Annual growth velocity (cm/yr)	8.0 ± 2.1		4.8 ± 1.7		P<0.01
Change in SDS for weight	0.2 ± 0.3		-0.3 ± 0.3		P<0.01
Bone age (yrs)	8.5 ± 3.4		9.5 ± 2.8		n/s
Tanner score	1.9 ± 0.7		2.2 ± 1.0		n/s
Glomerular filtration rate (ml/min)	61± 13 (change from baseline p=n/s)		67± 19 (change from baseline p=n/s)		
Biochemical markers	baseline	final	baseline	final	
Serum calcium (mg/dl)	9.8 ± 0.7	10±0.6	9.4 ± 0.5	9.6 ± 0.7	
Serum phosphorous (mg/dl)	4.8 ±0.8	4.8 ± 0.7	4.5 ± 0.8	4.2 ± 0.7	
Serum osteocalcin (ng/ml)	24 ±2.7	24 ± 0.3	20 ± 2.3	17 ±1.7	
Serum parathyroid hormone (pg/ml)	55 ±5.0	55 ± 5.3	38 ±4.0	34 ± 2.5	
Serum alkaline phosphate (IU/I)	239 ± 9.0	255 ± 9.0	225 ± 9.0	198 ± 6.4	
Serum 1,25-dihydroxyvitamin D (pg/ml)	43 ± 4.3	52 ± 4.7	39 ± 3.3	50 ± 3.1	
Bone histomorphology	baseline	final	baseline	final	
Bone area (%)	20 ± 2.6	21 ± 4.0	20 ± 4.8	22 ± 6.4	
Osteoid area (%)	8.8 ± 4.0	7.9 ± 1.8	6.1 ± 2.5	8.2 ± 2.3	
Eroded perimeter (%)	5.4 ± 4.8	4.0 ± 2.2	2.2 ± 1.7	3.0 ± 1.5	
Bone formation rate (μm2/mm2 per day)	266 ± 212	348 ± 304	262 ± 180	390 ± 232	
SDS for bone mass at lumber spine based on chronological age	-0.1 ± 1.6	-0.1 ± 1.3 (p=n/s)	-1.7 ± 0.9	-2.1 ± 1.0 (p<0.5)	
SDS for bone mass corrected for height-age	1.1 ± 1.3	0.7 ±0.8 (change from baseline p=ns)	0.01 ± 1.0	-0.3 ±1.2 (p<0.05 change from baseline)	
Comments Baseline serum levels of calcium, phosphorous, parathyroid hormone, alkaline phosphate, osteocalcin, and 1,25-dihydroxyvitamin D did not differ between patients given rhGH and untreated controls. Values remained unchanged after 12 months follow-up in both groups. IGF-I baseline values were similar between groups (actual values not given), and did not change from baseline in the untreated group. Change from baseline was significant for the treated group					

(p<0.001), although subgroup analysis indicated that this was only in the subgroup of patients with normal rates of bone formation, who experienced an increase in serum IGF-I levels of 54 ± 25% after 3 months and 98 ± 35% after 12 months of rhGH (p<0.05). Serum IGF-I levels remained unchanged in patients with adynamic bone, and values did not differ from those obtained in the untreated group. Cumulative dose of predisone did not differ between groups. Two patients with normal rates of bone formation experienced acute rejection episodes after 3 and 12 months of rhGH therapy. One was associated with non-compliance to immunosuppressive medications. Both episodes reversed after treatment with methylprednisolone. No rejection episodes in untreated pts.
Methodological comments Allocation to treatment groups: statistician who had no information about patients' clinical or biochemical characteristics randomized to treatment groups depending on their initial bone histological finding. Details of randomization procedure not given. Not stratified by height etc. Blinding: control group did not receive placebo injections Comparability of treatment groups: p=n/s for difference in age at baseline Method of data analysis: Not ITT as 2 pts who withdrew were excluded from analysis. Unpaired T-tests were used to compare changes from baseline. Sample size/power calculation: sample size estimated with 80% power to detect differences in group means and a 2-group comparison that required 20 pts per group. Appears to have been based on bone formation rates in a previous study, and it is not clear what the primary outcome for the present study is. Attrition/drop-out: 2 withdrawals: 1 in gp1 due to glucose intolerance after 3 months (which resolved in stopping treatment). 1 in gp2 due to being assigned to control group. 2 gp1 pts also failed to undergo 2nd bone biopsy.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	unknown
2. Was the treatment allocation concealed?	unknown
3. Were the groups similar at baseline in terms of prognostic factors?	adequate
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	partial
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 4/10/08	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
Broyer et al., 1996 [107] Country international	1. daily s.c. injection of GH (1 IU/kg/week) 2. no treatment	Target population: children who had received a kidney transplant Number of Participants: Total: n=203	Primary outcomes: GFR Secondary outcomes: transplant rejections; GV; HtSDS

Study design open label RCT Number of centres: multicentre Funding: Pharmacia & Upjohn	Duration of treatment:1 year of randomised treatment, followed by 1 year of GH treatment for both groups (only year 1 randomised data included here) Other interventions used: not reported	1. n=106 2. n=97 Sample attrition/dropout: 23 excluded from analysis of renal function; 49 excluded from analysis of growth Inclusion criteria for study entry: ≥ 12 months since transplantation; 2 ht measurements over last 6 mths; ht SDS <-2 or growth velocity below the 25th centile; GFR ≥ 20ml/min/1.73m2; normal serum thyroid hormone levels; testicular volume <8ml or breast development <B2 Exclusion criteria: ht velocity ≥ 75th centile, dialysis therapy, any form of malignancy or treatment with GH during past 12 mths.	Nb data only extracted where reported separately for prepubertal children Method of assessing outcomes: auxological and biochemical assessments every 3 mths. GFR measured by insulin clearance, or creatinine clearance (Morris method) Length of follow-up: 1 year (later follow-up not data extracted as not randomised)

Characteristics of participants:			
Mean ± SD	1 IU/kg/week GH	No treatment	P Value
Boys/girls	71/35	72/25	
Age, years	12.6 ± 3.4	12.1 ± 3.1	
Proportion prepubertal (%)	53	63	
Yrs since transplantation	3.6 ±2.3	3.2 ± 2.4	
Proportion cadaver donors (%)	81	86	
Height SDS	-3.2 ±1.4	-3.1 ±1.1	
Height velocity before treatment (cm/yr)	3.6 ±2.2	4.0 ±2.1	
GFR (insulin)(ml/min/1.73m2)	48 ±27	48 ±26	
GF (Morris) (ml/min/1.73m2)	51 ±21	51 ±2.1	
Rejection episodes prior to study (n)			
0-1 episode	69	63	
2-4 episodes	30	32	
5-8 episodes	7	1	
Comments N not clear for patient groups at baseline			
Results			
Mean ± SD change from baseline	1 IU/kg/week GH (n=28)	No treatment (n=30)	P Value
Change in GV (cm/yr)	3.7 ±1.6	0.3± 1.6	P<0.0001
Change in HtSDS	+0.6 ± 0.3	+0.1 ± 0.3	P<0.0001

Comments
Primary outcome (GFR) and other outcomes not data extracted as not reportedly separately for prepubertal children.
Methodological comments Allocation to treatment groups: randomised centrally, but no further details given Blinding: open label Comparability of treatment groups: no p values given. Appear to be similar, although control group contained 10% more prepubertal patients than treatment group and no. of patients with a high no. of acute rejections was higher in the GH-treated patients (7 vs. 1) Method of data analysis: no information given Sample size/power calculation: not stated Attrition/drop-out: 23 excluded from analysis of renal function (treatment occurred without randomisation, GFR<20ml/min/1.73m2; transplantation<12 mths before study entry; non-compliance); 49 excluded from analysis of growth (abnormal thyroid function, growing too well (or not being short enough) before the study, previous growth not documented).

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: AT, LB		Date: 24/10/2008	Version: checked
Reference and Design	Intervention	Participants	Outcome measures
Fine et al., 2004 [108] Country: USA Study design: RCT Number of centres: 17 Funding: Genentech	1. GH 0.05 mg/kg/day s.c. 2. placebo in equivalent volume dose adjusted every 3 months for change in weight Duration of treatment: 2 years treatment was discontinued at renal transplantation,	Target population: pre-pubertal growth-retarded children with CRF Number of Participants: Total: n=125 1. n=82 2. n=43 Sample attrition/dropout: grp 1 13 in year 1, 13 in year 2; grp2 12 in year 1, 3 in year 2 Inclusion criteria for study entry:	Primary outcomes: not stated Secondary outcomes: GV HtSDS Height age (HA) Bone age (BA) Cumulative ΔHA - ΔBA Weight gain Triceps skin-fold thickness (TSF) mid-arm muscle circumference (MAMC)

	significant adverse event, or when BA >15 years for boys and >14 years for girls and growth rate was < 2cm/yr. Treatment was paused if a patient's height percentile exceeded the Tanner target percentile for mid-parental height (4/82 grp1, 11/42 grp 2). Other interventions used: dialysis was permitted as required; multivitamins, vitamin D analog and various other therapies were permitted as required.	Irreversible renal insufficiency Creatinine clearance > 5 and < 75 ml/min/1.73 m2 height < third percentile for chronologic age bone age < 10 yr for girls and < 11 yr for boys prepubertal status (Tanner stage I) exclusion criteria: evidence of a specific cause for growth failure other than CRF inability to obtain accurate height measurements use of corticosteroids or other medications that influence growth diabetes mellitus, active malignant disease or treatment of a malignant disease within past year use of any other investigational drug therapy within 2 months of randomisation.	Method of assessing outcomes: anthropometric measurements made by same observer every 3 months; radiologic evaluation of bone age every 6 months. Length of follow-up: 2 years

Characteristics of participants:

Mean ± SD	GH 0.05 mg/kg/day (n=82)	Placebo (n=43)	P Value
Age, years	6.0 ± 3.9	5.7 ± 3.6	
Sex	21 female; 61 male	14 female; 28 male	
Height age	4.0 ± 2.9	3.8 ± 2.8	
Bone age	4.2 ± 3.0	4.2 ± 2.9	
HtSDS	-2.9 ± 0.9	-2.9 ± 1.0	
Standardized height	-2.94 ± 0.86 (n=55)	-2.82 ± 0.97 (n=27)	
IGF-I (μg/L)	121 ± 73 (n=47)	141 ± 94 (n=20)	
Fasting insulin (pmol/L)	70.3 ± 43.6 (n=40)	87.8 ± 71.1 (n=21)	
Postprandial insulin (pmol/L)	25.8 ± 26.8 (n=43)	30.1 ± 14.6 (n=19)	
Fasting glucose (mmol/L)	5.1 ± 1.1 (n=49)	5.0 ± 0.7 (n=24)	
Postprandial glucose (mmol/L)	5.3 ± 1.8 (n=37)	6.0 ± 1.7 (n=21)	
Hemoglobin A1c (%)	5.1 ± 0.9 (n=48)	5.4 ± 1.0 (n=24)	
Creatinine (μmol/L)	174 ± 111 (n=48)	173 ± 97 (n=24)	
Creatinine (mg/dl)	2.3 ± 1.5 (n=48)	2.3 ± 1.3 (n=24)	

Creatinine clearance (ml/sec/1.73m2)	0.55 ± 0.33 (n=48)	0.52 ± 0.31 (n=24)	
Creatinine clearance (ml/min/1.73m2)	32.8 ± 19.5 (n=48)	31.1 ± 18.3 (n=24)	
Blood urea nitrogen (mmol/L)	15.6 ± 6.6 (n=48)	16.0 ± 7.3 (n=24)	
Blood urea nitrogen (mg/dl)	43.6 ± 18.5 (n=48)	44.9 ± 20.5 (n=24)	
Results			
Mean ± SD	GH 0.05 mg/kg/day (n=82)	Placebo (n=43)	P Value
GV year 1 (cm/yr)	10.7 ± 3.1 (n = 55)	6.5 ± 2.6 (n = 27)	p< 0.00005
GV year 2 (cm/yr)	7.8 ± 2.1 (n = 55)	5.5 ± 1.9 cm, (n = 27)	p< 0.00005
HtSDS at year 2	–1.6 p< 0.00005 compared with baseline	–2.9 P=0.52 compared with baseline	
Roche-Wainer-Thissen predicted adult height at 2 years (cm)	+5.4	-0.4	p< 0.00005
Weight gain after 2 years (kg)	6.7 ± 2.2	4.6 ± 2.7	p = 0.0004
Triceps skin-fold thickness (mm)	-1.6 ± 2.6	+0.6 ± 3.8	p = 0.006
Mid-arm muscle circumference (cm)	2.1 ± 1.1	1.3 ± 1.2	p = 0.007
Change in BA at 2 years (years)	2.3 ± 0.7	1.6 ± 0.5	P=0.0001
Standardised height (1 year)	-1.93 ± 1.01 (n=55)	-2.90 ± 0.95 (n=27)	
Cumulative change in HA – change in BA (year 1)	0.28 ± 0.45 (n=43)	-0.04 ± 0.36 (n=21)	
Cumulative change in HA – change in BA (year 2)	0.15 ± 0.62 (n=43)	-0.12 ± 0.43 (n=21)	P=0.08
Standardised height (2 year)	-1.55 ± 1.16 (n=55)	-2.91 ± 1.04 (n=27)	P<0.00005
Height age (1 year)	4.5 ± 2.7 (n=43)	5.0 ± 3.2 (n=21)	
Height age (2 year)	5.6 ± 2.9 (n=43)	5.7 ± 3.3 (n=21)	P<0.00005
Bone age (1 year)	4.6 ± 2.6 (n=43)	5.2 ± 3.1 (n=21)	
Bone age (2 year)	5.8 ± 2.8 (n=43)	6.0 ± 3.2 (n=21)	P=0.0001
IGF-I (μg/L) year 1	286 ± 158 (n=47)	167 ± 97 (n=20)	P=0.0004
IGF-I (μg/L) year 2	244 ± 128 (n=47)	135 ± 80 (n=20)	P=0.0001
Fasting insulin (pmol/L) year 1	104.9 ± 54.5 (n=40)	76.9 ± 28.4 (n=21)	
Fasting insulin (pmol/L) year 2	80.9 ± 42.8 (n=40)	59.1 ± 34.6 (n=21)	P=0.03
Postprandial insulin (pmol/L) year 1	36.6 ± 29.0 (n=43)	27.7 ± 17.2 (n=19)	

Postprandial insulin (pmol/L) year 2	29.0± 20.7 (n=43)	27.2 ± 16.9 (n=19)	P=0.32
Fasting glucose (mmol/L) year 1	5.2 ± 0.6 (n=49)	5.2 ± 1.0 (n=24)	
Fasting glucose (mmol/L) year 2	5.0 ± 0.6 (n=49)	5.1 ± 0.7 (n=24)	P=0.70
Postprandial glucose (mmol/L) year 1	5.4 ± 1.1 (n=37)	5.1 ± 1.2 (n=21)	
Postprandial glucose (mmol/L) year 2	5.4 ± 1.1 (n=37)	5.5 ± 1.1 (n=21)	P=0.28
Hemoglobin A1c (%) year 1	5.0 ± 0.8 (n=48)	5.0 ± 0.8 (n=24)	
Hemoglobin A1c (%) year 2	4.9 ± 0.7 (n=48)	5.0 ± 0.8 (n=24)	P=0.33
Creatinine (µmol/L) Year 1	218 ± 163 (n=48)	192 ± 96 (n=24)	
Creatinine (µmol/L) Year 2	269 ± 205 (n=48)	219 ± 114 (n=24)	P=0.08
Creatinine (mg/dl) Year 1	2.9 ± 2.1 (n=48)	2.5 ± 1.3 (n=24)	
Creatinine (mg/dl) Year 2	3.5 ± 2.7 (n=48)	2.9 ± 1.5 (n=24)	P=0.08
Creatinine clearance (ml/sec/1.73m2) Year 1	0.55 ± 0.42 (n=48)	0.51 ± 0.33 (n=24)	
Creatinine clearance (ml/sec/1.73m2) Year 2	0.49 ± 0.35 (n=48)	0.48 ± 0.34 (n=24)	P=0.63
Creatinine clearance (ml/min/1.73m2) Year 1	32.8 ± 25.2 (n=48)	30.7 ± 19.9 (n=24)	
Creatinine clearance (ml/min/1.73m2) Year 2	29.3 ± 21.3 (n=48)	28.9 ± 20.4 (n=24)	P=0.63
Blood urea nitrogen (mmol/L) Year 1	16.1 ± 8.8 (n=48)	17.7 ± 8.7 (n=24)	
Blood urea nitrogen (mmol/L) Year 2	17.2 ± 8.7 (n=48)	15.9 ± 7.1 (n=24)	P=0.26
Blood urea nitrogen (mg/dl) Year 1	45.0 ± 24.5 (n=48)	49.7 ± 24.4 (n=24)	
Blood urea nitrogen (mg/dl) Year 2	48.2 ± 24.5 (n=48)	44.5 ± 20.0 (n=24)	P=0.26
Serum alkaline phosphatase level change from	120.1 ± 130.1 (n=48)	45.6 ± 90.0 (n=24)	P=0.014

baseline (IU/L) Year 1			
Serum alkaline phosphatase level change from baseline (IU/L) Year 2	Not reported		p=n/s
Comments Mean fasting insulin levels changed significantly in GH patients between baseline and 12 months (p=0.0005) but not between baseline and 24 months. Changes in placebo group were not significant. Postprandial insulin levels also significant for GH group between baseline and year one (p=0.0089) but not significant between baseline and 24 months. Changes from baseline in placebo group were not significant. No significant change in haemoglobin A1c or thyroxine or thyroid-stimulating hormone in either group at either time period. Biochemical measurements: There was no significant difference in the variation in the serum calcium, phosphorous, triglyceride, or cholesterol levels between the two groups during the first 2 years of treatment.			
Adverse Effects No differences between groups in year 1. Year 2 asthma or wheezing in 8 of 55 GH patients and none of placebo. All episodes preceded by upper respiratory tract infections. "No clinically significant side effects were associated with rhGH treatment." During the 1st 12 months, 19 of 82 patients had low titer GH antibodies (i.e. anti-GH antibody serum binding by radioimmunoassay at least twice background values after 10-fold dilution), but over 2 years there was no significant difference in growth rate between patients who acquired anti-growth hormone antibodies and those who did not.			
Methodological comments Allocation to treatment groups: No information on randomisation except performed to place 2/3 in treatment and 1/3 in placebo and to maintain balance in age, sex, standardised height, degree of renal function, and primary renal disease Blinding: placebo used in equivalent volume, but no further detail given. Comparability of treatment groups: IGF-I and fasting insulin levels were higher in the placebo group, but were not reported to have been significantly different. Method of data analysis: Between- and within-group comparisons were made with 2-tailed t tests; $p<0.05$ was considered statistically significant. Many outcome measures are only presented for patients who completed both years of the study. Not ITT Sample size/power calculation: not reported Attrition/drop-out: GH: 13 year 1, 13 year 2. Placebo: 12 year 1, 3 year 2. 41% of total withdrawals were due to renal transplant, 24% requested removal, 15% non-compliance.			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	adequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate

9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 9/09/2008	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Hokken-Koelega et al., 1991 [104] Country: international Study design cross-over RCT Number of centres: multicentre Funding: Novo-Nordisk A/S Denmark	1. 4 IU/m2 biosynthetic human GH daily subcutaneous injection, followed by cross-over to placebo 2. placebo followed by cross-over to biosynthetic human GH daily subcutaneous injection Duration of treatment: 6 months in each arm of the study Other interventions used: phosphate binding medication, calcium supplements and 1,25-(OH)2 vitamin D.	Target population: prepubertal children with CRF and severe growth retardation Number of Participants: Total: 20 1. 8 2. 8 Original assignment not stated Sample attrition/dropout: 4 left due to kidney transplantation Inclusion/exclusion criteria for study entry: Chronic renal failure ≥ 1 year Creatinine clearance below 20 ml/min/1.73m2 Height SDS for age < -1.88 and HV for age < 25th percentile Prepubertal (Tanner stage I) Bone age < 10 years for girls and 12 years for boys No evidence of growth retardation cause other than CRF Normal thyroid function No osteodystrophy No previous treatment with anabolic steroids, sex steroids, or recombinant human erythropoietin.	Primary outcomes: not stated Secondary outcomes: GV; GV SDS; bone age (yr)IGF-I and IFG-II plasma concentrations Method of assessing outcomes: Ht measured with a Harpenden stadiometer; bone-age calculated from X-rays at start of study and every 6 months. Length of follow-up:12 months

Characteristics of participants:

Median, range	4 IU/m2hGH/placebo (n=8)	Placebo/4 IU/m2hGH (n=8)	P Value
Age, years	8.7 (4.4 to 11.3)	8.6 (4.4 to 16.0)	
Sex	6 male, 2 female	4 male, 4 female	
Bone age, years	7.4 (3.7 to 10.2)	7.5 (3.7 to 10.6)	
HtSDS	-2.3 (-3.9 to -1.8)	-2.7 (-5.6 to -2.0)	
GV (cm/6mo)	1.6 (0 to 3.0)	1.4 (0.2 to 2.6)	
Weight for Height (%)	98.2 (86.7 to 113.5)	101.5 (90.3 to 116.5)	
Mean (SD) GV (cm/6mo)	1.5 (0.7)	1.5 (0.5)	

6 mths pre-study				
Mean (SD) HV SDS 6 mths pre-study	-3.2 (1.4)		-2.9 (2.0)	
Mean (SD) bone age (yr) 6 mths pre-study	6.9 (2.3)		7.7 (2.6)	
Mean (SD) IGF-I ng/ml SDS for bone age	173 (135) 0.8 (2.7)		197 (94) 1.4 (1.6)	
Mean (SD) IGF-II ng/ml SDS for bone age	1160 (485) 2.5 (3.0)		1178 (483) 3.4 (4.0)	
Mean (SD) IGFBP-3 ng/ml SDS for bone age	5429 (1352) 3.2 (1.1)		6559 (2552) 4.2 (2.1)	
Mean (SD) IGFBP-1 ng/ml SDS for bone age	195 (126) 30 (20)		190 (115) 29 (17)	

Results

Outcomes	4 IU/m2hGH/placebo (n=8)		Placebo/4 IU/m2hGH (n=8)		Overall mean effect of GH minus effect of placebo
	After 6mths GH	After 6mths placebo	After 6mths placebo	After 6mths GH	
Mean (SD) GV (cm/6mo)	5.2 (1.2)	1.5 (0.4)	2.4 (1.0)	4.4 (1.6)	2.9 [95% CI 2.3, 3.5] (p<0.0001)
Mean (SD) HV SDS	6.9 (2.4)	-3.0 (1.6)	-0.5 (3.2)	5.0 (4.5)	7.7 (p<0.0001)
Mean (SD) bone age (yr)	7.0 (1.9)	7.6 (1.7)	8.0 (2.6)	8.4 (2.8)	-0.01
Mean (SD) IGF-I ng/ml SDS for bone age	264 (168) 2.6 (2.0)	160 (104) -0.2 (1.5)	160 (95) 0.3 (1.6)	268 (120) 2.9 (2.0)	106 2.7 (p<0.0001)
Mean (SD) IGF-II ng/ml SDS for bone age	1174 (361) 2.8 (2.8)	983 (336) 0.9 (2.2)	1192 (340) 3.4 (2.4)	1346 (492) 4.6 (3.4)	172 1.6
Mean (SD) IGFBP-3 ng/ml SDS for bone age	7708 (2323) 5.0 (1.3)	6102 (1892) 3.7 (1.3)	6501 (1988) 3.9 (1.4)	8706 (2275) 5.2 (1.4)	1906 1.3 (p<0.0001)
Mean (SD) IGFBP-1 ng/ml SDS for bone age	119 (95) 16.4 (16.8)	185 (119) 27.1 (22.4)	215 (106) 32 (19.5)	140 (90) 20 (16.6)	-70 (p<0.0001) -11.2 (p<0.0001)

Comments

For growth velocity, there was no significant carry-over effect (-0.04 cm/6mths, p=0.94). Period check was -0.9cm/6 months (p<0.06).

Adverse Effects

Serum alkaline phosphate was significantly increased during GH treatment, but returned to pre-treatment levels when GH therapy was replaced by placebo (p<0.0001). There was no significant change in parathyroid hormone concentration during either treatment schedule. Thyroid function

was normal.
Methodological comments Allocation to treatment groups: States randomly and blindly assigned, but no further details given Blinding: stated to be double blind Comparability of treatment groups: Similar at baseline, although IGF-I and IGFBP-3 were higher in group 2 at baseline Method of data analysis: Baseline height expressed as SDS for chronological age compared with Dutch reference data. Height velocity expressed as SDS for chronological age compared with references derived from Infant-childhood-puberty model. Not ITT. Paper states that statistical methods appropriate for cross-over trials were used, but no further details were given. Treatment effects were calculated and tested after taking into account any period effect. Sample size/power calculation: no information in paper Attrition/drop-out: 4 children left the study to have kidney transplants, 3 at 6 mths and 1 at 7 mths.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	unknown
2. Was the treatment allocation concealed?	unknown
3. Were the groups similar at baseline in terms of prognostic factors?	reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	unknown
6. Was the care provider blinded?	unknown
7. Was the patient blinded?	unknown
8. Were the point estimates and measure of variability presented for the primary outcome measure?	adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	adequate

Reviewers: AT, LB		Date: 16/9/08	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Hokken-Koelega et al., 1996 [105] Country: International Study design cross-over RCT Number of centres: multicentre Funding: Novo Nordisk A/S	1. 4 IU/m2 GH / placebo daily subcutaneous injection 2. placebo / 4 IU/m2 GH daily subcutaneous injection Duration of treatment: 6 months in each arm Other interventions used: immunosuppressive therapy	Target population: prepubertal children after renal transplant Number of Participants: Total: n=11 1. n=6 2. n=5 Sample attrition/dropout: none Inclusion criteria for study entry: Post-renal transplant (≥ 12 months) Stable condition without rejection episodes (≥ 12 months) Height SDS for age < -1.88 and HV for age < 50th percentile OR HtSDS above –1.88 with HV < 25th percentile Prepubertal (Tanner stage I) Bone age < 10 years for girls and 12 years for boys Prednisone dose ≤ 0.25 mg.kg/day ≥ 6 months No evidence of growth retardation cause other than following renal transplant Normal thyroid function and acid-base balance No previous treatment with sex steroids	Primary outcomes: not stated Secondary outcomes: HV; HVSDS; Bone age; GFR; ERPF; IGF-I measures; insulin and other biochemical markers Method of assessing outcomes: same investigator examined children at enrolment and every 3 months; height measured with a Harpenden stadiometer until 3 consecutive readings within 0.2cm; GV references derived from Infant-childhood-Puberty model; Dutch reference data used for baseline HtSDS; bone age determined from wrist x-rays Length of follow-up: 12 months

Characteristics of participants:

Median, range	4 IU/m2 GH / placebo (n=6)	Placebo / 4 IU/m2 GH (n=5)	P Value
Age, years	12.1 (9.1 to 18.7)	11.1 (8.3 to 14.9)	
Sex	5 male/ 1 female	4 male / 1 female	
HtSDS	-3.0 (-7.6 to -1.2)	-2.6 (-3.6 to -2.1)	
GV (6mo)	1.4 (0.5 to 2.6)	0.8 (0.6 to 1.8)	
BMI SDS	3.1 (-1.1 to 4.2)	1.3 (-0.2 to 3.7)	
Glomerular filtration rate ml/min/1.73m2	62 (56-81)	38 (19-74)	
Bone age (yr)	9.5 (7.9 – 11.5)	7.5 (5.2 – 10.5)	

Results

Outcomes	4 IU/m2 GH / placebo (n=6)	Placebo / 4 IU/m2 GH (n=5)	Overall mean effect

							of GH minus effect of placebo
Mean (SD)	prestudy	After 6mths GH	After 6 mths placebo	Pre-study	After 6mths placebo	After 6 mths GH	
HV cm/6mths	1.5 (0.7)	5.3 (1.0)	1.5 (0.9)	1.0 (0.5)	1.9 (0.7)	3.9 (1.3)	+2.9 [95% CI 1.9, 3.9] (p<0.0001)
HVSDS	-1.7 (1.8)	9.1 (2.9)	-1.3 (2.9)	-3.3 (0.9)	-0.4 (1.7)	5.3 (4.0)	+8.0 (p<0.0001)
Bone age, yrs	9.5 (1.7)	9.7 (1.4)	10.5 (2.2)	7.7 (2.2)	8.0 (2.1)	8.1 (1.2)	-0.5
GFR ml/min/1.73 m2	66 (13)	80 (30)	64 (1)	44 (22)	49 (22)	47 (38)	+5.5
ERPF ml/min/1.73 m2	261 (75)	254 (87)	264 (77)	173 (79)	191 (62)	184 (86)	-15.6
IGF-I ng/ml SDSBA	280 (121) 0.9 (1.6)	594 (180) 5.4 (2.8)	240 (143) 1.0 (2.5)	274 (89) 2.8 (1.8)	321 (94) 3.4 (0.5)	488 (237) 6.4 (1.9)	228 +3.7 (p<0.0001)
IGF-II ng/ml SDSBA	759 (114) 0.5 (0.9)	799 (186) 1.1 (1.7)	689 (31) 0.0 (0.4)	728 (349) 0.9 (3.2)	898 (56) 2.2 (1.2)	900 (63) 2.3 (1.0)	73 +0.5
IGFBP-3 ng/ml SDSBA	4902 (1099) 2.8 (1.8)	7457 (2088) 4.5 (1.5)	5681 (1588) 3.7 (2.9)	5787 (1037) 3.8 (0.7)	6228 (2193) 3.9 (1.5)	8495 (2921) 5.3 (1.5)	1698 +0.9
IGFBP-1 ng/ml SDSBA	52 (32) 4.7 (4.6)	52 (23) 4.6 (3.5)	71 (43) 7.5 (6.3)	83 (40) 9.7 (6.8)	62 (28) 6.7 (4.9)	43 (35) 5.1 (5.2)	-19 -2.1
Cholesterol mM/l	6.4 (1.1)*	6.0 (1.0)*	6.5 (1.8)*	6.3 (0.7)*	6.5 (0.7)*	6.2 (0.6)*	-0.3
LDL mM/l	4.0 (1.4)	3.2 (0.6)	4.0 (2.3)	3.7 (1.0)	4.1 (0.9)	3.7 (0.7)	-0.5
Apolipoprotein A1 mg/dl	155 (22)	163 (29)	130 (45)	171 (52)	151 (18)	141 (25)	+10
Apolipoprotein B mg/dl	110 (33)	91 (18)	113 (40)	111 (28)	112 (20)	115 (27)	-9
Fructosamine mM/l	282 (40)	296 (16)	277 (36)	338 (59)	313 (62)	312 (37)	+8
OGTT							
Glucose mM/l							
Fasting	4.7 (1.2)	5.3 (0.9)	5.1 (1.1)	5.2 (0.3)	4.5 (0.5)	4.8 (0.3)	+0.3

integrated	738 (163)	784 (165)	691 (79)	943 (249)	846 (143)	854 (168)	+55
Insulin μU/mL							
Fasting	20 (14)	38 (12)	22 (14)	12 (5)	19 (15)	17 (8)	+7
integrated	2481 (1006)	4582 (3042)	3648 (1643)	2319 (1019)	2349 (444)	4267 (1092)	+1532 (p<0.05 GH vs. placebo)

Comments
*p<0.05 GH vs. placebo
HVSDS is for chronological age; SDSBA =SDS for bone age; OGTT=oral glucose tolerance test
ERPF is effective renal plasma flow
For HV, there was no significant carry-over effect (0.5cm/6 months, p=0.30). Period effect was 0.9cm/6 months (p=0.06).

Cholesterol and other outcomes above were compared against controls. Not data extracted as not part of randomised study.

Adverse Effects
None of the patients had an acute rejection episode during the study.
No serious AE

Methodological comments
Allocation to treatment groups: states randomly and blindly assigned to groups, but no further details given.
Blinding: no details provided
Comparability of treatment groups: similar at baseline (although bone age 2 years higher in group 1)
Method of data analysis: Paper states that statistical methods appropriate for cross-over trials were used. Ref cited, but no further details given. Treatment effects were calculated and tested after taking into account any period effect. ANOVA used to test influence of baseline variables. Correlations were tested by Spearman non-parametric test. ITT analysis performed.
Sample size/power calculation: not stated
Attrition/drop-out: all children completed the study

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Adequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Adequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: AT, LB		Date: 29/09/08	Version: final
Reference and Design	Intervention	Participants	Outcome measures
Powell et al., 1997 [106] Country USA Study design: multicentre, open label RCT Number of centres: 26 Funding: Genentech Inc; government grants	1. 0.05 mg/kg/day s.c. RhGH 2. no treatment Duration of treatment: 1 year Other interventions used:	Target population: prepubertal children with chronic renal failure Number of Participants: Total: 69 entered, 44 analysed 1. n=30 2. n=14 Sample attrition/dropout: 20 left (12 ESRF; 6 entered puberty; 1 allergic to rhGH; 1 drowned); 4 grp1 and 1 grp2 completed study but were excluded as they had insufficient serum for the 0 and 12 month protein assays Inclusion criteria for study entry: Irreversible renal insufficiency (GFR > 10 and < 40 ml/min/1.73 m2 Height < 5th percentile for age Age > 2.5 years Ability to stand for height measurement Bone age < 10 for girls and 11 for boys Tanner stage I Exclusion criteria for study entry: Serum albumin <2.5 g/dl; receiving medications which influence growth; presence of illness affecting growth; diabetes mellitus; presence or past history of malignancy	Primary outcomes: not specified Secondary outcomes: height gain; HtSDS; Bone age; Mid-arm muscle circumference (MAMC); Triceps skinfold thickness (TSF); Weight gain; various IGF measures; insulin; ALS;GHBP Method of assessing outcomes: anthropometric measurements taken at 0, 3 and 12 months; height measured using wall-mounted stadiometer; bone age determined by a left hand and wrist radiograph at 0 and 12 months Length of follow-up: 1 year

Characteristics of participants:

Mean ± SD	0.05 mg/kg/day RhGH (n=30)	No treatment (n=14)	P Value
Sex	83% male	86% male	
GFR ml/min/1.73m2	27.5 ± 8.9	27.6 ± 8.8	
Age, years	5.6 ± 2.0	5.7 ± 2.6	
Bone age, years	4.0 ± 1.5 (n=27)	4.2 ± 1.8	

Height SDS	-2.7 ± 0.7	-2.7 ± 0.8	
Weight for HtSDS	0.0 ± 1.3	-0.2 ± 1.5	
MAMC cm	14.1 ± 1.6 (n=29)	14.4 ± 2.8	
TSF mm	7.9 ± 3.2 (n=29)	8.5 ± 3.2	
IGF-I nM	15 ± 10	10 ± 5	
IGF-I SDS	-0.7 ± 1.3	-1.2 ± 1.0	
Free IGF-I pM	71 ± 41 (n=17)	141 ± 94 (n=9)	P=0.029
IGF-II nM	100 ± 29	101 ± 41	
IGF-II SDS	1.2 ± 1.2	1.1 ± 1.3	
Insulin pM b	19 ± 14	52 ± 66	P=0.021
Total IGF nM	115 ± 34	111 ± 45	
IGFBP-1 nM	18 ± 9	17 ± 21	
IGFBP-1 SDS	2.4 ± 0.6	2.1 ± 1.4	
IGFBP-2 nM a	50 ± 17	51 ± 26	
IGFBP-3 nM b	130 ± 50	109 ± 25	
IGFBP-3 SDS c	1.7 ± 2.0	0.7 ± 1.1	
ALS nM	207 ± 81	179 ± 40	
GHBP pM	183 ± 104	144 ± 104 (n=12)	
GHBP SDS	0.4 ± 1.7	0.0 ± 1.3 (n=12)	
Comments MAMC=mid-arm muscle circumference; TSF=tricep skinfold thickness a Values > normal range (22 ± 11), p<0.001 b Values not different from normal range (98 ± 17) c Values > normal range (-0.2 ± 0.7) p=0.013			
Results Mean ± SD change from 0-12 months	0.05 mg/kg/day RhGH (n=30)	No treatment (n=14)	P Value
Bone age, years	1.0 ± 0.3 (n=27)	0.9 ± 0.4 (n=13)	P=0.5282
Height gain (cm)	9.1±2.8	5.5±1.9	p < .0001
Weight gain (kg)	3.5 ± 1.5	2.2 ± 1.0 kg	p = 0.007
Height SDS	0.8 ± 0.5	0.0 ± 0.3	P<0.0001
Weight for HtSDS	0.4 ± 0.7	0.4 ± 0.5	P=0.8703
MAMC cm	1.2 ± 0.9 (n=29)	-0.2 ± 1.7 (n=13)	P=0.0015
TSF mm	-1.9 ± 2.5 (n=29)	0.9 ± 1.2 (n=13)	P=0.0003
IGF-I nM	No actual values presented – only small diagram which is hard to read accurately. Not data extracted (but could go back and do if required)		P<0.006
IGF-I SDS	0.2 ± 1.0	No change from baseline – no values reported	P<0.006
Free IGF-I pM	No actual values presented – only small diagram which is hard to read accurately. Not data extracted (but could go back and do if required)		P<0.0464
IGF-II nM			P<0.006
IGF-II SDS	2.1 ± 1.3	No change from baseline – no values reported	P<0.006
Insulin pM	No actual values presented – only small diagram which is hard		P<0.017

