A systematic review of the routine monitoring of growth in children of primary school age to identify growth-related conditions

D Fayter, J Nixon, S Hartley, A Rithalia,
G Butler, M Rudolf, P Glasziou, M Bland,
L Stirk and M Westwood

June 2007

HTA

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:
HTA Despatch   Email: orders@hta.ac.uk
c/o Direct Mail Works Ltd   Tel: 02392 492 000
4 Oakwood Business Centre   Fax: 02392 478 555
Downley, HAVANT PO9 2NP, UK   Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

A systematic review of the routine monitoring of growth in children of primary school age to identify growth-related conditions

D Fayter,[1]* J Nixon,[1] S Hartley,[1] A Rithalia,[1]
G Butler,[2] M Rudolf,[3] P Glasziou,[4] M Bland,[5]
L Stirk[1] and M Westwood[1]

[1] Centre for Reviews and Dissemination, University of York, UK
[2] Institute of Health Sciences, University of Reading, UK
[3] Department of Community Paediatrics, University of Leeds, UK
[4] Department for Primary Health Care, University of Oxford, UK
[5] Department of Health Sciences, University of York, UK

* Corresponding author

Published June 2007

NIHR Health Technology Assessment Programme

The Health Technology Assessment (HTA) programme, now part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the costs, effectiveness and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The research findings from the HTA Programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA Programme is needs-led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, the public and consumer groups and professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA Programme then commissions the research by competitive tender.

Secondly, the HTA Programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Thirdly, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer-reviewed by a number of independent expert referees before publication in the widely read monograph series *Health Technology Assessment*.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned by the HTA Programme as project number 04/09/02. The contractual start date was in April 2005. The draft report began editorial review in April 2006 and was accepted for publication in November 2006. As the funder, by devising a commissioning brief, the HTA Programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health.

G

Abstract

A systematic review of the routine monitoring of growth in children of primary school age to identify growth-related conditions

D Fayter,[1]* J Nixon,[1] S Hartley,[1] A Rithalia,[1] G Butler,[2] M Rudolf,[3] P Glasziou,[4] M Bland,[5] L Stirk[1] and M Westwood[1]

[1] Centre for Reviews and Dissemination, University of York, UK

[2] Institute of Health Sciences, University of Reading, UK

[3] Department of Community Paediatrics, University of Leeds, UK

[4] Department for Primary Health Care, University of Oxford, UK

[5] Department of Health Sciences, University of York, UK

* Corresponding author

Objectives: To clarify the role of growth monitoring in primary school children, including obesity, and to examine issues that might impact on the effectiveness and cost-effectiveness of such programmes.

Data sources: Electronic databases were searched up to July 2005. Experts in the field were also consulted.

Review methods: Data extraction and quality assessment were performed on studies meeting the review's inclusion criteria. The performance of growth monitoring to detect disorders of stature and obesity was evaluated against National Screening Committee (NSC) criteria.

Results: In the 31 studies that were included in the review, there were no controlled trials of the impact of growth monitoring and no studies of the diagnostic accuracy of different methods for growth monitoring. Analysis of the studies that presented a 'diagnostic yield' of growth monitoring suggested that one-off screening might identify between 1:545 and 1:1793 new cases of potentially treatable conditions. Economic modelling suggested that growth monitoring is associated with health improvements [incremental cost per quality-adjusted life-year (QALY) of £9500] and indicated that monitoring was cost-effective 100% of the time over the given probability distributions for a willingness to pay threshold of £30,000 per QALY. Studies of obesity focused on the performance of body mass index against measures of body fat. A number of issues relating to human resources required for growth monitoring were identified, but data on attitudes to growth monitoring were extremely sparse. Preliminary findings from economic modelling suggested that primary prevention may be the most cost-effective approach to obesity management, but the model incorporated a great deal of uncertainty.

Conclusions: This review has indicated the potential utility and cost-effectiveness of growth monitoring in terms of increased detection of stature-related disorders. It has also pointed strongly to the need for further research. Growth monitoring does not currently meet all NSC criteria. However, it is questionable whether some of these criteria can be meaningfully applied to growth monitoring given that short stature is not a disease in itself, but is used as a marker for a range of pathologies and as an indicator of general health status. Identification of effective interventions for the treatment of obesity is likely to be considered a prerequisite to any move from monitoring to a screening programme designed to identify individual overweight and obese children. Similarly, further long-term studies of the predictors of obesity-related co-morbidities in adulthood are warranted. A cluster randomised trial comparing growth monitoring strategies with no growth monitoring in the general population would most reliably determine the clinical effectiveness of growth monitoring. Studies of diagnostic accuracy, alongside evidence of effective treatment strategies, could provide an alternative approach. In this context, careful consideration would need to be given to target conditions and intervention thresholds. Diagnostic accuracy studies would require long-term follow-up of both short and normal children to determine sensitivity and specificity of growth monitoring.

Contents

Glossary 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. In some cases, usage differs in the literature, but the term has a constant meaning throughout this review.

Glossary

Accuracy   The degree of agreement between a measured value and the true value (see Precision).

Auxologist   A specialist in the study of human growth.

Clinical effectiveness   The extent to which a medical intervention achieves health improvements in real practice settings.

Cluster randomised controlled trial Randomised controlled trial in which randomisation is by group (e.g. clinical practice, hospital) rather than by individual participant.

Diagnostic case–control study   A study in which a group of people known to have the target condition (cases) and a group known to be free of the target condition (controls) receive the diagnostic intervention and are then compared with the established diagnosis.

Diagnostic cohort study   A study in which a group of people, representative of those who would receive the diagnostic intervention in clinical practice, receive both the diagnostic intervention and the reference standard.

Diagnostic yield   The diagnostic yield of a monitoring programme is the number of new cases detected out of the number of children measured. The incremental yield is the number of new cases that would not have been identified without the monitoring programme.

External validity   The extent to which the effects observed in a study can be expected to apply in routine clinical practice, i.e. to people who did not participate in the study.

Growth monitoring   The process of checking, observing or keeping track of height or weight measurement for a specific period of time or at specified intervals.

Incidence   The number of new cases of a specific condition occurring during a certain period in a specified population.

Index test   The diagnostic test whose performance is being measured.

Internal validity   The degree to which a result (of a measurement or study) is likely to be true and free of bias (systematic errors).

Internally derived threshold   The diagnostic/screening threshold (e.g. 95th centile) is calculated using data from the study population rather than from an external source.

Precision   A measure of how consistently the result is determined by repeated determinations without reference to any 'true' value (see Accuracy).

Prevalence   The proportion of people in a population who have a given disease or attribute at a given point in time.

Reference standard   Considered to be the best available method for establishing the presence or absence of the condition of interest.

Screening   A health service in which members of a defined population, who do not necessarily perceive they are at risk of a disease or its complications, are asked a question or offered a test, to identify those individuals who are more likely to be helped than harmed by further tests or treatment.

Measures of diagnostic test performance

Below is a summary of the measures of diagnostic test performance used in the review, and how these are calculated.

	Present	Absent
Test result +	*a*	*b*
−	*c*	*d*

True positives (TP)   Correct positive test results: *a* – number of diseased persons with a positive test result.

True negatives (TN)   Correct negative test results: *d* – number of non-diseased persons with a negative test result.

False positives (FP)   Incorrect positive test results: *b* – number of non-diseased persons with a positive test result.

False negatives (FN)   Incorrect negative test results: *c* – number of diseased persons with a negative test result.

Sensitivity $a/(a + c)$ – Proportion of people with the target disorder who have a positive test result.

Specificity $d/(b + d)$ – Proportion of people without the target disorder who have a negative test result.

List of abbreviations

%BF	percentage body fat	DEXA	dual-energy X-ray absorptiometry
BIA	bioelectrical impedance analysis	EVPI	expected value of perfect information
BMI	body mass index		
CBA	cost–benefit analysis	GH	growth hormone
CDC	Centers for Disease Control and Prevention	GHD	growth hormone deficiency
		HEED	Health Economic Evaluation Database
CEA	cost-effectiveness analysis		
CEAC	cost-effectiveness acceptability curve	HSDS / HtSDS	height standard deviation score
CI	confidence interval	ICER	incremental cost-effectiveness ratio
CRD	Centre for Reviews and Dissemination	IGF	insulin-like growth factor
CRF	chronic renal failure	IHQL	Index of Health-related Quality of Life
CRI	chronic renal impairment		
CUA	cost–utility analysis	IOTF	International Obesity Task Force