Total IGF nM	to read accurately. Not data extracted (but could go back and do if required)		P<0.011
IGFBP-1 nM			P<0.017
IGFBP-1 SDS			P<0.017
IGFBP-2 nM			n/s
IGFBP-3 nM			P<0.011
IGFBP-3 SDS	4.0 ± 3.2	No change from baseline – no values reported	P<0.011
ALS nM	No actual values presented – only small diagram which is hard to read accurately. Not data extracted (but could go back and do if required)		P<0.011
GHBP pM			n/s
GHBP SDS	No actual values presented – only small diagram which is hard to read accurately. Not data extracted (but could go back and do if required)		n/s
Comments 10 healthy children (80% male; mean age 7.4 ± 2.7 years) provided serum samples for control values for IGFBP-2 and IGFBP-3 measurements.			
Adverse Effects - Not reported			
Methodological comments Allocation to treatment groups: randomised 1:2, no information on method of randomisation. Groups balanced for age, gender, height, GFR at baseline and nature of primary renal disease Blinding: open label Comparability of treatment groups: Free IGF-I and insulin were statistically significantly higher in control group, otherwise groups were similar. 10 healthy children (80% male; mean age 7.4 ± 2.7 years) provided serum samples for control values for IGFBP-2 and IGFBP-3 measurements. Mean age for control children was approximately 2 years older than for the randomised children. Method of data analysis: not ITT. Data presented as mean ± SD but converted to log10 values for statistical analysis. ANCOVA used to test differences between groups; $p \leq 0.05$ considered significant. Multiple regression analysis used to analyse effect of multiple variables on change in HtSDS, but not data extracted here. Sample size/power calculation: not reported, and primary outcome not clearly defined. Attrition/drop-out: 20 left (12 ESRF; 6 entered puberty; 1 allergic to rhGH; 1 drowned); 4 grp1 and 1 grp2 completed study but were excluded as they had insufficient serum for the 0 and 12 month protein assays			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	adequate

SGA Data extraction forms

Reviewers: LB, AT		Date: 10/11	Version: Checked
Reference and Design	Intervention	Participants	Outcome measures
Lagrou et al., 2008 [110] Country: Belgium and Luxembourg Study design: RCT Number of centres: 11 Funding: Belgian Study Group for Paediatric Endocrinology/ GH provided by Pfizer	1. GH 0.066mg/kg·day 2.Untreated (did not receive placebo injections) Duration of treatment: 2 years Other interventions used: None stated	Target population: Prepubertal children born small for gestational age Number of Participants: Total: 40 1.20 2. 20 Sample attrition/dropout:1 treated patient dropped out due to family problems Inclusion/exclusion criteria for study entry: Inclusion criteria: Birth weight and or length below -2 SD for gestational age Chronological age between 3 and 8 years Current height below -2.5 SD Height velocity SDS below +1.0 SD during the last 6-18 months Exclusion criteria: Gestational age <34 weeks Endocrine disease including GH deficiency, severe chronic disease, Turner, Noonan or Down Syndrome or other genetically confirmed syndromes Chromosomal abnormalities Bone disease Current or previous irradiation therapy Current or previous (up to 18 months before inclusion) treatment with glucocorticoids Severe mental retardation (IQ ≤ 50)	Primary outcomes: Height velocity Secondary outcomes: Height SDS, WtSDS, BMI SDS, Head circumference SDS, perception of short stature (not data extracted), perception of changes in height and physical appearance (not data extracted), perceptions of changes in psychosocial functioning (not data extracted) Method of assessing outcomes: Standard auxological assessment of height, weight and head circumference measurements every 6 months, calculated using British references. Psychological assessments performed at start of study and after 2 years of follow up (not data extracted). Length of follow-up: 2 years

Characteristics of participants:			
	GH 0.066mg/kg·day (n=20)	Untreated (n=20)	P Value

Birth WtSDS	-2.7 ± 0.9	-2.6 ± 0.8	ns
Gestational age	37.3 ± 2.1	38.2 ± 1.6	ns
Age, years	5.5 ± 1.6	5.1 ± 1.3	ns
HtSDS	-3.3 ± 0.6	-3.2 ± 0.9	ns
WtSDS	-3.8 ± 1.3	-3.9 ± 1.4	ns
BMI (SDS)	-1.7 ± 1.1	-2.0 ± 1.5	ns
Head circumference (SDS)	-2.7 ± 1.4	-2.8 ± 1.6	ns
Results			
Outcomes mean ± SD	GH 0.066mg/kg·day (n=20)	Untreated (n=19)	P Value
HtSDS	-1.9 ± 0.7	-3.1 ± 0.9	<0.001
WtSDS	-2.3 ± 1.2	-3.7 ± 1.5	<0.01
BMI (SDS)	-1.5 ± 1.1	-2.0 ± 1.5	ns
Head circumference (SDS)	-2.0 ± 1.4	-2.8 ± 1.5	<0.05
Adverse Effects: Tolerance only discussed in terms of perceptions of the injection by parents and children. No AE reported or discussed.			
Methodological comments Allocation to treatment groups: States randomised taking into account: gender, chronological age, WtSDS and study centre, no further details Blinding: No details given, untreated participants not given placebo injections Comparability of treatment groups: Authors report no differences in the auxological parameters between groups at baseline Method of data analysis: Differences of continuous variables between subgroups were evaluated by Students unpaired t test or by the Mann-Whitney U test as appropriate. The level of significance of difference was set at $p<0.05$ Sample size/power calculation: Based on 0.8 power to detect a significant difference (p=0.05) 20 subjects in each group were required assuming a difference of 2 cm/year in GV and a standard deviation of 2.2 cm/ year Attrition/drop-out: 1 treated patient dropped out due to family problems. Data for untreated group is for 19 after 2 years, no explanation of this			

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Inadequate

Reviewers: LB, AT		Date:10/11	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
Carel et al., 2003 [111] Country: France Study design: RCT Number of centres: Not stated Funding: Sanofi – Synthélabo	1. Daily GH injections: 0.2 IU/kg·d (0.067 mg/kg·d) 2. No treatment Duration of treatment: Until reached adult height (AH). The mean duration of treatment was 2.7 ± 0.6 yr. Other interventions used: None stated	Target population: Children born small for gestational age Number of Participants: Total:168 1.112 2. 56 Sample attrition/dropout: For treatment: Group1: n=21 Group 2: n=23 For analysis: Group 1: n=10 Group 2: n=9 Inclusion/exclusion criteria for study entry: Inclusion criteria: Birth length < – 2 SDS for gestational age (GA) and term > 30 weeks At study inclusion, height ≤ -2.5 for age or less Chronological age (C.A.) >10.5 yr for girls and >12.5 yr for boys Bone age (B.A.) ≥9 yr for girls and ≥10 yr for boys Peak plasma GH concentration after pharmacological stimulation at least 10µg/l to exclude GH deficiency Tanner stage I or II with testicular volume <8ml or uterus length <50mm Exclusion criteria: Chromosomal abnormalities in girls Constitutional bone diseases, any chronic disease interfering with growth Steroid or sex steroid treatment Dysmorphic syndromes other than Russell-Silver No catch-up growth criteria	Primary outcomes: Adult height (AH) SDS Secondary outcomes: Gain in SD units between height at inclusion and adult height Method of assessing outcomes: Follow up visits were every 3 months for the treated group, and every 6 months for the control group and the following data recorded: height, weight, chronological age, pubertal stage, dose and tolerance. BA analysed yearly. Length of follow-up: Criteria for stopping treatment/ follow up were <1 cm growth over the last 6 months, and a bone age of ≥15y for girls, and ≥16y for boys. Only 4% of patients met this criteria when treatment was stopped, so authors considered treatments to be almost complete for analytical purposes if growth velocity was 2 cm or less over the last 6 months, or bone age was ≥13y for girls, and ≥15y for boys Patients who had discontinued follow up before reaching AH were contacted later for a final AH measurement. Those who had not reached AH

		were specified	were maintained in the analysis without correction

Characteristics of participants:			
	Daily GH injections: 0.2 IU/kg·d (0.067 mg/kg·d) (n=102)	Untreated (n=47)	P Value
Target height	-1.2 ± 0.9	-0.9 ± 1.0	
Duration of pregnancy (wk)	39 ± 2	39 ± 2	
Birth length (SDS)	-2.8 ± 0.8	-3.1 ± 1.0	<0.05
Birth WtSDS	-1.8 ± 0.8	-1.9 ± 0.8	
Age (yr)	12.7 ± 1.4	12.8 ± 1.6	
Height (cm)			
HtSDS	-3.2 ± 0.7	-3.2 ± 0.6	
WtSDS	-1.9 ± 0.7	-2.2 ± 0.6	
Growth velocity (cm/yr)			
Bone age (yr)	10.6 ± 1.4	10.8 ± 1.6	
Pubertal (Tanner stage II)	22%	21%	
Comments: 4 patients had Russell – Silver syndrome. Growth velocity and height (cm) were not detailed for the groups as a whole, but for boys and girls within the group separately.			
Results			
Outcomes	Daily GH injections: 0.2 IU/kg·d (0.067 mg/kg·d) (n=91)	Untreated) (n=33)	P Value
At inclusion: age (yr)	12.6 ± 1.5	12.9 ± 1.4	
At inclusion: Height SDS	-3.2 ± 0.6	-3.2 ± 0.6	
At inclusion: Height (cm)	Not reported for whole group	Not reported for whole group	
At end of treatment: age (yr)	15.7 ± 1.5	Not reported	
At end of treatment: Height SDS	-2.1 ± 1.0	Not reported	
At end of treatment: Height (cm)	Not reported for whole group	Not reported	
At AH measurement: age (yr)	Not reported	Not reported	
At AH measurement: Height SDS	-2.1 ± 1.0	-2.7 ± 1.0	0.005
At AH measurement: Height (cm)	Not reported for whole group	Not reported for whole group	
At AH measurement: Total height gain (cm)	26 ± 7	22 ± 6	0.005
At AH measurement: Total height gain (SDS)	1.1 ± 0.9	0.5 ± 0.8	0.002
At AH measurement: Difference from target HtSDS	-0.9 ± 1.2	-1.7 ± 1.2	0.005
Comments: A difference of 0.6 SDS was observed in final height between the control and treated groups (95%CI 0.2-0.9) (A difference of 0.4 was observed at baseline, unclear if this is accounted for in finding the 0.6 result significant). The measurements above that have not been reported for the whole group are reported in the paper separately for boys and girls.			
Adverse Effects: 44% of treated patients reported AE, 10% having 4 or more events. The most			

frequently reported events involved the respiratory system (19%), osteomuscular system (14%), central nervous system (9%), and digestive tract (8%). Authors state that all of these were mild, reversible, benign conditions unlikely to be related to GH treatment. 16 AE recorded in 14 treated patients were considered severe: trauma, psychiatric symptoms, abdominal symptoms, otitis, asthma, varicocele, striae, and migraine. Again, authors state these are unlikely to be related to GH treatment. 2 were causally related to treatment: 1 slipped capital ephiphysis after 1.5 yrs of treatment and had one single seizure episode 10 mins after 1st injection.
Methodological comments Allocation to treatment groups: Allocation sequence generated centrally and faxed to participants Blinding: Group assignment was not masked, and the treated group was twice as large as the control group Comparability of treatment groups: There is a significant difference in birth length between the treated and untreated group, with the treated group being longer than the untreated group (P 0.04). On other characteristics the groups appear to be broadly similar. Method of data analysis: Means and SD values are presented. Mann-Whitney U test to compare groups. An α risk of 5% was set as the significance threshold. Not ITT. Sample size/power calculation: Not reported. Attrition/drop-out: Four patients in the treatment group were excluded from analysis due to severe diseases interfering with growth, (sickle cell anemia, pulmonary hypertension, type 1 neurofibromatosis and severe prematurity). Five patients assigned to the treatment group refused GH treatment but remained in the study and were analysed as part of the control group. Fifteen patients left the study early (14 in control and 1 in the treated group). Treatment was completed in 4/102 patients and almost complete in 64/102. The reasons for interrupting treatment early were: growth rates considered insufficient by patient/physician (n=12), weariness with the treatment (n= 10), loss to follow up (n=5), satisfaction with height (n=2), local intolerance (n=1), and striae attributed to the treatment by the patient (n=1). In addition some of the investigators wrongly considered that the treatment duration was limited to 3 yr and stopped the treatment early (n=unclear). 102 treated and 47 control patients are included in the analysis. Authors state that group reassignments or protocol deviations concerned 12 and 5 patients followed to AH in the treated and control groups respectively. Appear to have been significant problems with attrition for various reasons, appears to be fully described. Group assignment was not blinded, and despite the study being randomised and centrally allocated, the treatment group is twice as large as the control: either this was 2:1 randomisation (this is not reported), or large numbers of the control group dropped out after randomisation, or possibly swapped to the treatment group: this is unclear.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Inadequate
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Inadequate
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: LB, AT		Date: 10/11	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
De Schepper et al., 2007 [109] Country: Belgium Study design: RCT Number of centres: 8 Funding: Belgian Study Group for Paediatric Endocrinology/ Pfizer	1. High dose growth hormone (GH): 66 ± 3 µg/kg s.c. once daily. Adjusted every 6 months to body weight 2. Untreated (did not receive placebo injections) Duration of treatment: 2 years Other interventions used: None stated	Target population: Children born short for gestational age Number of Participants: Total: 40 (25) 1.11 2. 14 Sample attrition/dropout: The trial cohort was reduced from 40 to 25 based on the availability of the same absorptiometry apparatus to assess body composition in a homogenous fashion across 8 centres. No anthropometric differences were detectable between the study population and the non-included sub cohort (authors state, no data reported) Inclusion/exclusion criteria for study entry: Inclusion criteria: Birth weight, length or both <-2 SD for gestational age (G.A.) Current height <-2.5 SD Height velocity <+1 SD in the last 6-18 months Age between 3 and 8 yrs at study start Exclusion criteria: Premature birth (G.A <34 wks) Evidence for endocrine or bone disease Severe chronic disease Turner, Noonan, Down or other genetic syndrome Irradiation treatment Current or previous glucocorticoid treatment Severe cognitive dysfunction (est. I.Q. <50)	Primary outcomes: None clearly stated Secondary outcomes: Height and WtSDS, anthropometric and absorptiometric characteristics Method of assessing outcomes: Study participants seen every 3 months, height measured with Harpenden stadiometer, and weight with electronic scale. Mid upper arm circumference and four skin folds were measured at study start and after 1 and 2 years. Length of follow-up:2 years

Characteristics of participants:			
	High dose growth hormone (GH) (n=11)*	Untreated (n=14)	P Value

Age, years	5.1 ± 1.6	5.1 ± 1.4	
G.A (weeks)	37 ± 3	38 ± 2	
Birth WtSDS	-2.4 ± 0.8	-2.5 ± 0.8	
Birth length (SDS)	-3.1 ± 0.6	-2.9 ± 0.7	
Mid-parental height**	-0.9 ± 0.8	-0.8 ± 0.7	
Height SDS	-3.3 ± 0.7	-3.2 ± 1	
WtSDS	-3.5 ± 1.2	-3.6 ± 1.5	
Subscapular skinfold (mm)	5.4 ± 1.1	6.4 ± 2.1	
Triceps skinfold (mm)	7.9 ± 1.4	8.3 ± 2.1	
Subscapular/Triceps	0.7 ± 0.2	0.8 ± 0.2	
Sum skinfolds (mm)	22.1 ± 3	24.3 ± 6	
Body fat fraction (%)	12.9 ± 2.1	14.1 ± 3.6	
MUAMA (cm)	12.8 ± 2.5	14.1 ± 3.5	
MUAFA (cm)	5.5 ± 1.1	5.7 ± 1.7	
Lean mass (kg)	10 ± 3	9.9 ± 2.2	
Fat mass (kg)	2.3 ± 0.5	2.5 ± 0.9	
Lean mass (%)	78 ± 4	77 ± 5	
Fat mass (%)	15 ± 3	20 ± 5	
Trunk fat (kg)	0.7 ± 0.3	0.8 ± 0.4	
Limb fat (kg)	1.1 ± 0.3	1.2 ± 0.5	
Trunk fat/Limb fat	0.6 ± 0.2	0.6 ± 0.2	
Trunk fat/Leg fat	0.8 ± 0.3	0.8 ± 0.3	

* Not significant for baseline comparisons between groups **[father's HtSDS + mother's HtSDS] divided by 2

Results

Outcomes	High dose growth hormone (GH) (n=11)*		Untreated (n=14)		P Value*
	1 year	2 years	1 year	2 years	
Height SDS	-2.1 ± 0.7[a]	-1.7 ± 0.7[a, d]	-3.1 ± 1[b]	-3 ± 1[b]	<0.0001
WtSDS	-2.4 ± 1.3[a]	-1.8 ± 1[a, d]	-3.5 ± 1.4	-3.4 ± 1.6[b]	<0.0001
Subscapular skinfold (mm)	4.7 ± 0.8[b]	5.1 ± 1	5.7 ± 1.8[b]	6 ± 2.1	ns
Triceps skinfold (mm)	4.9 ± 1.5[a]	5.5 ± 2.1[a]	8.2 ± 2.3	7.9 ± 2.4	<0.001
Subscapular/Triceps	1 ± 0.3[c]	1 ± 0.3[a, e]	0.7 ± 0.2	0.8 ± 0.2[h]	0.001
Sum skinfolds (mm)	16.6 ± 3.4[a]	18.1 ± 5[c]	22.4 ± 5.8[b]	22.9 ± 6.8	<0.005
Body fat fraction (%)	9.1 ± 2.1[a]	10.1 ± 3[c]	13.3 ± 3.5	13.4 ± 3.5	<0.005
MUAMA (cm)	15.2 ± 2.9[a]	17 ± 2.7[a, f]	13.3 ± 2.3[c]	14.1 ± 2.9[a, g]	<0.005
MUAFA (cm)	3.6 ± 1.2[a]	4.3 ± 1.9[c, g]	5.8 ± 2	5.7 ± 1.9	0.001
Lean mass (kg)	13.2 ± 3.4[a]	15.5 ± 3.4[a, d]	10.9 ± 2.4[a]	12.2 ± 2.5[a, d]	<0.0001
Fat mass (kg)	2.4 ± 0.7	2.9 ± 1[b, f]	2.8 ± 1.1[b]	3.1 ± 1.1[a, g]	ns
Lean mass (%)	82 ± 3[c]	82 ± 3[b]	77 ± 6	77 ± 5	<0.05
Fat mass (%)	15 ± 3[c]	15 ± 2[b]	20 ± 6	20 ± 5	<0.05
Trunk fat (kg)	0.9 ± 0.3	1 ± 0.3[b]	0.9 ± 0.5	1 ± 0.6[a]	ns
Limb fat (kg)	0.9 ± 0.5	1.3 ± 0.7[f]	1.4 ± 0.6[b]	1.5 ± 0.6	<.05
Trunk fat/Limb fat	1 ± 0.5[c]	0.9 ± 0.3[c, e]	0.6 ± 0.2	0.7 ± 0.2[h]	<0.0001
Trunk fat/Leg fat	1.5 ± 0.7[c]	1.3 ± 0.4[c, e]	0.8 ± 0.3	0.9 ± 0.3[h]	<0.0001