List of abbreviations *continued*

ISS	idiopathic short stature	PPP	Purchasing Power Parities
JH	juvenile hypothyroidism	PSSRU	Personal Social Services Resource Unit
MPH	mid-parental height	PWS	Prader–Willi Syndrome
MPHD	multiple pituitary hormone deficiency	QALY	quality-adjusted life-year
NCHS	National Center for Health Statistics	QoL	quality of life
NHANES	National Health and Nutrition Examination Study	QUADAS	Quality Assessment of Diagnostic Accuracy Studies
NHS EED	NHS Economic Evaluation Database	RCT	randomised controlled trial
		SD	standard deviation
NICE	National Institute for Health and Clinical Excellence	SIGLE	System for Information in Grey Literature
NSC	National Screening Committee	StHA	Strategic Health Authority
NTIS	National Technical Information Service	TS	Turner's syndrome
		TSF	triceps skinfold
PCT	Primary Care Trust	WTP	willingness to pay

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or at the end of the table.

Executive summary

Background

Assessment of children's height and weight is well established as an indicator of their general health and well-being. Monitoring height and weight to identify growth disorders, including obesity, might be a useful exercise. The current role of growth monitoring is unclear and uncertainties exist as to the most appropriate age(s) to measure and the measurement strategies to adopt.

Objectives

The aim of this review was to clarify the role of growth monitoring, including obesity, and to examine issues that might impact on the effectiveness and cost-effectiveness of such programmes. The objectives were to:

- determine the detection rate, age at diagnosis and route to diagnosis of the target growth-related conditions in the UK population
- determine the clinical effectiveness of growth monitoring in terms of the age of diagnosis and management/outcome of children
- determine the diagnostic performance of growth monitoring strategies for the identification of growth related conditions
- evaluate any evidence of human resource requirement of growth monitoring programmes
- evaluate any evidence of attitudes of children, parents and health care professionals to growth monitoring
- determine the likely cost-effectiveness of routine growth monitoring.

Methods

Data sources

Studies were identified through extensive searches of electronic databases up to July 2005, handsearching of journals, scanning reference lists of included papers and consultation with experts in the field.

Study selection

Two reviewers independently screened titles and abstracts for relevance. Full papers of potentially relevant studies were assessed for inclusion by one reviewer and checked by a second. Published and unpublished studies in any language were eligible for inclusion.

Inclusion criteria

Separate inclusion criteria, relating to study design, participant characteristics and outcome measures, were derived for each objective.

Data extraction

Data extraction and quality assessment were performed using standardised forms. The quality of the included studies was evaluated using specially designed or standard checklists according to study type. Data extraction was checked by a second reviewer.

Data synthesis

Data were analysed separately for each of the phases of the review. Results were presented in tables and synthesised narratively. The performance of growth monitoring to detect disorders of stature and obesity was evaluated against National Screening Committee (NSC) criteria.

Results

Monitoring for stature-related disorders

Thirty-one studies were included in the review. There were no controlled trials of the impact of growth monitoring and no studies of the diagnostic accuracy of different methods for growth monitoring. Analysis of the studies that presented a 'diagnostic yield' of growth monitoring suggested that one-off screening might identify between 1:545 and 1:1793 new cases of potentially treatable conditions. Economic modelling suggested that growth monitoring is associated with health improvements [incremental cost per quality-adjusted life-year (QALY) of £9500] and indicated that monitoring was cost-effective 100% of the time over the given distributions for a willingness to pay threshold of £30,000 per QALY.

Monitoring for obesity

Studies of obesity focused on the performance of body mass index against measures of body fat. A number of issues relating to human resources of growth monitoring were identified, but data on attitudes to growth monitoring were extremely sparse. Preliminary findings from economic modelling suggested that primary prevention may be the most cost-effective approach to obesity management, but the model incorporates a great deal of uncertainty.

Conclusions

This review has indicated the potential utility and cost-effectiveness of growth monitoring in terms of increased detection of stature-related disorders. It has also pointed strongly to the need for further research.

Implications for policy and practice
Monitoring for stature-related disorders
Growth monitoring does not currently meet all NSC criteria. However, it is questionable whether some of these criteria can be meaningfully applied to growth monitoring given that short stature is not a disease in itself, but is used as a marker for a range of pathologies and as an indicator of general health status. There is a need to consider the extent to which it is appropriate to evaluate growth monitoring against NSC criteria. Those considering implementing growth monitoring programmes may need to consider whether the potential for earlier detection of stature-related disorders outweighs the lack of information on other relevant NSC criteria. It may be useful to consider the potential benefits of growth monitoring in the context of overall child health and the potential to detect other important, treatable disorders.

Monitoring for obesity
Identification of effective interventions for the treatment of obesity is likely to be considered a prerequisite to any monitoring programme designed to identify individual overweight and obese children. Similarly, further long-term studies of the predictors of obesity-related co-morbidities in adulthood are warranted; at present it is unclear how the target population of any monitoring programme should be defined. There is a need to consider these issues, and also the lack of data on the benefits and harms of monitoring,

before moving away from the current population-based approach to obesity monitoring.

Recommendations for research
Monitoring for stature-related disorders
The primary consideration for future research on growth monitoring is the establishment of clinical and cost-effectiveness. The clinical effectiveness of growth monitoring would be most reliably determined by a cluster randomised trial comparing growth monitoring strategies with no growth monitoring in the general population. Studies of diagnostic accuracy, alongside evidence of effective treatment strategies, could provide an alternative approach. In this context, careful consideration would need to be given to target conditions and intervention thresholds. Diagnostic accuracy studies would require long-term follow-up of both short and normal children to determine sensitivity and specificity of growth monitoring. Qualitative research in the following areas would provide additional information pertinent to NSC criteria: attitudes of children, parents and health professionals to growth monitoring; system barriers to implementation; methods of management and quality assurance; training and staffing needs; optimisation of coverage; and the effects of participant information.

Monitoring for obesity
In the absence of evidence of effective interventions, the value of monitoring children in order to identify those who are overweight or obese will remain questionable. Research to identify weight reduction strategies that are effective in children is therefore a priority. Of equal priority is research on the effectiveness of primary prevention as an alternative or complementary strategy. Long-term epidemiological studies to establish which children are at most risk of adverse outcomes of obesity in adulthood is also a high priority; these studies define the target population for any monitoring programme aiming to identify and treat children. Before implementation of any such monitoring programme, funding for UK research into the benefits and harms of monitoring for and treating obesity, including long-term outcomes, would be a priority. Should effective treatments for obesity be identified, the effectiveness of monitoring for obesity would be most reliably determined by a cluster randomised trial comparing monitoring strategies with no monitoring and with alternative preventative strategies.

Chapter 1

Background

Potential benefits of monitoring height and weight

Assessment of a child's height and weight as an indicator of health and well-being is well established and has been incorporated into paediatric practices in both developed and developing countries. In 1998 in the UK, a multidisciplinary group consisting of paediatricians, endocrinologists, public health professionals, GPs and nurses met to develop a consensus on growth monitoring (known as the Coventry Consensus).[1] The group established that the potential benefits of growth monitoring were the identification of treatable chronic disorders or diseases in apparently normal children, the provision of reassurance to parents, the provision of data to monitor children's health from a public health perspective and the provision of data for use in epidemiological research.[1]

The potential benefits of monitoring for the detection of childhood overweight and obesity need to be determined in the light of increasing prevalence in the developed world. In the UK, the prevalence of overweight or obese children aged 2–10 years rose from 22.7 to 27.7% between 1995 and 2003. The increase in obesity was most significant in children aged 8–10 years, rising from 11.2 to 16.5%.[2] A recent UK House of Commons Health Committee report has expressed particular concerns about the long-term health consequences of obesity in children.[3] The importance placed upon childhood obesity, in public health terms, is reflected by the Government's commitment to stemming its increase by 2010.[4]

Growth-related conditions

Routine monitoring of growth does not aim to detect a single pathology. There are, in fact, a number of conditions that may lead to decreased or increased growth rate and/or short (*Table 1*) or tall stature (*Table 2*), which potentially could be detected through growth monitoring. Conditions in which stature outside the normal range is often the only or most significant presenting feature are growth hormone deficiency (GHD) and Turner's

syndrome (TS), and it is these conditions which are used to justify growth screening in childhood.[1] However, there are a number of other conditions for which some new cases may be identified as a consequence of growth monitoring. Short stature may result from hypothyroidism, psychosocial deprivation, intrauterine growth retardation or other chronic but undetected illness. Tall stature is a feature of a number of syndromes, such as Marfan syndrome and Klinefelter syndrome, and may also be a sign of treatable endocrine disorders. Early detection and diagnosis of organic causes of abnormal growth are important to ensure that final adult height within a normal range is achieved.[5–8] Where possible, treatment will be provided for the underlying condition. In cases where the condition is not treatable, early diagnosis can allow for discussion with the family and child and counselling to minimise any psychological distress. However, children with a treatable cause of abnormal growth are frequently diagnosed at, or treatment is initiated at, a late age.[5,9–12]