Comments* difference between untreated and treated group (analysis of variance) unclear if this is totals over the 2 years of the study, including baseline measurements. [a] P<0.0005 paired t test or Wilcoxon rank test@ baseline – year 1, baseline – year 2 [b] P<0.05 [c] P< 0.005 [d] P<0.0005 paired t test or Wilcoxon rank test: year 1 – year 2 [e] Elevated for age [f] P<0.005 [g] P<0.5 [h] Normal for age MUAMA: mid upper arm muscle area; MUAFA: mid upper arm fat area GH treatment was accompanied by a gain of lean mass (P<0.0001) and by a centripetal redistribution of fat mass (P<0.0001) but not by an overall gain or loss of fat mass. The effects of high dose GH on adiposity are not readily detectable in the trunk and are essentially limited to the limbs.
Adverse Effects: Authors state that 'none had a noteworthy adverse event during the 2 years of study'
Methodological comments Allocation to treatment groups: States randomised, no information reported on allocation to groups. Original trial cohort was 40, this was reduced to 25 due to availability of equipment. Blinding: No information on blinding reported, untreated group did not receive placebo injections Comparability of treatment groups: Groups appear comparable at baseline: authors state there were no detectable baseline differences in the subgroups Method of data analysis: Results are expressed as mean ± SD. Repeated measures analysis of variance was used to test for differences between sub groups. The level of statistical significance was set at p<0.05 Sample size/power calculation: None reported Attrition/drop-out: 15 children from the original cohort were withdrawn due to issues with availability of measuring equipment – unclear at what stage this happened. No drop-outs are reported from the 25 included in the study, apart from this.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: LB, AT		Date: 12/11	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
de Zegher et al., 2002 [113] Country: UK and Belgium Study design:	1. High dose GH (100 µg/kg/d) 2. No treatment Duration of treatment: 2 years	Target population: Short children born small for gestational age Number of Participants: Total: 13 1.9	Primary outcomes: Not clearly stated Secondary outcomes: Growth response and its relationship to pre-treatment GH secretion

RCT Number of centres: 2 Funding: Pharmacia Ltd	Other interventions used: None stated	2.4 Sample attrition/dropout: Not reported Inclusion/exclusion criteria for study entry: Inclusion criteria: Birth weight/length <-2 SD for gestational age Current height <-3.0 SD Height velocity below 0.0 SD Age between 2 and 8 year Exclusion criteria: Identified syndrome other than Silver-Russell	(not data extracted) HtSDS, WtSDS, BMI SDS, GV (cm/yr) Method of assessing outcomes: Overnight GH profiles and GH stimulation tests at baseline (not data extracted), intravenous glucose tolerance tests were performed at baseline, yearly on GH treatment and 3 months post-GH treatment. Height, weights and body mass index converted to SDS using current UK reference Length of follow-up: 2 years

Characteristics of participants:

	High dose GH (100 µg/kg/d) (n=9)	No treatment (n=4)	P Value
Age (yr)	6.3 (4.0-8.0)	4.7 (2.3 -6.3)	
Height SDS	-3.6 (-5.5 - -2.8)	-3.1 (-3.4 – -2.8)	
WtSDS	-4.5 (-7.2 – -2.6)	-3.8 (-5.5 – -2.7)	
BMI SDS	-2.3 (-5.0 - -0.7)	-2.0 (-4.2 - -0.1)	
Height velocity (cm/yr)	5.1 (4.0 – 6.8)	6.4 (5.3 – 7.5)	

Results are presented as means and ranges

Results

Outcomes	High dose GH (100 µg/kg/d) (n=9)	No treatment (n=4)	P Value
Age (yr) (year 1)	7.2 (5.0 to 8.8)	5.7 (3.3 to 7.3)	
Age (yr) (year 2)	8.2 (6.0 to 9.9)	6.5 (4.3 to 8.3)	
Height SDS (year 1)	-2.4 (-4.6 to -1.4)*	-3.0 (-3.3 to -2.7)	
Height SDS (year 2)	-1.8 (-3.9 to -0.5)*	-3.0 (-3.3 to -2.5)	
WtSDS (year 1)	-2.9 (-4.7 to -1.7)*	-4.0 (-5.4 to -3.2)	
WtSDS (year 2)	-2.1 (-3.6 to -0.9)*	-3.8 (-4.8 to -3.2)	
BMI SDS (year 1)	-1.6 (-3.8 to -0.8)*	-2.3 (-3.9 to -1.3)	
BMI SDS (year 2)	-1.2 (-3.4 to -0.4)*	-2.1 (-2.9 to -1.4)	
Height velocity (cm/yr) (year 1)	11.0 (7.4 to 13.3)	Not reported	
Height velocity (cm/yr) (year 2)	8.5 (6.3 to 10.2)	5.6 (4.4 to 6.8)	

Comments: Authors state that GH treated children showed significant increments in HtSDS, WtSDS and BMI SDS over 2 yr (all P<0.0001). Untreated SGA children remained on their height, weight and BMI SD levels. * P<0.0001 from baseline

Glucose and insulin metabolism markers not data extracted as reported for the treated group, no results reported for controls. Authors state that compared to baseline levels, children in the treated group showed significant increases in fasting levels of insulin (year 1 P=0.003, year 2 P=0.0002) and decreases in insulin sensitivity (year 1 P=0.003, year 2 P=0.0002).
Adverse Effects: Not reported/discussed. No child showed impaired glucose tolerance.
Methodological comments Allocation to treatment groups: Randomised on a 2:1 basis, no further details Blinding: No details given. No placebo used. Comparability of treatment groups: Groups appear similar Method of data analysis: Means and ranges are presented. Changes in height/ weight, glucose and insulin parameters analysed using paired t tests. ITT Sample size/power calculation: Not reported Attrition/drop-out: Not reported

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Unknown
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Adequate
10. Were withdrawals and dropouts completely described?	Adequate

Reviewers: LB, AT		Date:11/11	Version: Checked
Reference and Design	Intervention	Participants	Outcome measures
de Zegher et al., 1996 [112] Country: Belgium Study design: Open-labelled RCT Number of centres: Multi Funding: Support from Pharmacia Peptide	1. Growth hormone (GH) 0.2 IU/kg/day s.c. 2. GH 0.3 IU/kg/day s.c. 3. Untreated Duration of treatment: 2 years Other interventions used: None stated	Target population: Children born small for gestational age Number of Participants: Total:54 1.20 2.21 3.13 Sample attrition/dropout: Group 1: n=2 Group 2: n=1 Group 3: n=1 Inclusion/exclusion criteria for study entry: Inclusion criteria: Birth weight/length < -2 SD	Primary outcomes: Not clearly stated Secondary outcomes: Height, HtSDS, GV, GV SDS, WtSDS, weight gain, BMI and BMI SDS, serum IGF-I, IGF-II, IGFBP-3, osteocalcin Method of assessing outcomes: Study visits including history, auxological evaluation, bone age determination, and dose adjustment were scheduled every 6 months. Biochemical examinations

Hormones		for gestational age Height SDS for age < -2.5 Height velocity SDS for age <+1 Chronological age between 2 and 8 Serum GH concentration >10µg/ L after exercise, glucagon or insulin tolerance test Available growth data concerning the period preceding the start of the study Exclusion criteria: Endocrine disorders Turner or Downs syndromes Previous or concomitant irradiation or anabolic steroid therapy Severe chronic disease Severe mental retardation	were performed yearly All bone ages were read according to Tanner-Whitehouse II method Height SDS for bone age was used as an index of final height prognosis Length of follow-up:2 years

Characteristics of participants:

	GH 0.2 IU/kg/day (n=20)	GH 0.3 IU/kg/day (n=19)	Untreated (n=13)	P Value
Birthweight (g)	2082.0 ± 139.0	1842.0 ± 115.0	1996.0 ± 136.0	ns
Birthlength (cm)	42.3 ± 1.1	42.5 ± 0.9	42.1 ± 1.1	ns
Chronological age (yr)	5.4 ± 0.5	5.1 ± 0.4	4.9 ± 0.5	ns
Bone age (yr)	4.5 ± 0.5	3.7 ± 0.5	3.7 ± 0.5	ns
Height SDS	-3.5 ± 0.2	-3.7 ± 0.2	-3.4 ± 0.3	ns
Height velocity (cm/yr)	6.6 ± 0.4	7.0 ± 0.5	6.7 ± 0.7	ns
Height velocity SDS	-0.9 ± 0.2	-0.7 ± 0.3	-0.6 ± 0.3	ns
Weight (kg)	13.2 ± 0.9	12.3 ± 0.7	12.0 ± 0.8	ns
WtSDS	-2.5 ± 0.2	-2.9 ± 0.2	-2.8 ± 0.2	ns
BMI	14.0 ± 0.4	13.8 ± 0.4	13.5 ± 0.4	ns
BMI SDS	-1.8 ± 0.4	-1.8 ± 0.3	-2.0 ± 0.4	ns
Serum IGF-I (µg/L)	107.0 ± 15.0	108.0 ± 14.0	108.0 ± 21.0	ns
Serum IGF-II (µg/L)	557.0 ± 44.0	748.0 ± 60.0	699.0 ±103.0	ns
Serum IGFBP-3 (mg/L)	3.34 ± 0.33	3.36 ± 0.38	3.35 ± 0.38	ns
Serum osteocalcin (µg/L)	69.0 ± 3.0	69.0 ± 2.0	63.0 ± 3.0	ns

Results are mean ± SEM. The 52 participating children were considered to have no specific syndrome (n=33), Silver-Russell syndrome (n=10), Fetal Alcohol syndrome (n=4), Dubowitz syndrome (n=3), 4p-syndrome (n=1) or Lacrimo-auriculo-dento-digital syndrome (n=1).

Results

Outcomes at 2 years,	GH 0.2 IU/kg/day	GH 0.3	Untreated	P Value

unless otherwise stated	(n=20)	IU/kg/day (n=19)	(n=13)	
Gain in bone age (yr)	1.35 ± 0.16	1.33 ± 0.24	0.84 ± 0.07	<0.001 treated vs. untreated
Height velocity (cm/yr) (Year 1)	11.5 ± 0.4	12.0 ± 0.4	Not reported	
Height velocity (cm/yr)	10.2 ± 0.2	11.0 ± 0.4	5.7 ± 0.3	<0.001untreated vs. treated; <0.05 group 1 vs. group 2
Height velocity SDS (Year 1)	5.3 ± 0.3	5.8 ± 0.4.	Not reported	
Height velocity SDS	4.3 ± 0.3	5.2 ± 0.4	-0.9 ± 0.3	<0.001 untreated vs. treated
Gain in HtSDS	2.1 ± 0.1	2.5 ± 0.1	0.2 ± 0.1	<0.001 untreated vs. treated
Gain in HtSDS for bone age	1.0 ± 0.2	1.2 ± 0.4	0.0 ± 0.3	<0.05 untreated vs. treated
Weight gain (kg)	6.9 ± 0.6	7.8 ± 0.5	3.6 ± 0.4	<0.001 untreated vs. treated
Gain in WtSDS	1.3 ± 0.1	1.8 ± 0.1	0.4 ± 0.1	<0.001 untreated vs. group 1; <0.01 group 1 vs. group 2
Serum IGF-I (µg/L) (Year 1)	274 ± 30	392 ± 43	145 ± 23	<0.01 group 1 vs. untreated; <0.05 group 1 vs. group 2
Serum IGF-I (µg/L)	332 ± 29	655 ± 69	168 ± 46	<0.0001 group 1 vs. group2; <0.01 group 1 vs. untreated
Serum IGF-II (µg/L) (Year 1)	745 ± 72	944 ± 101	756 ± 108	
Serum IGF-II (µg/L)	834 ± 53	966 ± 56	881 ± 125	ns
Serum IGFBP-3 (mg/L) (Year 1)	5.37 ± 0.42	6.35 ± 0.44	3.88 ± 0.48	
Serum IGFBP-3 (mg/L)	6.10 ± 0.35	6.50 ± 0.52	4.00 ± 0.58	ns group1 vs. 2; <0.001 untreated vs. group 1
Serum osteocalcin (µg/L) (Year 1)	89.4 ± 5.9	93.6 ± 9.9	59.9 ± 1.9	
Serum osteocalcin (µg/L)	100.0 ± 8.6	102.7 ± 9.8	72.5 ± 7.3	<0.05 untreated vs. group 1, ns

				group 1 vs. group 2
Comments: Results are mean ± SEM. Compliance: Over 2 years less than 10 injections were said to be missed in 36/38 children. In 2 children respectively, 3% and 8% of the injections were reportedly omitted. Children with and without specified syndromes appeared to present similar growth responses. The GV during the first year was higher than during the second year, both in group 1(11.5 ± 0.4 vs. 8.8 ± 0.2 cm/yr) and group 2 (12.0 ± 0.4 vs. 10.0 ± 0.3 cm/yr). After two years all untreated children still had a HtSDS <-2.2, whereas this was no longer the case for 35/38 treated children. BMI and BMI SDS remained similar in the three groups after 1 and 2 years. BMI of the study population is reported, not separately for the groups, or treated vs. untreated. Fasting serum insulin concentrations were twice as high (P=0.01) in treated children compared with untreated children both after 1 year (20.3 ± 2.2 mU/L vs. 10.6 ± 2.4 mU/L) and 2 years (18.9 ± 3.0 mU/L vs. 9.4 ± 1.3 mU/L) with no difference between the treated groups.				
Adverse Effects: Four serious AE, authors state conceivably not related to GH. One treated child received antibiotics for possible osteomyelitis of the distal tibia. 3 children hospitalised in relation to viral diseases: 1 untreated and 2 treated. Treatment was not interrupted. Cutaneous eczema was aggravated in one child in group 1, no treatment interruption. 3 treated children reported a possible increase in size or number of pigmented nevi, treatment was not interrupted. After 2 years, all haemoglobin AIC values were normal.				
Methodological comments Allocation to treatment groups: Stated to be weighted randomisation, no further details Blinding: Open label. Assessor for bone age blinded to chronological age and treatment randomisation Comparability of treatment groups: No significant differences at baseline Method of data analysis: Wilcoxon rank sum test used for differences between groups for growth variables, and Student's t test for biochemical markers. Statistically significant differences were considered to be obtained at P<0.05. Results are mean ± SEM. Not ITT Sample size/power calculation: None reported Attrition/drop-out: Two children allocated to 0.3 IU/kg did not start. Two children dropped out of the study for psychosocial reasons, one control after the start visit and one child from Group 1 after 19 months				

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	Adequate
5. Were outcome assessors blinded to the treatment allocation?	Partial
6. Was the care provider blinded?	Unknown
7. Was the patient blinded?	Inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	Adequate
9. Did the analyses include an ITT analysis?	Inadequate
10. Were withdrawals and dropouts completely described?	Adequate

Data extraction form for primary studies

Reviewers: AT,LB		Date: 8/07/09	Version: checked
Reference and Design	Intervention	Participants	Outcome measures

Phillip et al., 2009 [114] Country multinational Study design RCT Number of centres multicentre Funding: Novo Nordisk	1. rhGH 0.033 mg/kg/d 2. rhGH 0.1 mg/kg/d 3. untreated Duration of treatment: 2 years, but data only extracted for year 1 as control group received rhGH in year 2. Other interventions used: none	Target population: 3-8 year olds with persistent short stature born SGA Number of Participants: Total: n=151 1. n=51 2. n=51 3. n=47 Sample attrition/dropout: n=2 Inclusion/exclusion criteria for study entry: birth weight/length ≤ -2 SDS; HtSDS ≤ -2.5 SDS; GV SDS ≤ 0 during last 3 months; parental height ≥ -2 SDS; normal response to GH test.	Primary outcomes: measurement of height during 2 years. Treatment effect was additional height gain compared with untreated children Secondary outcomes: HtSDS; IGF-I; IGFBP-3; glucose; insulin Method of assessing outcomes: Harpenden stadiometer; sex-adjusted target height calculated based on national references; BA assessed by X-ray; HtSDS calculated using appropriate population references by country Length of follow-up: 1 year

Characteristics of participants:				
Mean ± SD	rhGH 0.033 mg/kg/d (n=51)	rhGH 0.1 mg/kg/d (n=51)	No treatment (n=47)	P Value
M:F, %	55:45	47:53	51:49	nr
Birth length, cm	44.3 ± 5.3	44.6 ± 4.3	43.9 ± 5.0	nr
Birth weight, kg	1.9 ± 0.6	2.0 ± 0.6	2.0 ± 0.6	nr
Gestational age, wks	36.9 ± 3.6	37.6 ± 3.3	37.5 ± 3.2	nr
Target HtSDS	-0.9 ± 0.6	-0.8 ± 0.6	-0.9 ± 0.8	nr
Height, cm	99.0 ± 9.3	98.9 ± 9.0	99.2 ± 7.9	nr
Height SDS	-3.1 ± 0.5	-3.2 ± 0.7	-3.1 ± 0.5	nr
Age, years	5.5 ± 1.5	5.5 ± 1.4	5.6 ± 1.4	nr
Bone age, years	4.7 ± 1.8	4.9 ± 1.8	5.0 ± 1.9	nr
Bone age: chronological age	0.8 ± 0.2	0.8 ± 0.2	0.8 ± 0.2	nr
IGF-I, ng/ml	116.7 ± 59.4	145.9 ± 92.3	130.0 ± 84.1	nr
IGFBP-3, µg/l	3.2 ± 0.9	3.5 ± 0.9	3.4 ± 1.1	nr
IGF-I SDS	-1.4 ± 0.6	-1.1 ± 0.9	-1.2 ± 1.0	nr
Fasting glucose, mmol/l	4.6 ± 0.6	4.7 ± 0.6	4.6 ± 0.4	nr
Fasting insulin µIU/ml	3.1 ± 2.8	2.7 ± 1.9	2.8 ± 3.3	nr
HbA_{1c}, %	5.2 ± 0.4	5.2 ± 0.3	5.1 ± 0.4	nr
Results at year 1				
Mean ± SD	rhGH 0.033 mg/kg/d (n=51)	rhGH 0.1 mg/kg/d (n=51)	No treatment (n=45)	P Value

Height SDS	-2.3 ± 0.6	-1.8 ± 0.8	-3.0 ± 0.6	nr
Change in HtSDS	0.8 ± 0.3	1.4 ± 0.4	0.1 ± 0.3	nr
Additional height gain (cm)	3.3 ± 0.2, 95% CI 2.9-3.7	6.5 ± 0.2 95% CI 6.0-6.9	n/a	nr
IGF-I, ng/ml	345.6 ± 177	594.3 ± 221	176.3 ± 107	nr
IGFBP-3, µg/l	4.8 ± 1.1	6.1 ± 1.4	3.9 ± 1.1	nr
IGF-I SDS	0.9 ± 1.9	3.3 ± 2.1	-0.9 ± 1.2	nr
Fasting glucose, mmol/l	4.8 ± 0.5	5.0 ± 0.5	4.8 ± 0.6	nr
Fasting insulin µIU/ml	5.3 ± 3.5	8.9 ± 5.0	4.1 ± 6.3	nr
HbA_{1c}, %	5.3 ± 0.4	5.3 ± 0.2	5.2 ± 0.4	nr
Adverse events Only reported for overall 2 year study, so treatment arms are different (no control arm). The majority (349/358, 73.5%) of AE were mild to moderate in severity, and the most common events (57%) were childhood infections. 16 serious AE were reported, 3 of which were likely to be related to rhGH (convulsions, epilepsy, papilloedema – all stabilised/resolved after rhGH discontinued).				
Methodological comments Allocation to treatment groups: randomised 1:1 to double-blind treatment in the two rhGH groups or to a control group that was untreated in the first year and received rhGH in the second. A computer-controlled, centralised system was used to assign treatment. Blinding: Bone age assessed centrally by clinicians blinded to subject's characteristics (other than gender) and treatment. Comparability of treatment groups: Similar at baseline, but no p vals reported. Method of data analysis: mixed effects model (ANCOVA) used where effects of age, sex and treatment duration were included. Tests were 2-side F tests, performed at the 5 % significance level. Sample size/power calculation: At least 50 patients per group were required to detect a difference in height gain of 0.75 cm between the two rhGH groups with a power of 90% and a significance level of 0.05. To allow for comparison with the 3rd group, and allowing for a dropout rate of 20%, 180 patients were required to be enrolled. Attrition/drop-out: 2 randomised patients missing from analysis. Reasons not given.				