Growth hormone deficiency

Growth hormone (GH) is produced by the pituitary gland and stimulates the growth of bones throughout childhood and adolescence. An absence or insufficient production of growth hormone therefore leads to slowed growth and results in short stature. GHD can be present at birth or may be acquired due to disease or injury. It may occur as part of a multiple pituitary hormone deficiency (MPHD). Estimations of the prevalence of GHD range from one in 3500 to one in 7000.[13] A UK study in 1977 suggested a prevalence among 9-year-olds of approximately one in 4000.[7] If left untreated, GHD results in a greatly reduced adult height, an average of 4.7 standard deviations (SDs) below the mean.[14] At the age of 5 years, most cases of children with GHD would be expected to be below the 0.4th percentile for height.[15] GHD can be successfully treated with GH. Recent studies have shown males attaining a mean final height –0.9 SD below the mean (range –2.6 to –0.6) and females –1.21 (range –2.8 to –0.4), and optimisation of the treatment regime should further improve this in the future.[14] The final height achieved appears to be influenced by genetic potential and also the

2

TABLE 1 Conditions of interest associated with short stature

Condition	Overview	Prevalence	Growth as a marker	Age of diagnosis	Treatment	Implications for adulthood
GHD	Insufficient production of GH by the pituitary gland. May be isolated, part of MPHD, or consequence of another disease	Around 1:4000 births	Short stature is the most obvious feature of isolated GHD. It is estimated that 80% of 5-year-old cases would fall under the 0.4th centile for height	NR	Treated with GH. Mean final height SD is –0.9 for males and –1.21 for females	Untreated GHD would result in very short stature: –4.7 SDs (range –3.9 to –6.1)
TS	Chromosomal disorder leading to short stature and infertility. Also associated with characteristic physical features	Around 1:2500 live female births	Short stature may be the only obvious feature in childhood. It is estimated that 50% of 5-year-old cases would fall below the 0.4th centile for height	40% may be diagnosed before the age of 1. Diagnosis during childhood (40%) and adolescence (20%) is usually due to growth retardation	GH treatment can allow TS patients to achieve height in the normal range (within 1 SD), especially if started early before the use of oestrogen to induce puberty	Untreated, height would be below 2 SDs. As adults are infertile, early diagnosis is also important to allow psychological preparation for this issue
Juvenile hypothyroidism	Acquired during childhood (as opposed to congenital hypothyroidism, which is screened for at birth), insufficient thyroid hormone leads to a range of symptoms including reduced growth rate	1:1450 in under 22s	Short stature may be presenting feature	NR	Treated with thyroxine, which leads to catch-up growth	
Psychosocial short stature	Psychosocial stress suppresses GH secretion	No data	Growth rate slows, which may lead to short stature	NR	Moving the child away from the stressful environment may bring growth back to normal	

NR, not reported.

TABLE 2 *Conditions of interest associated with tall stature*

Condition	Overview	Prevalence	Growth as a marker	Age of diagnosis	Treatment	Implications for adulthood
Marfan syndrome	Genetic disorder of connective tissue. Up to 30% of cases are sporadic, with no family history	1 in 3000–5000	Tall stature is one of a number of characteristic features in children with Marfan syndrome	NR	Cardiovascular complications are common, and these need to be monitored and managed	Lifespan may be reduced due to cardiovascular complications. Early diagnosis is therefore important to monitor these
Klinefelter syndrome	Chromosomal disorder leading to a range of symptoms, the most common of which is infertility	1 in 500–1000 live male births	Tall stature is one of a number of characteristic features in boys with Klinefelter syndrome	Diagnosed from 18 months. School-age children are often diagnosed due to learning disabilities, adults due to infertility	Testosterone therapy can improve masculinisation	Adults are infertile. Although this cannot be treated, early diagnosis would allow psychological preparation for this issue
Premature sexual maturation	Development of secondary sexual characteristics under 8 years in girls and 9 years in boys	No data	Children are generally tall	NR	Drugs can be used to suppress the onset of puberty, allowing a more normal course of development and final height	If untreated, adult height tends to be shorter than average

NR, not reported.

height at which treatment starts and the height at onset of puberty.[16] Therefore, early diagnosis of GHD to maximise final height is likely to be important.