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	adequate
2. Was the treatment allocation concealed?	unclear
3. Were the groups similar at baseline in terms of prognostic factors?	reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	partial
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	inadequate

SHOX-D Data extraction forms

Reviewers: AT, LB		Date: 8/10/08	Version: Final
Reference and Design	Intervention	Participants	Outcome measures
Blum et al., 2007 [48] Country international (14 countries) Study design RCT Number of centres 33 Funding: Eli Lilly and co.	1. daily s.c. injection of 50 µg GH 2. no treatment 3. daily s.c. injection of 50 µg GH Duration of treatment: 2 years Other interventions used:	Target population: prepubertal children with SHOX-D Number of Participants: Total: 68 patients had SHOX gene deletions or mutations, of which 52 with SHOX-D enrolled. A further 26 (grp 3) with TS were enrolled as an additional GH arm. 1. n=27 2. n=25 3. n=26 [not data extracted as not per protocol] Sample attrition/dropout: 1 Inclusion criteria for study entry: Confirmed SHOX-D; age ≥ 3 years; prepubertal (Tanner stage 1); height < 3rd percentile or <10th percentile with HV < 25th percentile; bone age <10yrs (boys) or <8 yrs (girls); <9 yrs (TS girls); no GH deficiency or resistance; no chronic disease; no growth-influencing medications.	Primary outcomes:1st year GV Secondary outcomes: comparison between treatment effects in SHOX-D and TS patients [not data extracted as not per protocol]; AE Method of assessing outcomes: height, IGF-I and IGFBP-3 measured at baseline, 3 mths, 6 mths, then at 6mth intervals for remainder of the 2 years; left hand and wrist x-rays for bone age performed at baseline, 1 yr and 2yr – assessed centrally using Greulich and Pyle method; glucose and routine blood analysis at baseline and 1st year. Length of follow-up: 2 years

Characteristics of participants:

Mean ± SD, unless otherwise stated	SHOX-D grp1 50 µg GH (n=27)	SHOX-D grp2 No treatment (n=25)	P Value gp 1 vs. gp2
Complete deletion of SHOX gene, n	18	16	
Partial gene deletions, n	2	2	
Point mutations, n	7	7	
Female/male (%)	52/48	56/44	
LWS/ISS phenotype (%)	56/40	44/56	0.689
Chronological age (yr)	7.5 ± 2.7	7.3 ± 2.1	0.914
Bone age (yr)	6.6 ± 2.8	6.5 ± 2.0	0.928
Bone age – chronological age	-1.0 ± 0.9	-0.8 ± 0.8	0.809
Bone age SDS	-1.2 ± 1.1	-1.0 ± 1.0	0.641
Height SDS	-3.3 ± 1.0	-3.3 ± 0.8	0.111

Target HtSDS	-1.3 ± 1.0	-1.5 ± 0.9	0.013
Body mass index SDS	0.2 ± 0.9	0.6 ± 0.9	0.147
IGF-I SDS	-0.8 ± 1.0	-0.9 ± 1.0	0.521
IGFBP-3 SDS	0.6 ± 1.3	0.1 ± 1.1	0.058
Results			
Mean ± SD, unless otherwise stated	SHOX-D grp1 50 µg GH	SHOX-D grp2 No treatment	P Value gp 1 vs. gp2
Baseline HV (cm/yr)	4.8 ± 0.3 (n=18)	5.0 ± 0.5 (n=14)	0.721
Baseline HV SDS	-1.2 ± 0.3 (n=12)	-1.0 ± 0.6 (n=10)	0.605
Baseline HtSDS	-3.3 ± 0.2 (n=27)	-3.2 ± 0.2 (n=24)	0.822
1st year HV (cm/yr)	8.7 ± 0.3 (n=27)	5.2 ± 0.2 (n=24)	<0.001
1st year HV SDS	3.0 ± 0.3 (n=25)	-0.7 ± 0.2 (n=22)	<0.001
1st year ht SDS	-2.6 ± 0.2 (n=27)	-3.1 ± 0.2 (n=24)	<0.001
2nd year HV (cm/yr)	7.3 ± 0.2 (n=27)	5.4 ± 0.2 (n=24)	<0.001
2nd year HV SDS	2.3 ± 0.3 (n=27)	-0.4 ± 0.1 (n=22)	<0.001
2nd year ht SDS	-2.1 ± 0.2 (n=27)	-3.0 ± 0.2 (n=24)	<0.001
2nd year ht gain (cm)	16.4 ± 0.4 (n=27)	10.5 ± 0.4 (n=24)	<0.001
Catch up of bone age	1.34 ± 0.07	1.1 ± 0.09	P=0.161
Adverse events			
	SHOX-D grp1 50 µg GH (n=27)	SHOX-D grp2 No treatment (n=25)	P Value gp 1 vs. gp2
At least 1 treatment-emergent AE (mostly common childhood illnesses)	85%	68%	
Arthralgia	3	2	
Gynecomastia (males)	1 (n=12 males)	0 (n=12 males)	
Increased number of cutaneous nevi	2	0	
Recurrent otitis media	1	1	
Scoliosis	1	0	
diabetes	0	0	
Comments 41% of GH treated SHOX-D patients reached a height within the normal range for age and gender (>-2.0SDS), compared with only 1 pt in the control group. For the GH treated SHOX-D patients, 1st year GV was somewhat greater for males (9.3 ± 0.5 cm/yr) than for females (8.4 ± 0.5 cm/yr), the baseline to second-year change in GV was very similar. Subgroup analysis for ISS phenotype vs. LWS phenotype presented but not data extracted as not per protocol. IGF-I SDS were in the low-normal range in each of the study groups at baseline and remained there for the untreated group. In the GH treated group, values increased to the upper-normal range. IGF-I concentrations exceeded +2 SDS at least once during GH treatment in 10 (37%) of pts and no untreated patients. IGFBP-3 SDS at baseline were closer to the normal mean that the corresponding IGF-I SDS in both study groups and increased to the upper-normal range in the treated group. There was a strong relationship between IGF-I SDS and IGFBP-3 SDS values during GH treatment, such that no subject had an IGF-I SDS in the upper tertile with an IGFBP-3 SDS in the lower tertile. No significant changes in thyroid function. No serious AE were reported for subjects with SHOX-D			
Methodological comments			

Allocation to treatment groups: After stratification by sex and according to presence or absence of LWS, patients were randomized on a 1:1 basis. No further details given Blinding: blood analyses were carried out in a central facility. Open label Comparability of treatment groups: similar at baseline Method of data analysis: Height SDS calculated using a central European reference Sample size/power calculation: Not reported Attrition/drop-out: One subject who discontinued with no post-baseline height data was excluded from the efficacy analyses; all pts were included in the safety analyses. ANOVA used for between group differences.

Quality criteria for assessment of experimental studies

1. Was the assignment to the treatment groups really random?	Unknown
2. Was the treatment allocation concealed?	Unknown
3. Were the groups similar at baseline in terms of prognostic factors?	Reported
4. Were the eligibility criteria specified?	adequate
5. Were outcome assessors blinded to the treatment allocation?	partial
6. Was the care provider blinded?	inadequate
7. Was the patient blinded?	inadequate
8. Were the point estimates and measure of variability presented for the primary outcome measure?	adequate
9. Did the analyses include an ITT analysis?	inadequate
10. Were withdrawals and dropouts completely described?	inadequate

Appendix 5 List of excluded studies

Excluded due to wrong patient group: n=40

Arends NJ, Boonstra VH, Mulder PG, Odink RJ, Stokvis-Brantsma WH, Rongen-Westerlaken C et al. GH treatment and its effect on bone mineral density, bone maturation and growth in short children born small for gestational age: 3-year results of a randomized, controlled GH trial. Clinical Endocrinology 2003; 59(6):779-787.

Arends NJ, Boonstra VH, Hokken-Koelega AC. Head circumference and body proportions before and during growth hormone treatment in short children who were born small for gestational age. Pediatrics 2004; 114(3):683-690.

Argente J, Gracia R, Ibanez L, Oliver A, Borrajo E, Vela A et al. Improvement in growth after two years of growth hormone therapy in very young children born small for gestational age and without spontaneous catch-up growth: Results of a multicenter, controlled, randomized, open clinical trial. Journal of Clinical Endocrinology and Metabolism 2007; 92(8):3095-3101.

Arwert LI, Deijen JB, Witlox J, Drent ML. The influence of growth hormone (GH) substitution on patient-reported outcomes and cognitive functions in GH-deficient patients: a meta-analysis. Growth Hormone & Igf Research 2005; 15(1):47-54.

Attanasio AF, Shavrikova E, Blum WF, Cromer M, Child CJ, Paskova M et al. Continued growth hormone (GH) treatment after final height is necessary to complete somatic development in childhood-onset GH-deficient patients. Journal of Clinical Endocrinology & Metabolism 2004; 89(10):4857-4862.

Attanasio AF, Shavrikova EP, Blum WF, Shalet SM. Quality of life in childhood onset growth hormone-deficient patients in the transition phase from childhood to adulthood. Journal of Clinical Endocrinology & Metabolism 2005; 90(8):4525-4529.

Barton JS, Hindmarsh PC, Preece MA, Brook CGD. Blood-Pressure and the Renin-Angiotensin Aldosterone System in Children Receiving Recombinant Human Growth-Hormone. Clinical Endocrinology 1993; 38(3):245-251.

Boguszewski M, bertsson-Wikland K, Aronsson S, Gustafsson J, Hagenas L, Westgren U et al. Growth hormone treatment of short children born small-for-gestational-age: the Nordic Multicentre Trial. Acta Paediatrica 1998; 87(3):257-263.

Boonstra VH, Arends NJ, Stijnen T, Blum WF, Akkerman O, Hokken-Koelega AC. Food intake of children with short stature born small for gestational age before and during a randomized GH trial. Hormone Research 2006; 65(1):23-30.

Bundak R, Darendeliler F, GüNöZ H, Ba&#351, F, Saka N et al. Growth hormone treatment in short children with intrauterine growth retardation. Journal of pediatric endocrinology & metabolism : JPEM 2001; 14(3):313-318.

Butenandt O, Lang G. Recombinant human growth hormone in short children born small for gestational age. German Study Group. Journal of Pediatric Endocrinology 1997; 10(3):275-282.

Carrascosa A, Esteban C, Espadero R, Fernandez-Cancio M, Andaluz P, Clemente M et al. The d3/fl-growth hormone (GH) receptor polymorphism does not influence the effect of GH treatment (66 microg/kg per day) or the spontaneous growth in short non-GH-deficient small-for-gestational-age children: results from a two-year controlled prospective study in 170 Spanish patients. Journal of Clinical Endocrinology & Metabolism 2006; 91(9):3281-3286.

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(2):146-153.

Chatelain P, Job JC, Blanchard J, Ducret JP, Olivier M, Sagnard L et al. Dose-dependent catch-up growth after 2 years of growth hormone treatment in intrauterine growth-retarded children. Journal of Clinical Endocrinology and Metabolism 1994; 78(6):1454-1460.

Christiansen JS, Vahl N, Norrelund H, Jorgensen JO. Effects of GH replacement in young patients with childhood onset GH deficiency. International Journal of Clinical Practice 2002; Supplement.(126):32-36.

Conway GS, Szarras-Czapnik M, Racz K, Keller A, Chanson P, Tauber M et al. Treatment for 24 months with recombinant human GH has a beneficial effect on bone mineral density in young adults with childhood-onset GH deficiency. European Journal of Endocrinology 2009;

160(6):899-907.

Crabbe R, von HM, Engrand P, Chatelain P. Recombinant human growth hormone for children born small for gestational age: Meta-analysis confirms the consistent dose-effect relationship on catch-up growth. Journal of Endocrinological Investigation 2008; 31(4):346-351.

Czernichow P, Fjellestad-Paulsen A. Growth hormone in the treatment of short stature in young children with intrauterine growth retardation. Hormone Research 1998; 49 Suppl 2:23-27.

de ZF, bertsson-Wikland K, Wilton P, Chatelain P, Jonsson B, Lofstrom A et al. Growth hormone treatment of short children born small for gestational age: Metanalysis of four independent, randomized, controlled, multicentre studies. Acta Paediatrica, International Journal of Paediatrics, Supplement 1996; 85(417):27-31.

de ZF, Butenandt O, Chatelain P, bertsson-Wikland K, Jonsson B, Lofstrom A et al. Growth hormone treatment of short children born small for gestational age: reappraisal of the rate of bone maturation over 2 years and metanalysis of height gain over 4 years. Acta Paediatrica Supplement 1997; 423:207-212.

de ZF. Growth hormone treatment of short children born small for gestational age. Clinical Pediatric Endocrinology 1997; 6(SUPPL. 10):129-133.

Fine RN, Stablein D, Cohen AH, Tejani A, Kohaut E. Recombinant human growth hormone post-renal transplantation in children: a randomized controlled study of the NAPRTCS Kidney International 2002; 62(2):688-696.

Fjellestad-Paulsen A, Czernichow P, Brauner R, Bost M, Colle M, Lebouc JY et al. Three-year data from a comparative study with recombinant human growth hormone in the treatment of short stature in young children with intrauterine growth retardation. Acta Paediatrica 1998; 87(5):511-517.

Gram J, Hansen TB, Jensen PB, Christensen JH, Ladefoged S, Pedersen FB. The effect of recombinant human growth hormone treatment on bone and mineral metabolism in haemodialysis patients. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association European Renal Association 1998; 13(6):1529-1534.

Guest G, Berard E, Crosnier H, Chevallier T, Rappaport R, Broyer M. Effects of growth hormone in short children after renal transplantation. French Society of Pediatric Nephrology. Pediatric Nephrology 1998; 12(6):437-446.

Hokken-Koelega AC, Sas T, van PY. Effects of long-term growth hormone treatment on body composition, carbohydrate metabolism, blood pressure and lipids in short children born small for gestational age. Hormone Research 2003; 59 Suppl 1:138.

Ibanez L, Fucci A, Valls C, Ong K, Dunger D, de ZF. Neutrophil count in small-for-gestational age children: contrasting effects of metformin and growth hormone therapy. Journal of Clinical Endocrinology & Metabolism 2005; 90(6):3435-3439.

Ingulli E, Tejani A. An Analytical Review of Growth-Hormone Studies in Children After Renal-Transplantation. Pediatric Nephrology 1995; 9:S61-S65.

Jorgensen JO, Norrelund H, Vahl N, Juul A, Skakkebaek NE, Christiansen JS. Continuation of growth hormone therapy versus placebo in transition-phase patients with growth hormone deficiency: impact on body composition, insulin sensitivity, and thyroid function. Journal of Pediatric Endocrinology 2002; 15 Suppl 5:1355-1360.

Juul A, Andersson AM, Pedersen SA, Jorgensen JO, Christiansen JS, Groome NP et al. Effects of growth hormone replacement therapy on IGF-related parameters and on the pituitary-gonadal axis in GH-deficient males. A double-blind, placebo-controlled crossover study. Hormone Research 1998; 49(6):269-278.

Mauras N, Pescovitz OH, Allada V, Messig M, Wajnrajch MP, Lippe B et al. Limited efficacy of growth hormone (GH) during transition of GH-deficient patients from adolescence to adulthood: a phase III multicenter, double-blind, randomized two-year trial. Journal of Clinical Endocrinology & Metabolism 2005; 90(7):3946-3955.

Nguyen VT, Misra M. Transitioning of children with GH deficiency to adult dosing: Changes in body composition. Pituitary 2009; 12(2):125-135.

Norrelund H, Vahl N, Juul A, Moller N, Alberti KG, Skakkebaek NE et al. Continuation of growth hormone (GH) therapy in GH-deficient patients during transition from childhood to adulthood: impact on insulin sensitivity and substrate metabolism. Journal of Clinical Endocrinology & Metabolism 2000; 85(5):1912-1917.

Underwood LE, Attie KM, Baptista J, Genentech Collaborative Study Group. Growth hormone (GH) dose-response in young adults with childhood-onset GH deficiency: a two-year, multicenter, multiple-dose, placebo-controlled study. Journal of Clinical Endocrinology & Metabolism 2003; 88(11):5273-5280.