Isolated GHD of unknown cause is most likely to be picked up as a new case in childhood growth screening as MPHD would be likely to present before the age of 2 years with a collection of symptoms and GHD resulting from injury or other disease should be picked up at specialist follow-up.[1] It has been predicted that growth screening at the 0.4th centile of 100,000 5-year-olds would pick up approximately 20 out of the expected 25 cases of isolated GHD, some of whom may already have been diagnosed.[15]

Turner's syndrome
TS is a chromosomal abnormality affecting approximately 1 in 2500 live female births. It leads to short stature and infertility, and may be associated with a number of characteristic physical features. The route to diagnosis varies, with 30–50% of cases diagnosed prenatally, at birth or in the first year, mainly due to characteristic dysmorphic features, and the remainder in childhood (30–50%) or adolescence (20%), with short stature as the key to diagnosis in the majority of these cases.[11,17] The average height at adulthood if TS is left untreated is 143–147 cm, much more than 2 SD below the average normal female height.[14] At 5 years of age, approximately 50% of girls with TS would be expected to fall below the 0.4th percentile for height.[15] GH can be used to increase height, to within 1 SD of the normal mean.[14] The duration of GH treatment prior to induction of puberty with oestrogen appears to be an important predictor of final height.[14] Therefore, it is important to diagnose TS as early as possible. It has been predicted that screening 50,000 5-year-old girls at the 0.4th percentile would identify approximately nine out of an expected 16 undiagnosed cases (assuming that four were diagnosed at birth).[15]

Juvenile hypothyroidism (JH)
Hypothyroidism occurs when the thyroid gland fails to produce sufficient thyroxine, leading to a number of complications, including impaired physical and mental function or development. This condition can be present at birth (congenital hypothyroidism) or can be acquired later in life. In a UK analysis of thyroxine prescriptions, the prevalence of acquired primary hypothyroidism in people under the age of 22 years was estimated to be 1:1450.[18] In the UK, newborns are screened for congenital hypothyroidism, but if the condition

develops later in childhood one presenting feature may be short stature. Treatment with thyroxine resolves the symptoms and also leads to catch-up growth, although expected adult height may only be achieved if treatment is started early.[16]

Psychosocial short stature
The link between social deprivation and short stature is well established.[19] In cases of severe psychological stress in childhood, GH secretion may be suppressed, leading to growth arrest and short stature.[19] In these cases, moving the child away from the stressful environment can lead to rapid catch-up growth.[20] Early detection of these cases is important, to avoid exposure to further stress and to enable child protection procedures to be put in place. It is unclear, however, what additional proportion of these children would be picked up through growth screening programmes rather than through usual clinical care, or indeed how many such cases exist.[7]

Other conditions associated with short stature
There are a number of other conditions associated with short stature, including skeletal abnormalities (e.g. achondroplasia, hypochondroplasia), multi-symptomatic syndromes [e.g. Down's syndrome, Noonan syndrome, Prader–Willi syndrome (PWS)] chronic conditions (chronic renal disease, coeliac disease, Crohn's disease) and intrauterine growth retardation. Almost all cases with these chromosome abnormalities would present with characteristic features other than short stature, and/or before school age. Therefore, it is considered that childhood growth screening is unlikely to contribute greatly to the diagnoses of these conditions. However, other systemic conditions may not present with obvious physical features. Of particular interest is the well-documented occurrence of childhood cases of coeliac disease that present with slowed growth and short stature in the absence of the expected gastrointestinal complaints.[7] Coeliac disease can be readily treated with a gluten-free diet, and therefore its early detection has obvious benefit. Conditions for which GH treatment is licensed in the UK are also of particular interest; these include chronic renal disease and PWS.[21]

Conditions associated with tall stature
There are a number of conditions for which increased growth rate or tall stature may be a feature.

Marfan syndrome is a genetic disorder of connective tissue affecting about one in 5000

individuals. Up to 30% of cases are sporadic, with no family history. There are many symptoms, among which is a tall, lanky frame in childhood. The average height of children with Marfan syndrome exceeds the 95th centile from the age of 3 years.[22] The condition was considered by the Coventry Consensus as a possible justification for height screening at the 99.6th centile, but it was concluded that an insufficient percentage of children with Marfan syndrome would fall above the normal height range.[15] Although it is important to diagnose Marfan syndrome early in order to manage the cardiovascular complications associated with it, treatment for height reduction is not usually necessary as adult height tends not to be excessive.[23]

Klinefelter syndrome is a chromosomal condition affecting between one in 500 and one in 1000 live male births.[23] In childhood, boys with Klinefelter syndrome are generally taller than average, but final height tends to overlap with the upper part of the normal range. Adults are infertile. Testosterone treatment can be used to manage height and to improve masculinisation.