Vahl N, Juul A, Jorgensen JO, Orskov H, Skakkebaek NE, Christiansen JS. Continuation of growth hormone (GH) replacement in GH-deficient patients during transition from childhood to adulthood: a two-year placebo-controlled study. Journal of Clinical Endocrinology & Metabolism 2000; 85(5):1874-1881.

Vimalachandra D, Hodson EM, Willis NS, Craig JC, Cowell C, Knight JF. Growth hormone for children with chronic kidney disease. Cochrane Database of Systematic Reviews: Reviews. Cochrane Database of Systematic Reviews 2006 Issue 3. Chichester (UK): John Wiley & Sons, Ltd; 2006.

Vimalachandra D, Craig JC, Cowell CT, Knight JF. Growth hormone treatment in children with chronic renal failure: A meta-analysis of randomized controlled trials. Journal of Pediatrics 2001; 139(4):560-567.

Whitman BY, Myers S, Carrel A, Allen D. A treatment/control group study of growth hormone treatment: Impact on behavior - A preliminary look. Endocrinologist 2000; 10(4 SUPPL. 1):31S-37S.

Willemsen RH, Arends NJ, Bakker-van Waarde WM, Jansen M, van Mil EG, Mulder J et al. Long-term effects of growth hormone (GH) treatment on body composition and bone mineral density in short children born small-for-gestational-age: six-year follow-up of a randomized controlled GH trial. Clinical Endocrinology 2007; 67(4):485-492.

Wilton P, bertsson-Wikland K, Butenandt O, Chaussain JL, de ZF, Jonsson B et al. Growth hormone treatment induces a dose-dependent catch-up growth in short children born small for gestational age: a summary of four clinical trials. Hormone Research 1997; 48 Suppl 1:67-71.

Excluded due to study design: n=27

Bannink EM, van Pareren YK, Theunissen NC, Raat H, Mulder PG, Hokken-Koelega AC. Quality of life in adolescents born small for gestational age: does growth hormone make a difference? Hormone Research 2005; 64(4):166-174.

Boguszewski MC, de ZF, bertsson-Wikland K, Nordic Study Group for Growth Hormone Treatment in SGA Children and the Belgian Study Group for Pediatric Endocrinology. Serum leptin in short children born small for gestational age: dose-dependent effect of growth hormone treatment. Hormone Research 2000; 54(3):120-125.

Carroll PV, Drake WM, Maher KT, Metcalfe K, Shaw NJ, Dunger DB et al. Comparison of continuation or cessation of growth hormone (GH) therapy on body composition and metabolic status in adolescents with severe GH deficiency at completion of linear growth. Journal of Clinical Endocrinology & Metabolism 2004; 89(8):3890-3895.

Coutant R, Carel JC, Letrait M, Bouvattier C, Chatelain P, Coste J et al. Short stature associated with intrauterine growth retardation: final height of untreated and growth hormone-treated children. Journal of Clinical Endocrinology & Metabolism 1998; 83(4):1070-1074.

Crompton C. Recombinant human growth hormone (r-hGH) treatment in children. Nephrology 2005; 10(SUPP. 5):S224-S230.

de ZF, Du Caju MV, Heinrichs C, Maes M, De SJ, Craen M et al. Early, discontinuous, high dose growth hormone treatment to normalize height and weight of short children born small for gestational age: results over 6 years. Journal of Clinical Endocrinology & Metabolism 1999; 84(5):1558-1561.

de ZF, Hokken-Koelega A. Growth hormone therapy for children born small for gestational age: height gain is less dose dependent over the long term than over the short term. Pediatrics 2005; 115(4):e458-e462.

de ZF, bertsson-Wikland K, Wollmann HA, Chatelain P, Chaussain JL, Lofstrom A et al. Growth hormone treatment of short children born small for gestational age: growth responses with continuous and discontinuous regimens over 6 years. Journal of Clinical Endocrinology & Metabolism 2000; 85(8):2816-2821.

Donaldson MDC. Growth hormone therapy in Turner syndrome - Current uncertainties and future strategies. Hormone Research 1997; 48(SUPPL. 5):35-44.

Festen DA, Vissert TJ, Otten BJ, Wit JM, Duivenvoorden HJ, et al. Thyroid levels in children with Prader-Willi syndrome before and during growth hormone treatment. Clinical Endocrinology 2007; 67:449-456.

Fine RN, Brown DF, Kuntze J, Wooster P, Kohaut EE. Growth after discontinuation of recombinant human growth hormone therapy in children with chronic renal insufficiency. The Genentech Cooperative Study Group. Journal of Pediatrics 1996; 129(6):883-891.

Fine RN, Kohaut E, Brown D, Kuntze J, Attie KM. Long-term treatment of growth retarded children with chronic renal insufficiency, with recombinant human growth hormone. Kidney International 1996; 49(3):781-785.

Fine RN. Long-term use of recombinant human growth hormone (r-hGH) in children with chronic renal insufficiency (CRI). Clinical Pediatric Endocrinology 1997; 6(SUPPL. 10):81-84.

Fine RN, Attie KM, Kuntze J, Brown DF, Kohaut EC. Recombinant human growth hormone in infants and young children with chronic renal insufficiency. Genentech Collaborative Study Group. Pediatric nephrology (Berlin, Germany) 1995; 9(4):451-457.

Hokken-Koelega A, Mulder P, De JR, Lilien M, Donckerwolcke R, Groothof J. Long-term effects of growth hormone treatment on growth and puberty in patients with chronic renal insufficiency. Pediatric Nephrology 2000; 14(7):701-706.

Hokken-Koelega ACS. Growth hormone treatment in children before and after renal transplantation. Journal of Pediatric Endocrinology & Metabolism 1996; 9:359-364.

Juul A, Vahl N, Jorgensen JO, Christiansen JS, Sneppen SB, Feldt-Rasmussen U et al. Consequences of stopping growth hormone (GH) therapy in young GH deficient patients with childhood onset disease. Growth Hormone & Igf Research 1998; 8 Suppl A:15-19.

Mahan JD, Warady BA, Fielder P, Gipson DS, Greenbaum L, Juarez-Congelosi MD et al. Assessment and treatment of short stature in pediatric patients with chronic kidney disease: A consensus statement. Pediatric Nephrology 2006; 21(7):917-930.

Mauras N, Attie KM, Reiter EO, Saenger P, Baptista J. High dose recombinant human growth hormone (GH) treatment of GH-deficient patients in puberty increases near-final height: a randomized, multicenter trial. Genentech, Inc., Cooperative Study Group. Journal of Clinical Endocrinology & Metabolism 2000; 85(10):3653-3660.

Maxwell H, Rees L. Randomised controlled trial of recombinant human growth hormone in prepubertal and pubertal renal transplant recipients. British Association for Pediatric Nephrology 79(6):481-487.

Mehls O, Broyer M, bertsson-Wikland K, Allen B, Baur L, Beetz R et al. Growth response to recombinant human growth hormone in short prepubertal children with chronic renal failure with or without dialysis. Acta Paediatrica, International Journal of Paediatrics, Supplement 1994; 83(399):81-87.

Monson JP. Indications for GH replacement in adolescents and young adults. Journal of Endocrinological Investigation 2005; 28(5 Suppl):52-55.

Nissel R, Ucur E, Mehls O, Haffner D. Final height after long-term treatment with recombinant human growth hormone (R-HGH) in children with uremic growth failure. Nephrology Dialysis Transplantation 2006; 21:367-368.

Querfeld U, Haffner D, Wuhl E, Wingen AM, Wolter K, Friedrich B et al. Treatment with growth hormone increases lipoprotein(a) serum levels in children with chronic renal insufficiency. European Journal of Pediatrics 1996; 155(10):913.

Rosenfeld RG, Attie KM, Frane J, Brasel JA, Burstein S, Cara JF et al. Growth hormone of Turner's syndrome: Beneficial effect on adult height. Journal of Pediatrics 1998; 132(2):319-324.

Simon D, Leger J, Fjellestad-Paulsen A, Crabbe R, Czernichow P. Intermittent recombinant growth hormone treatment in short children born small for gestational age: Four-year results of a randomized trial of two different treatment regimens. Hormone Research 2006; 66(3):118-123.

Wilton P, Gunnarsson R. Clinical experience with Genotropin in growth hormone deficient children. Acta Paediatrica Scandinavica - Supplement 1988; 343:95-101.

Excluded due to wrong intervention: n=4

Rosenfeld RG. Acceleration of growth in Turner syndrome patients treated with growth hormone: summary of three-year results. Journal of Endocrinological Investigation 1989; 12(8 Suppl 3):49-51.

Rosenfeld RG, Hintz RL, Johanson AJ, Sherman B. Results from the first 2 years of a clinical trial with recombinant DNA-derived human growth hormone (somatrem) in Turner's syndrome. Acta Paediatrica Scandinavica - Supplement 1987; 331:59-69.

Rosenfeld RG, Frane J, Attie KM, Brasel JA, Burstein S, Cara JF et al. Six-year results of a randomized, prospective trial of human growth hormone and oxandrolone in Turner syndrome. Journal of Pediatrics 1992; 121(1):49-55.

Wilson DM, Frane JW, Sherman B, Johanson AJ, Hintz RL, Rosenfeld RG. Carbohydrate and lipid metabolism in Turner syndrome: effect of therapy with growth hormone, oxandrolone, and a combination of both. J Pediatr 1988; 112(2):210-7.

Excluded due to wrong outcomes: n= 4

Gravholt CH, Leth-Larsen R, Lauridsen AL, Thiel S, Hansen TK, Holmskov U et al. The effects of GH and hormone replacement therapy on serum concentrations of mannan-binding lectin, surfactant protein D and vitamin D binding protein in Turner syndrome. European Journal of Endocrinology 2004; 150(3):355-362.

Lagrou K, Vanderfaeillie J, Froidecoeur C, Thomas M, Massa G, Tenoutasse S et al. Effect of 2 years of high-dose growth hormone therapy on cognitive and psychosocial development in short children born small for gestational age. European Journal of Endocrinology 2007; 156(2):195-201.

Ross JL, Feuillan P, Kushner H, Roeltgen D, Cutler GB, Jr. Absence of growth hormone effects on cognitive function in girls with Turner syndrome.[see comment]. Journal of Clinical Endocrinology & Metabolism 1997; 82(6):1814-1817.

Whitman BY, Myers S, Carrel A, Allen D. The behavioral impact of growth hormone treatment for children and adolescents with Prader-Willi syndrome: a 2-year, controlled study. Pediatrics 2002; 109(2):E35.

Reason for exclusion – repeat publication with no new randomised data:

Lindgren AC, Ritzen EM. Five years of growth hormone treatment in children with Prader-Willi syndrome. Swedish National Growth Hormone Advisory Group. Acta Paediatrica Supplement 1999; 88(433):109-111.

Reason for exclusion – conference paper pre-2006:

Fine RN, Kohaut EC, Frane JW, Perlman AJ. Multicenter Randomized Double-Blind Placebo-Controlled Study of Recombinant Human Growth-Hormone (R-hGH) in Children with Chronic-Renal-Failure (Crf). Clinical Research 1993; 41(2):A283.

Reason for exclusion – previous HTA report

Bryant J, Cave C, Mihaylova B, Chase D, McIntyre L, Gerard K et al. Clinical effectiveness and cost-effectiveness of growth hormone in children: a systematic review and economic evaluation. Health Technology Assessment (Winchester, England) 2002;6:1-168.

Reason for exclusion – conference paper abstract

Christensen T, Buckland AG, Bentley A, Djuurhus C, Wing C. Economic evaluation of somatropin (Norditropin) for the treatment of short children born small for gestational age (SGA). Value in Health 2008;11:A223.

Reason for exclusion – children of short stature – not part of scope.

Lee JM, Davis MM, Clark SJ, Hofer TP, Kemper AR. Estimated cost-effectiveness of growth hormone therapy for idiopathic short stature. Archives of Pediatrics & Adolescent Medicine 2006;160:263-9.

Reason for exclusion – Disease specific QoL measure used

Abs R, Mattsson AF, Bengtsson BA, Feldt-Rasmussen U, Goth MI, Koltowska-Haggstrom M et al. Isolated growth hormone (GH) deficiency in adult patients: baseline clinical characteristics and responses to GH replacement in comparison with hypopituitary patients. A sub-analysis of the KIMS database. Growth Hormone & Igf Research 2005;15:349-59.

Bannink EM, van Pareren YK, Theunissen NC, Raat H, Mulder PG, Hokken-Koelega AC. Quality of life in adolescents born small for gestational age: does growth hormone make a difference? Hormone Research 2005;64:166-74.

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.

Deijen JB, Arwert LI, Witlox J, Drent ML. Differential effect sizes of growth hormone replacement on Quality of Life, well-being and health status in growth hormone deficient patients: a meta-analysis. Health & Quality of Life Outcomes 2005;3:63.

Dixon S, McEwan P, Currie CJ. Estimating the health utility of treatment in adults with growth hormone deficiency. Journal of Outcomes Research 2003;7:1-12.

Koltowska-Haggstrom M, Hennessy S, Mattsson AF, Monson JP, Kind P. Quality of life assessment of growth hormone deficiency in adults (QoL-AGHDA): comparison of normative reference data for the general population of England and Wales with results for adult hypopituitary patients with growth hormone deficiency. Hormone Research 2005;64:46-54.

Koltowska-Haggstrom M, Mattsson AF, Monson JP, Kind P, Badia X, Casanueva FF et al. Does long-term GH replacement therapy in hypopituitary adults with GH deficiency normalise quality of life? European Journal of Endocrinology 2006;155:109-19.

Lagrou K, Xhrouet-Heinrichs D, Massa G, Vandeweghe M, Bourguignon JP, De SJ et al. Quality of life and retrospective perception of the effect of growth hormone treatment in adult patients with childhood growth hormone deficiency. Journal of Pediatric Endocrinology 2001;14 Suppl 5:1249-60.

Malik IA, Foy P, Wallymahmed M, Wilding JPH, MacFarlane IA. Assessment of quality of life in adults receiving long-term growth hormone replacement compared to control subjects. Clinical Endocrinology 2003;59:75-81.

McKenna SP, Doward LC, Alonso J, Kohlmann T, Niero M, Prieto L et al. The QoL-AGHDA: an instrument for the assessment of quality of life in adults with growth hormone deficiency. Quality of Life Research 1999;8:373-83.

McMillan CV, Bradley C, Gibneyt J, Russell-Jones DL, Sonksent PH. Evaluation of two health status measures in adults with growth hormone deficiency. Clinical Endocrinology 2003;58:436-45.

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.

Saller B, Mattsson AF, Kann PH, Koppeschaar HP, Svensson J, Pompen M et al. Healthcare utilization, quality of life and patient-reported outcomes during two years of GH replacement therapy in GH-deficient adults--comparison between Sweden, The Netherlands and Germany. European Journal of Endocrinology 2006;154:843-50.

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 comment]. Journal of Clinical Endocrinology & Metabolism 1998;83:1134-42.

Sheppard L, Eiser C, Davies HA, Carney S, Clarke SA, Urquhart T et al. The effects of growth hormone treatment on health-related quality of life in children.[see comment]. Hormone Research 2006;65:243-9.

Reason for exclusion – Mixed patient group of adults and children

McMillan CV, Bradley C, Gibney J, Healy ML, Russell-Jones DL, Sonksen PH. Psychological effects of withdrawal of growth hormone therapy from adults with growth hormone deficiency. Clinical Endocrinology 2003;59:467-75.

Reason for exclusion – Review article

Petrou S, McIntosh E. Measuring the benefits of growth hormone therapy in children: A role for preference-based approaches? Archives of Disease in Childhood 2008;93:95-7.

Reason for exclusion – Unclear whether adult or child onset
Suzukamo Y, Noguchi H, Takahashi N, Shimatsu A, Chihara K, Green J et al. Validation of the Japanese version of the Quality of Life-Assessment of Growth Hormone Deficiency in Adults (QoL-AGHDA). Growth Hormone & Igf Research 2006;16:340-7.

Appendix 6 List of eligible abstracts

The following conference abstracts were identified in searches and were of relevance to the review, but did not contain sufficient information to be included.

Gardner M, Boshart M, Carron L, Sandberg D. Effects of growth hormone in childhood on quality of life endpoints: a systematic review. Paediatric Acadamic Societies (PES) Conference, Baltimore. May 2009.

Salgin B. Effect of growth hormone treatment on insulin secretion and sensitivity in relation to growth of children born small for gestational age. Horm Res 2008; 70 (suppl 1); 76.

Phillip M, Lebl J, Steensberg A, Kappelgaard, A-M, Ibanez L. Metabolic parameters during growth hormone treatment in short children born small for gestatioal age. Horm Res 2008; 70 (suppl 1); 100.

Appendix 7 List of ongoing studies

Searches identified 2 relevant RCTs which are ongoing:

Study NCT00190658 aims to compare the mean first year GV of somatropin-treated prepubertal patients with SHOX-D with the GV of a control group of untreated prepubertal patients with

SHOX-D. Both groups will be compared to a somatropin-treated group of girls with Turner syndrome. Sponsor: Eli Lilly and Company Estimated end date: December 2010.

Study NCT00625872 focuses on the effect of a one year Somatropin treatment (0.035 mg/kg/d or 0.067 mg/kg/d) in short children born SGA on neuromuscular function and cognitive performance. Height gain and growth velocity are included as secondary outcome measures. Inclusion Criteria are birth length- and/or birth weight-SDS adjusted to gestational age < -2.0, current height-SDS < -2.5 and parental adjusted height-SDS below -1, growth velocity SDS < 0 during the last year before inclusion. Sponsor: Pfizer; end date: not reported.

Appendix 8 Critique of industry submissions (clinical effectiveness)

Six of the seven manufacturers submitted reports to NICE, and these are briefly appraised below. Please see Section 4.3 for a discussion of the economic models and results included in the manufacturers' submissions.

SHTAC review of clinical effectiveness in Eli Lilly's submission

Comprehensiveness of ascertainment of published studies

- The MS uses the Novo Nordisk systematic review, which did not include SHOX. The MS states (p.13) that the evidence for SHOX came from Lilly's databases, i.e. there was no systematic review for this. The conditions listed as inclusion criteria for SGA include intrauterine growth retardation, which was not in the NICE scope. The comparator was clearly stated to be 'no treatment'. However, the inclusion criteria also state that active-controlled RCTs were included. This is then contradicted by the exclusion criteria which state that studies comparing somatropin with other treatments known or presumed to affect growth would be excluded.
- The MS clearly reports search dates, search strategies and databases searched.
- Enough detail was provided for the searches to be reproducible.
- The MS does not present information on searches for ongoing studies.
- Conference proceedings were excluded from the review.
- The MS includes a separate search for QoL data in adolescents and adults.