Excessive height may also have an endocrine cause. Children with precocious sexual maturation, that is, development of secondary sexual characteristics before the age of 8 years in girls and 9 years in boys, are taller than their peers, although their final adult height will generally be short due to early skeletal maturation.[24] If diagnosed early enough, the condition can be treated with gonadotrophin-releasing hormone analogues to delay puberty.[23]

Other rare syndromes, such as Sotos syndrome and Beckwith–Wiedemann syndrome, are associated with excessive height in childhood, but have characteristic dysmorphic features which are more likely to prompt diagnosis.[23] Other rare, treatable, endocrine causes of tall stature in childhood include pituitary gigantism and hyperthyroidism.

Normal short stature
In screening for short stature, most children identified would be classed as normal variants, with no underlying pathology or psychosocial issues causing their reduced height. Some children may have constitutional delay in growth and/or puberty, where they are growing at a normal rate but their bone age is younger than their actual age, meaning that they would eventually achieve a normal height, but at a later date than their peers. Their parents have usually experienced similar

growth patterns, making them readily identifiable. If the child's parents are also short, then the child may be considered to have genetic or familial short stature. In some cases, however, no explanation is found for the child's short stature, and this is termed idiopathic short stature.

Although GH is not licensed in the UK for use in children with normal short stature, it has been estimated that approximately 1.2% of these children receive GH treatment.[21] GH treatment has been found to add between 2 and 7 cm to the final height of children with normal short stature, bringing it just within 2 SDs of the mean normal height.[21] Although it has been suggested that untreated short children suffer psychologically due to their short stature, the evidence for this remains unclear.[25,26]

Normal tall stature
A height screen would also identify the normal tall variants, for whom there is no underlying medical cause, and as with short stature, tall stature may be constitutional. Normal tall stature is not generally viewed as a disadvantage, but children seeking treatment can be given sex steroids to accelerate skeletal maturity.[23]

Conditions associated with obesity
A growth screening programme considering obesity may identify some children with an underlying medical cause for their obesity in addition to those with nutritional obesity.

Medical causes of obesity
Medical conditions associated with obesity include PWS, Bardet–Biedl syndrome, hypothyroidism and Cushing's disease, all of which are rare and occur in conjunction with slowed growth and/or short stature. This contrasts with nutritional obesity, which is more often associated with tall stature. Endocrine causes of obesity, such as hypothyroidism and Cushing's disease, can be effectively treated.

Nutritional obesity
Most obese children do not have an underlying medical condition. Childhood obesity is a complex condition with many possible contributing factors ranging from food consumption and activity to policies relating to the advertising and the labelling of foods.[27] The main contributing factor to the observed marked increase in childhood obesity in recent years is thought to be societal changes relating to nutrition and an increased sedentary lifestyle of children.[28] Evidence suggests variations in levels of obesity across different socio-

economic groups, geographical locations, and gender.[2,29] Co-morbidities of obesity in childhood include type 2 diabetes, hypertension, dyslipidaemia, emotional and behavioural problems, asthma and sleep apnoea.[28–30] Obesity in childhood is highly likely to persist into adulthood and the cardiovascular effects of childhood obesity have been linked with morbidity and mortality in adulthood.[30] Although it appears important to identify children who are overweight or obese as early as possible, evidence on the effectiveness of measures to prevent and treat childhood obesity is limited. A Cochrane systematic review of treatments for obese children found limited success for some intensive lifestyle programmes, interventions involving diet, exercise and behavioural components; benefit was often demonstrated only for the short term.[31] Pharmacological agents may also be of some benefit in treating certain obese children, although most are associated with side-effects, and none are licensed for use in children in the UK.[32] Prevention is an alternative strategy to tackle the problem of childhood nutritional obesity. A Cochrane systematic review of interventions to prevent childhood obesity, mainly school-based, found that nearly all were successful in improving diet or physical activity, and although improvements in measures of obesity were generally not seen, the majority of studies had a duration of less than 1 year.[33] This is an area of ongoing concern and research.