Searches identified:

The MS contains a summary of the included trials, but there is no tabulation of details such as study type, treatment arms etc. The review included the following RCTs:

- GHD: 5 placebo/no treatment-controlled RCTs (mostly during transition phase) – Jorgensen 2002 [excluded by SHTAC as mean age=20], Underwood 2003 [excluded by SHTAC as mean age=23.8], Drake 2003, Shalet 2003 and Mauras 2005 [all excluded by SHTAC as pts had completed linear growth]. The manufacturer included 6 other studies which were either dosing studies or compared two different versions of somatropin.
- TS: 9 RCTs (3 placebo controlled: Gravholt 2002, 2005; Quigley 2002; all included by SHTAC) and 6 other studies (Bannick 2006, van Pareren 2003, Sas 2001 [all excluded by SHTAC as dose studies], Davenport 2007, Johnston 2001, CGHAC 2005 [all included by SHTAC]).
- CRI: 4 RCTs: de Graaf 2003 [SHTAC excluded as this is analysis of body proportions in an RCT that we have already included for height and body composition outcomes – Hokken Koelega 1991], Hertel 2002[SHTAC excluded as compares 2 doses, no placebo arm], Sanchez 2002[included by SHTAC], Fine 2002 [SHTAC excluded as includes pubertal children, with no separate data analysis].
- SGA – 20 RCTs identified, of which 6 had placebo or no treatment as control arm (Boguszweski 1998, Butenadt 1997, Arends 2003 2004, Boonstra 2006) [SHTAC excluded these as patient group did not meet our criteria] van Pareren 2003 [SHTAC excluded as this is a follow up of a dose-response study].
- *PWS not relevant for this drug*
- SHOX – not included in systematic review. Reported data comes from GDFN study (n=78), Blum et al 2007 (SHTAC included this).
- None of the additional studies met SHTAC's inclusion criteria.

Clinical Analysis:

- The MS also reports observational studies, in particular data from the KIGS database.
- Given that the manufacturer included a range of studies which did not meet SHTAC's inclusion criteria, it is not possible to compare their conclusions with SHTAC's.
- The MS did not include a meta analysis or indirect comparison.
- The MS includes a short narrative summary of the included trials for each disease, but there is no overall tabulation of the included studies' characteristics or results, and no quality assessment of the trials.
- The MS uses the same outcome measures as the SHTAC review.

- The MS reports more detail on AE from observational studies in addition to the limited information available in the RCTs.

Interpretation:

- The MS does not present any tabulated data from the studies included in the systematic review; there is simply a short narrative summary of each disease. It is therefore not possible to assess whether or not the manufacturer's analysis is supported by data in the included trials.

Key issues:

- The manufacturer's systematic review included a broad range of studies, for example dosage studies, which did not meet their own inclusion criteria.
- Very little detail is presented for the included studies (e.g. patient characteristics, treatment arms, length of study) and there is no tabulation of data. The manufacturer's conclusions seem to be based on both trials which met their inclusion criteria and those which clearly did not (e.g. dosage studies).

SHTAC review of clinical effectiveness in Novo Nordisk's submission

Comprehensiveness of ascertainment of published studies

- Databases searched and the dates of searches are specified. Searches were conducted from the date of the original NICE appraisal – w/c 28th August 2008, and from 1996 to w/c 28th August 2008 for SGA (not included in the last review).
- Search strategies are supplied in the appendices
- Search strategies are detailed and appear reproducible
- Novo Nordisk does not appear to have searched for other ongoing studies, but do report on two ongoing studies, specifically of Norditropin – NESGAS and NordiNet IOS.
- Conference proceedings were not searched for and are listed in the exclusion criteria.

Clinical Analysis:

- Novo Nordisk did not include PWS or SHOX. Uncontrolled trials were included. For long term effects of rhGH treatment i.e final height/adult height/ near adult height open-label extension studies were 'deemed to be appropriate as the length of the RCTs was likely to be

too short to capture the long term treatment effect'. Dose-response trials have been included. In the case of SGA these from the majority of the submission.

- SGA: Novo Nordisk have included 21 studies. None of these were included in SHTAC's MTA. Exclusions in the SHTAC MTA were on the basis of patient group not meeting the inclusion criteria, or on design as 14 of the 21 were dose response studies. The five studies included in our MTA were not included in the Novo Nordisk submission. Novo Nordisk also included open-label extension studies.
- GHD: Novo Nordisk have included 13 studies. One of these is the GHD study included in SHTAC's MTA. Eight are transition phase studies – these are not included in SHTAC's systematic review. Four are dose response studies and therefore excluded from the MTA. Two are biosimilars compared with their reference product.
- TS: Novo Nordisk discuss the Turner Cochrane Review. 23 studies were included, including the six included in SHTAC's MTA. The remaining studies were dose response, with the exception of one which compared once versus twice daily injections.
- CRI: Novo Nordisk have included nine studies, five of which were included in SHTAC's MTA. Of the four excluded from the MTA, two were dose response studies, one was excluded on patient group
- Nothing in the excluded reasons indicates why all of SHTAC's included SGA papers are excluded.

Conclusions

SGA: It is not possible to compare the conclusions as the studies included in the two reviews are so different.

GHD: Again the conclusions are difficult to compare as Novo Nordisk include transition phase studies, which SHTAC excluded from the main systematic review as patients had completed linear growth; dose response studies; and studies comparing biosimilars to their reference drug. Novo Nordisk's conclusions tend to be based on dose–response studies, and how far an outcome/ result is dose-dependant.

TS: Novo Nordisk concludes that height is improved in a 'dose-dependant' manner: The SHTAC MTA does not include dose-response studies or consider dose issues. SHTAC has concluded that there is evidence of improved body composition and height outcomes in girls with TS; this needs to be weighed against issues of quality of reporting and size of trials.

CRI: Height conclusions are dose related, and body composition 'does not appear to be negatively influenced by rhGH therapy'.

- Outcome measures are broadly similar.
- Additional adverse event rates from KIGS and NCGS databases are included in an appendix

Interpretation:

SGA: Conclusions do not appear to fully reflect Novo Nordisk's analyses, although the analysis contains few results, and is a broad summary in itself. Very few of the points discussed in the analyses compare treated and untreated groups, predominantly focusing on dose-response or differences in the treated group from baseline.

TS: Apart from height outcomes, few results are reported, and again the focus is often on dose-related effects. The summary somewhat overstates the evidence presented.

CRI: Conclusions do appear to match analyses, although again few detailed results are presented. Novo Nordisk does not comment on quantity/ quality of research available to support their conclusions.

GHD: Novo Nordisk considered transition phase studies alongside non-transition phase studies for height and other outcomes, but separately for biochemical/ body composition markers. The authors then summarise that the transition phase studies may lead to an underestimation of growth in children with growth hormone deficiency. Other conclusions appear to match the analyses.

Quality is discussed to a degree in the results sections – it is mentioned for example if trials are short, or low in patient numbers. However this, or its possible effects on conclusions/ findings, is not referred to in the summary.

Key issues:

- The submission does not include the SGA papers included in SHTAC's review, but does include studies whose patients do not meet the birth length/WtSDS criteria and/or current HtSDS criteria included.
- Dose response studies are included for all conditions.

SHTAC review of clinical effectiveness in Pfizer's submission

Comprehensiveness of ascertainment of published studies

- The manufacturer supplied full details of the systematic review, specifying dates and databases searched.

- Search strategies were supplied.
- Enough detail was provided for the searches to be reproducible.
- Inclusion criteria differed from that used by SHTAC in that cohort, observational, and retrospective studies were included. The manufacturer's inclusion criteria defined children as being <16 years old, whereas SHTAC included those up to 18 since they may still be growing and thus able to benefit from rhGH treatment. The manufacturer did not specify what the comparator should be (NICE's final scope indicates that this should be treatment without somatropin).
- The manufacturer restricted the review to only those studies which used Genotropin, or were sponsored by Pfizer. They excluded studies which used a competitor's brand of somatropin. However, they also report the results of the Novo Nordisk full systematic review – see SHTAC assessment of the Novo Nordisk MS for more details.
- The MS does not report ongoing studies.
- The MS does not state whether or not they searched for conference proceedings.

Searches identified (studies for Genotropin):

- GHD: 3 RCTs and 17 observational studies. None of the 3 RCTs met our inclusion criteria. Coelho et al. (2008) compared 2 doses of genotropin; Romer et al. (2007) compared omnitrope with genotropin; Dorr et al. (2003) compared genotropin delivered via 2 different devices.
- TS: 1 RCT and 8 observational studies: The single RCT by Johnston (2001) was also included in the SHTAC review.
- PWS: 12 RCTs (3 from previous appraisal) and 6 observational studies. One of these (Festen 2007) is not included in our review as it is not a fully randomised study (children were stratified by age, and only the under 12s were randomized. Older children were all given rhGH, but results were not reported separately for the randomised patients). Two of the studies included by the manufacturer have been combined by SHTAC, as they report data from the same RCT (Festen et al. 2008 and de Lind van Wijngaarden 2009 (cited as Roderick et al 2009 in the MS).
- CRI: no new RCTs, 3 observational studies. The submission only discusses the Broyer study from the previous review, and not the others SHTAC included as these weren't Genotropin.
- SGA: 13 RCTs, 10 observational studies. Of the 13 RCTs, only 5 reported treatment vs. no treatment/placebo. SHTAC excluded the review by Lagrou (2007) as its outcomes did not meet our inclusion criteria. We also excluded the reviews by Bundak 2001 and Carracosa

2006 as their patient groups did not match our criteria. We included the de Schepper 2008 study and the de Zegher 2002 studies.

- None of the manufacturer's included studies reported QoL as an outcome measure
- The MS also includes a summary of the Novo Nordisk systematic review. Please see SHTAC's appraisal of that submission for further details.

Clinical Analysis:

- The manufacturer has only included RCTs of its own brand of somatropin, so it is not possible to compare their findings directly with SHTAC's.
- GH and SGA RCTs – the MS and SHTAC reviews included different RCTs, so it is not possible to compare the evidence reported. The RCTs included for GHD were not placebo/no treatment controlled.
- PWS – the MS includes two studies (Roderick et al 2009 and Festen et al 2008) which appear to be the same RCT – SHTAC has treated these as one RCT to avoid double-counting.
- Given that the manufacturer included a range of studies which did not meet SHTAC's inclusion criteria, and focussed only on studies of their own product, it is not possible to compare their conclusions directly with SHTAC's.
- The MS did not include a meta analysis or indirect comparison. Results are presented in tables and there is a narrative synthesis for each disease area.
- The MS uses the same outcome measures as the SHTAC review.
- The MS includes data from the KIGS database, which is not included in the SHTAC review of clinical effectiveness as it is observational data. Additional adverse event data from the KIGS database is presented on p.97 of the MS.

Interpretation:

- The manufacturer's interpretation of the clinical data in the RCTs matches their analyses.
- There are separate sections discussing the results of RCTs and of observational studies.
- Data from observational studies have not been checked by SHTAC.

Key issues:

- The manufacturer's systematic review included dose comparison studies for GHD, which SHTAC excluded.
- Many of the studies included for the manufacturer's review of SGA studies were excluded by SHTAC as their patients did not meet our inclusion criteria.

SHTAC review of clinical effectiveness in Merck Serono's submission

Comprehensiveness of ascertainment of published studies

- The MS uses the SHTAC review conducted in 2002 [5] and the systematic review conducted by Novo Nordisk for studies published since then (see Novo Nordisk critique) for the licensed indications for Saizen (GHD, TS, CRI and SGA).

Searches identified:

- Studies identified and reported are all those from the previous SHTAC report (RCTs and non RCTs reporting FH) plus RCTs published since then identified by the Novo Nordisk review.
- GHD: No additional RCTs were reported for GHD although an additional one is included in the SHTAC MTA (Mauras 2005).
- TS: 4 RCTs (Johnston 2001; CGHAC 2005; Quigley 2002; Davenport 2007). However, the MS did not identify 2 RCTs included in the SHTAC MTA (both Gravholt 2005).
- CRI: 3 RCTs (de Graaf 2003; Fine 2002; Sanchez 2002). Two of these (de Graaf and Fine) are not included in the SHTAC MTA review because they do not meet our inclusion criteria. One RCT (Fine 2004) is not included in the MS but meets the SHTAC MTA inclusion criteria and is therefore included in that.
- SGA: 4 RCTs (Buttenandt 1997; Boguszewski 1998; Arends 2004; Van Pareren 2003). These do not match the studies identified in the SHTAC MTA (from which they are excluded on the basis of patient group and study design).
- The MS does not identify any RCTs that meet the inclusion criteria of the SHTAC MTA which are not already included.

Clinical Analysis:

- Evidence reported is broadly similar to the SHTAC MTA in that it uses RCTs in the original SHTAC report; some discrepancies on RCTs since that time and on the extra indication SGA.
- Narrative synthesis is somewhat selective. All included studies are tabulated, but only height results are reported.
- MS also includes some non-systematic review data on psychological outcomes and body composition, and long term data from the KIGS observational database.
- Conclusions are generally similar to the SHTAC MTA.

GHD: The MS has used the previous SHTAC review so conclusions on growth are similar but no data no lean body mass/biochemical markers.

TS: Conclusions are broadly similar to the SHTAC MTA in terms of growth and lean body mass.

CRI: Conclusions broadly similar to the SHTAC MTA in terms of growth; no statement on other outcomes.

SGA: Conclusions broadly similar to the SHTAC MTA in terms of growth; no statement on other outcomes.

- Growth outcomes measures are same as the SHTAC MTA.

Interpretation:

- Overall MS interpretation of the clinical data matches the MS analyses although the MS relies heavily on the previous SHTAC report. The new evidence is not really synthesised except for SGA which includes studies not in the SHTAC MTA. Conclusions are based on selective statements and focus on height outcomes.
- MS states that new data has ‘not materially changed the understanding of the efficacy of GH in children’.

Questions:

- The major areas of discrepancy compared with the SHTAC MTA relate to studies omitted from the MS (GHD 1; TS 2, CRI 1 and SGA 5).

SHTAC review of clinical effectiveness in Ipsen Limited's submission

Comprehensiveness of ascertainment of published studies

- The databases and dates searched are specified.
- Search strategies were supplied and appear comprehensive enough to be reproducible.
- Ongoing studies were not searched for or reported in this submission.
- Conference proceedings were excluded.
- This review includes CRI, GHD and TS, 'somatropin' as intervention, including products from other manufacturers, and published and available in full studies in the English language. Exclusion criteria given but reasons for individual studies' exclusions not stated.

- Assessment of article quality looks at allocation concealment, patient blinding, investigator blinding, baseline differences of the experimental groups and 'completeness of follow up' (assume withdrawals?). Did not appear to assess if there was an intention-to-treat analysis, care-provider blinding.

Clinical Analysis: For the results of the systematic review, we are referred to the submission prepared by Novo Nordisk. Studies are not referenced in the text. No conclusions in this submission, apart from on the limitations of RCTs for final height data, and the subsequent need to rely on observational studies (i.e. KIGS database) for this. The number of studies for each condition reporting certain outcomes is given, but the results are in the Novo Nordisk submission and not detailed in the Ipsen submission.

- Manufacturer has included 11 GHD studies; most appear to be transition phase studies.
- MS states 9 TS studies found.
- MS states 4 CRI studies.
- Limited new data on final height from RCTs so appear to have included observational studies for this outcome. However, no references are given in the text so cannot check
- The MS states that 'there are limited data available on the effect of GH on height in RCTs [therefore] use of observational data from...KIGS was appropriate.' This appears to have been employed to inform the economic model.
- A 'rapid appraisal of the literature' was undertaken by Eli Lilly for QoL – referred for this to Eli Lilly submission – 'impact of short stature in adults' due to lack of data on children and quality of life.
- No conclusions stated here: referred to Novo Nordisk submission.
- There are no indirect comparisons included here.
- No outcome results are reported here, but those outcomes reported in the included studies reflect those in the SHTAC review.

GHD: 4/8 studies reporting 'AE found that a higher dose was associated with a greater incidence of AEs and/or serious AEs'. The remaining studies reported no differences between groups. Only 1 study in the SHTAC review reported AE, with a slightly higher percentage in the GH group experiencing these. Only one event in each group was thought to be study drug related: edema in GH and sluggishness in placebo. MS reports AEs thought to be related to study drug.

CRI: 3 studies in the MS report AE, 1/3 reported higher number of SAEs (serious AE) related to GH therapy compared with no treatment, another study reported SAEs that were 'therapy-related'. SAEs related to therapy reported here include diabetes mellitus, hypertension and injection pain. This is not reflected in the studies included in SHTAC review. Difficulty with comparisons as there are no references in the text.

- TS: A greater incidence of AE in the GH group was reported in two out of four studies reporting AEs in the SHTAC review. In the MS one study showed GH to be associated with 'greater incidence of treatment emergent AEs'. No major differences between the groups were found in the other studies in the MS
- No references are given for these studies and AEs, no proportions/ means are reported, just these general results.

Interpretation:

- No interpretation included here – referred to the Novo Nordisk submission.

Key Issues:

Inclusion of observational data to inform final height differs from SHAC review

Studies not referenced here – can't cross-check with SHTAC review. See Novo Nordisk submission for further details.

SHTAC review of clinical effectiveness in Sandoz's submission

Comprehensiveness of ascertainment of published studies

- The submission did not include a systematic review, so there were no details of search strategies, databases or dates searched.

Searches identified:

- The MS includes details of 2 phase III studies: AQ-study and LYO-study. Neither meets SHTAC's inclusion criteria; AQ-study compares different doses of Omnitrope with a reference product, and LYO-study is a non-comparative trial.

Clinical Analysis:

- The evidence reported in the Sandoz submission is from trials specific to their biosimilar product. The submission does not include any trials of rhGH vs. no treatment. It is therefore not possible to compare their submission with the evidence presented in the SHTAC systematic review.
- The submission uses the same outcome measures as the SHTAC review.
- The submission includes a summary of AE from the AQ-study and the LYO-study, neither of which was included in the SHTAC review. The manufacturer stated that the safety profiles of Omnitrope and Genotropin were comparable.

Interpretation:

- The manufacturer's interpretation of the clinical data matches their analyses.

Key issues:

The manufacturer presents evidence for the use of Omnitrope compared with other somatropin formulations, but does not present any information for its effectiveness compared with no treatment. The included studies did not meet SHTAC's inclusion criteria.

Appendix 9 Critical appraisal of manufacturers' economic evaluation

Table A1 Critical appraisal checklist of economic evaluation (Questions in this checklist based on Drummond and Jefferson, the NICE reference case, and the ISPOR checklist.)

	Item	MS
1	Is there a well defined question?	Yes
2	Is the patient group in the study similar to those of interest in UK NHS?	Yes
3	Is the correct comparator used that is routinely used in UK NHS?	Yes
4	Is the study type and modelling methodology reasonable?	Yes
5	Is an appropriate perspective used for the analysis?	Yes
6	Is the health care system or setting comparable to UK?	Yes
7	Is the effectiveness of the intervention established based on a systematic review?	No
8	Is the model structure appropriate and does it fit with the clinical theory of the	Yes

	disease process?	
9	Are assumptions reasonable and appropriate?	Yes
10	Are health benefits measured in QALYs using a standardised and validated generic instrument from a representative sample of the public?	Yes
11	Are the resource costs used reasonable and appropriate for the UK NHS?	Yes
12	Are the health states and parameters used in the model described clearly and are they reasonable and appropriate for the UK NHS?	?
13	Is an appropriate discount rate used?	Yes
14	Has the model been validated appropriately?	?
15	Is sensitivity analysis undertaken and presented clearly?	Yes

Yes / No / ? (unclear or partially true)

Appendix 10 Critical appraisal of Sandoz MS (cost-effectiveness)

This appendix describes a critical appraisal of the cost effectiveness section of the Sandoz MS. The submission attempts a cost-minimisation analysis comparing Omnitrope with Genotropin (which was defined as the reference product) in patients with GHD and TS, rather than a cost effectiveness analysis. There is no indication that a systematic review of clinical evidence has been undertaken. The cost effectiveness analysis according to NICE guidance[138] was not presented.

Appraisal of the manufacturer cost effectiveness analysis

A summary of the manufacturer's submission compared with the NICE reference case requirements is given in Table A2.

Table A2 Assessment of Sandoz submission against NICE reference case requirements

NICE reference case requirements	**Included in submission**
Decision problem: as per the scope developed by NICE	✗[#]
Comparator: no treatment alternative	✗[#]
Perspective on costs: NHS and PSS	✓[†]

Perspective on outcomes: all health effects on individuals	×[+]
Type of economic evaluation: cost-effectiveness analysis	×
Synthesis of evidence on outcomes: based on a systematic review	No evidence synthesis
Measure of health benefits: QALYs	×
Description of health states for QALY calculations: use of a standardised and validated generic instrument	×
Method of preference elicitation for health state values: choice based method (e.g. TTO, SG, not rating scale)	×
Source of preference data: representative sample of the public	×
Discount rate: 3.5% p.a. for costs and health effects	×
Notes (✓=yes; × = no; ? = uncertain; N/A=not applicable): [#] scope states that rhGH (somatropin) be compared with no treatment alternative. The cost comparison includes only omnitrope and genotropin. [†] only costs of pharmaceuticals omnitrope and genotropin are included in cost-comparison [+] the MS does not include an economic evaluation according to the NICE guidance. Patient outcomes (either observed or the final outcomes) are not included in the health economics part of the MS	

Summary of general concerns

The MS did not comply with NICE's recommended structure[138] and did not estimate QALYs or present cost-effectiveness analysis. The MS attempted a cost-minimisation analysis implicitly suggesting that treatment with Omnitrope is equally effective as treatment with Genotropin (in terms of additional height in children with GHD and TS) but is associated with less cost to the NHS. Due to the number of uncertainties it is not clear whether this assertion is justified. In particular, there was limited clinical efficacy data to support the non-inferiority of Omnitrope compared to Genotropin. The only head-to-head RCT comparing Omnitrope with Genotropin was of insufficient duration and might not have been designed as a non-inferiority trial. The MS did not include any clinical evidence in relation to licensed indications other than GHD. Without clinical evidence that unequivocally demonstrated the non-inferiority of Omnitrope in comparison with Genotropin, the results of a cost-minimisation analysis can not be confirmed.

The results of the cost-comparison reported in the MS were not comparable with the results of cost-effectiveness analysis reported in the submissions by Pfizer, Eli Lilly, Ipsen, Novo Nordisk and Merck Serono because Sandoz have not presented results either as an estimated incremental cost per QALY or as an incremental cost per extra cm gained, and the reported cost was neither a life-time cost nor the cost per duration of treatment (until near-adult height is achieved).

Appendix 11 Quality of life from HSE 2003

The Health Survey for England database was reanalysed in a similar way to Christensen and colleagues for adults aged older than 18 years. The HSE 2003 contains variables for height (estht) and EQ5-D (eqmean). Incomplete records were omitted. For those with complete records (n = 13321), the HSE 2003 data had mean adult height for males of 175 cm (SD 7.2) and mean adult height for females of 161 cm (6.8). There were 50 observations less than – 3 SDS or greater than 3 SDS (ie 0.4%) and 617 observations less than -2 SDS or greater than 2 SDS (4.6%).

An analysis was completed to see the effect of different ages on QoL scores using a subset of people of age 18 to 49 years and over 50 years old. QoL score for all ages was 0.86, age 18 - 49 years QoL had mean 0.91 (SD = 0.18) and age 50+ yrs QoL had mean 0.8 (SD = 0.26). The QoL in the younger category was significantly better than for the older category and so it is logical to estimate the EQ5-D for each of these age groups.

For the SDS <-3 there were few individuals in this group and the estimates are highly variable. In addition the majority of these individuals are in the older age group (mean age 72 years). It is therefore more logical to fit the distribution to all data and use this in the model.

Table A3 Frequency of individuals at different ages and HtSDS in HSE 2003

	Age 18 to 49 years		Age 50+ years	
SDS	n	Eqmean	n	Eqmean
<-3.0	5	0.85	24	0.63
-3 to <-2.5	6	0.75	62	0.70
-2.5 to <-2.0	42	0.88	161	0.73
-2 to <-1.5	140	0.85	397	0.78
-1.5 to <-1.0	475	0.91	798	0.79
-1.0 to <-0.5	845	0.90	1133	0.78
-0.5 to <0	1331	0.90	1288	0.82
0 to <0.5	1485	0.91	1029	0.81
0.5 to <1.0	1288	0.91	707	0.83
1.0 to <1.5	837	0.91	368	0.84

1.5 to <2.0	431	0.91	152	0.85
2.0 to <2.5	201	0.92	41	0.84
2.5 to <3.0	42	0.89	12	0.83
>3.0	20	0.98	1	0.90

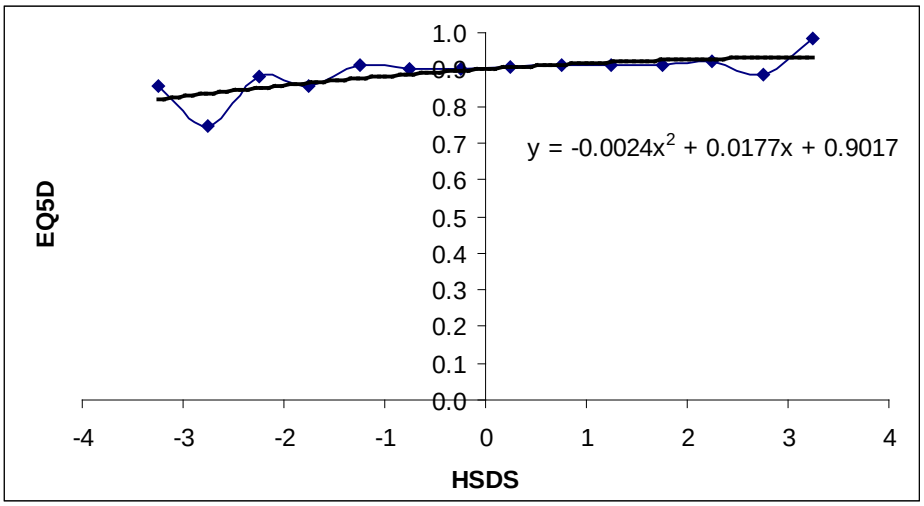


Figure A1 Relationship between height (HtSDS) and EQ-5D score for adults aged 18 to 50 years in HSE 2003

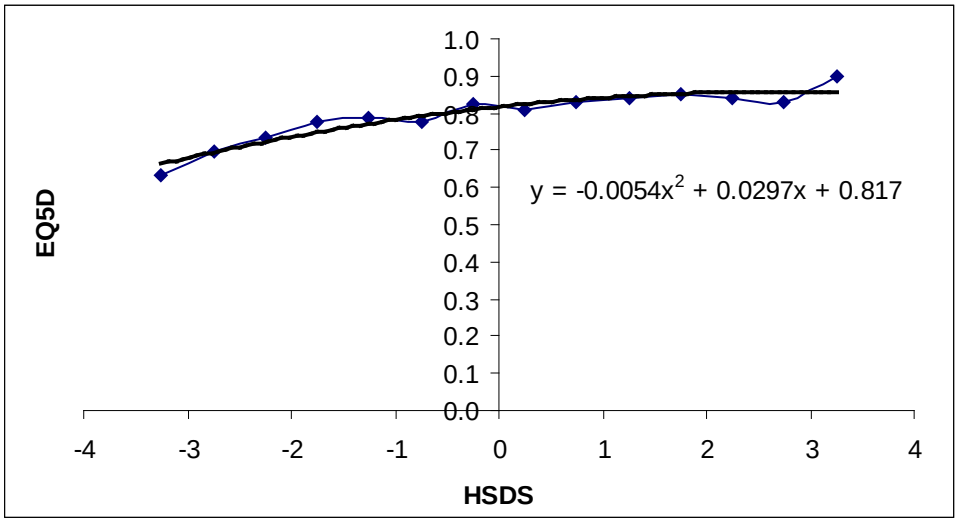


Figure A2 Relationship between height (HtSDS) and EQ-5D score for adults aged older than 50 years in HSE 2003

Table A4 Quality of life from fitted values

Age	Fitted QoL score
18-49 years	$-0.0024x^2 + 0.0177x + 0.9017$
> 50 years	$-0.0054x^2 + 0.0297x + 0.817$

Appendix 12 Input parameters for probabilistic sensitivity analysis.

The distribution assigned to each variable included in the probabilistic sensitivity analysis and the parameters of the distribution are reported in this appendix.

Health state utility

The utility increments for HtSDS below -2.0, between -2.0 HtSDS and 0 HtSDS and above 0 HtSDS were sampled using estimated standard errors. These were derived from an assumption that a variation of plus or minus 20% was an appropriate confidence interval for the average utility gain. No other summary statistic was available. These were sampled using a normal distribution.

Table A5 Health state utility parameters and distribution

Health state utility	Mean	"Standard error"	Lower 95% CI	Upper 95% CI	Distribution
Below -2 HtSDS	0.061	0.0061	0.049	0.730	Normal
Between -2 and 0 HtSDS	0.010	0.0010	0.008	0.120	Normal
Above 0 HtSDS	0.002	0.0002	0.0016	0.0024	Normal

Compliance

The compliance of the model was based on the on the range of 69% to 95% compliance estimated in the compliance review conducted by Merck Serono. The estimated "standard errors" for compliance was derived from this range, as this was thought to provide the best estimate of variability due to lack of other summary data.

Table A6 Compliance parameters and distribution

	Mean	"Standard error"	Alpha	Beta	Distribution
Compliance	0.85	0.085	14.150	2.497	Beta

Height standard deviations

The reported mean HtSDS were taken from the applicable RCTs and KIGS data for both the treated and untreated groups consistent with the basecase analysis. The standard errors were calculated for each mean HtSDS, except for PWS where there was no mean reported. In this case a median value was assumed to adequately represent the mean. A standard deviation of 1 was used to estimate the standard error for PWS. This is consistent with the level of dispersion reported for the other conditions. The HtSDS were simulated using the normal distributions. See table below for mean and standard errors for each condition:

Table A7 HtSDS parameters and distribution

Condition	HtSDS	Mean	Standard error	Distribution
GHD	Treated baseline	-2.99	0.0134	Normal
	Treated end	-1.17	0.0216	Normal
	Untreated baseline	-2.99	0.0134	Normal
	Untreated end	-2.99	0.0216	Normal
TS	Treated baseline	-3.40	0.1152	Normal
	Treated end	-1.80	0.0206	Normal
	Untreated baseline	-3.40	0.1220	Normal
	Untreated end	-3.10	0.2294	Normal
PWS	Treated baseline	-2.00	0.2000	Normal
	Treated end	-0.50	0.2085	Normal
	Untreated baseline	-2.50	0.2132	Normal
	Untreated end	-2.60	0.2182	Normal
CRI	Treated baseline	-2.90	0.1214	Normal
	Treated end	-1.60	0.1925	Normal
	Untreated baseline	-2.90	0.0994	Normal
	Untreated end	-2.90	0.1525	Normal
SGA	Treated baseline	-3.10	0.0700	Normal
	Treated end	-2.30	0.0840	Normal
	Untreated baseline	-3.10	0.0729	Normal
	Untreated end	-3.00	0.0894	Normal
SHOX	Treated baseline	-3.30	0.1925	Normal
	Treated end	-2.10	0.0385	Normal
	Untreated baseline	-3.30	0.1600	Normal
	Untreated end	-3.00	0.0408	Normal

Starting age and treatment length

The starting age and treatment length were sampled using estimated "standard errors." These were derived from confidence intervals placed two years either side of the mean starting age and treatment length. This method was used instead of calculating the standard errors from the KIGs database. It was felt that the very small standard errors from KIGs did not reflect the possible variability in starting age and treatment length. These were sampled using normal distributions.

Table A8 Starting age and treatment length parameters and distribution

Starting age	Mean	"Standard error"	Lower 95% CI	Upper 95% CI	Distribution
GHD	9.0	1.020	7.0	11.0	Normal
TS	10.0	1.020	8.0	12.0	Normal
PWS	7.0	1.020	5.0	9.0	Normal
CRI	9.0	1.020	7.0	11.0	Normal
SGA	8.0	1.020	6.0	10.0	Normal
SHOXs	8.0	1.020	6.0	10.0	Normal
Treatment Length	Mean	"Standard error"	Lower 95% CI	Upper 95% CI	Distribution
GHD	7.0	1.0200	5.0	9.0	Normal
TS	6.0	1.0200	4.0	8.0	Normal
PWS	8.0	1.0200	6.0	10.0	Normal
CRI	5.0	1.0200	3.0	7.0	Normal
SGA	6.0	1.0200	4.0	8.0	Normal
SHOXs	7.0	1.0200	5.0	9.0	Normal

Childhood drug dose

The means for the childhood drug dose for all the conditions were the same as used in the base case analysis. The estimated "standard errors" attempted to express the appropriate variability of doses used in the KIGS database and also the maximum doses suggested in the BNF. These were sampled using normal distributions.

Table A9 Childhood drug dose parameters and distribution

Childhood dose	Mean	"Standard error"	Lower 95% CI	Upper 95% CI	Distribution
GHD	0.025	0.00255	0.020	0.030	Normal
TS	0.045	0.00255	0.040	0.050	Normal
PWS	0.035	0.00255	0.030	0.040	Normal
CRI	0.045	0.00255	0.040	0.050	Normal
SGA	0.035	0.00255	0.030	0.040	Normal
SHOXs	0.040	0.00255	0.040	0.050	Normal

Proportion of Males

The reported mean proportion of males for each condition was taken from the KIGS database for both the treated and untreated groups. This was consistent with the base-case analysis. The standard errors were calculated for each mean proportion of males and sampled using a normal distribution.

Table A10 Proportion of males parameters and distribution

Proportion of males	Mean	Standard error	Distribution
GHD	0.70	0.0100	Normal
TS	0.00	0.0000	Normal
PWS	0.50	0.0045	Normal
CRI	0.71	0.0040	Normal
SGA	0.596	0.0032	Normal
SHOXs	0.48	0.0019	Normal

Costs

Costs included in the PSA were those related to outpatient visits, nurse visits and monitoring tests. Drug costs were not varied in the PSA, but were included at values quoted in the BNF. Costs derived from NHS Reference Costs were sampled using estimated "standard errors". These assumed that a variation of plus or minus 25% was an appropriate confidence interval for the average reference costs. The estimated standard errors are shown in column 3 of the table below. Parameters for gamma distributions (shown in columns labelled Alpha and Beta) were derived using the means and estimated "standard errors". The simulated values were inflated to 2008/09 prices using appropriate inflation indices, as for the base case and deterministic sensitivity analyses.

Table A11 Costs parameters and distribution

Item	Mean	"Standard error"	Alpha	Beta	Distribution
Outpatient (first)	275.84	24.57	126.07	2.19	Gamma
Outpatient (subsequent)	127.97	11.40	126.07	1.02	Gamma
Specialist nurse	73.00	6.50	126.07	0.58	Gamma
District nurse	64.00	5.70	126.07	0.51	Gamma
Blood test	51.00	4.54	126.07	0.40	Gamma
X-ray	28.64	2.55	126.07	0.23	Gamma
Pituitary function test	246.50	21.95	126.07	1.96	Gamma

Appendix 13 Weight tables for males and females by age (Western Europe KIGS).

Age (Year)	SGA weight (kg)		GHD weight (kg)		PWS weight (kg)		CRI weight (kg)		TS weight (kg)
	Male	Female	Male	Female	Male	Female	Male	Female	Female
0	4.00	3.0	6.01	5.63	4.00	3.00	4.00	3.00	3.00
1	6.00	5.7	8.40	7.96	9.41	8.37	8.14	6.60	7.03
2	8.07	8.48	10.18	9.81	10.96	10.15	10.42	9.60	10.19
3	10.10	10.04	12.18	11.98	14.48	12.08	12.39	11.77	11.91
4	11.13	11.39	13.97	13.63	17.67	15.92	14.26	13.13	13.80
5	13.63	13.62	15.72	15.41	20.55	20.00	16.24	15.22	15.56
6	15.58	15.79	17.79	17.49	23.37	23.18	17.98	18.15	17.67
7	17.96	17.86	20.15	19.76	26.96	26.64	20.14	19.33	20.20
8	20.06	19.86	22.76	22.41	31.48	29.42	22.42	21.47	23.14
9	22.27	22.45	25.4	25.42	35.82	33.94	24.92	23.41	26.57
10	24.93	24.83	28.5	28.79	40.95	41.24	27.49	26.42	30.04
11	27.73	28.52	31.74	32.02	44.46	44.29	30.49	30.17	34.05
12	31.08	31.71	35.00	35.99	51.70	47.49	34.08	34.78	38.47
13	34.53	35.36	39.28	40.26	57.96	52.80	37.43	37.27	42.33
14	38.89	38.22	44.40	44.19	63.80	56.84	41.15	39.80	46.00
15	44.33	40.27	49.91	47.72	69.02	59.07	44.84	41.03	49.05
16	49.04	43.05	54.47	49.97	74.43	56.32	48.70	41.15	51.47
17	53.50	47.03	58.5	53.38	74.14	61.15	50.4	42.66	52.53