Growth monitoring programmes

Current status

The Child Health Subcommittee of the UK National Screening Committee (NSC) recommended in 2004 that a single height and weight measurement should be taken at or around time of school entry and that the 0.4th centile for height should be used to initiate referral.[34] However, historically, routine growth monitoring practices have varied across the UK and continue to do so. There are differences in the policy, equipment and growth charts used.[35] Looking outside the UK, variations also exist across different countries.[36,37]

In 2004, an international group of physicians produced the Childhood Obesity Consensus statement,[32] which recommended screening all children where resources permit, directing overweight children [body mass index (BMI) > 85th centile] for weight management, counselling and testing obese children (BMI > 95th centile)

for co-morbidities, with specialist treatment if indicated. However, monitoring programmes for the detection of overweight and obese children are not currently in general use in the UK.

Routes to diagnosis

Community growth monitoring programmes do not identify all cases of growth-related conditions. The proportion of clinic-referred growth problems detected by routine growth monitoring in Leeds is approximately 25%[9] and in Manchester is approximately 36% (Shah SA, University of Manchester: personal communication, 2000). The remaining cases are detected using other routes of referral such as concerns raised by parents or health professionals during routine appointments.[9,38]

Measurement issues

Height measurement

Height varies naturally within any given population and an individual child's height is measured relative to a population norm for age and sex.[21] In addition, height also varies within specific subgroups of a population, for example across different ethnic groups or geographical locations.[39] Given that there is always a distribution of heights, a small proportion of children will be significantly taller or shorter compared with other children of the same age and sex, that is, they will fall outside the defined 'normal' range.[21] In the majority of these children there is no apparent medical condition. However, a small number of children will have an underlying pathology that may be treatable to allow normal height to be achieved, as described in the previous section.

Considerable debate exists regarding whether a single height measurement at school entry is the best way to identify growth-related disorders, or whether growth should be monitored over time. Single height measurements give an absolute assessment of stature and can only be used to identify children whose growth has been persistently, sufficiently slow compared with others of the same age and sex to result in a significant absolute difference at the time of measurement. Repeated height measurements over time allow for calculation of a growth rate and can be used to define abnormal growth in terms of a crossing of the height centiles, that is, they identify abnormality through the pattern of growth within the individual.[9] This approach may intuitively appear more useful. However, studies showing imprecision in the measurement of height suggest that short-term height velocity may not be

adequate to identify abnormal growth in routine growth monitoring[40,41] and there is an inherent normal cyclicity to height velocity which is not evident from a limited number of measurements.

The performance of height measurement may be improved if a child's height is corrected for parental height; some children may be incorrectly referred for further investigation if their height potential has not been considered. The potential or target height of a child can be calculated from the mid-parental height (MPH), or if the height of both parents is not available the height of one parent or sibling, and adjusting for the child's sex.[42] However, it has been argued that it is only appropriate to consider MPH when both parents are of normal stature and the calculation may be misleading in the assessment of short children.[43] It is also possible that children with underlying pathology, which may be familial or otherwise, may also have at least one short parent.[44]

A potential source of variation that can impact on the validity of all growth monitoring is the instrument used to measure height. The use of correctly installed and calibrated instruments should be mandatory in order to minimise instrument-generated error.[41] Several instruments are available to measure the height of children; these include stadiometers, Microtoises, rulers and wall charts.[41] For clinic or school use the Leicester Height Measure or Minimeter are considered to be reliable, provided that they are accurately installed.[45] At least two measures are needed to obtain an accurate measurement.[46]

Body fat measurement
Overweight and obesity are caused by an excessive accumulation of body fat. Percentage body fat (%BF) has therefore been regarded as the reference standard for definition of obesity. Several techniques are available to determine the amount of body fat present in an individual, and there is variation in their precision, costs and availability for use in routine practice. Body composition techniques such as dual-energy X-ray absorptiometry (DEXA), imaging, hydrodensitometry and bioelectrical impedance analysis (BIA) tend to be used only in research settings or for the validation of other measurement methods.

For the purposes of routine clinical practice, the method to assess overweight and obesity in children needs to be simple, relatively inexpensive and able to identify rapidly children with excess fat who are at risk of associated morbidity.[32,47] The most widely used and recommended measurement is the BMI, which describes a relative weight for height. BMI can only give an indirect estimate of total body fat and cannot provide a reliable prediction of outcome.[32] Since it is a ratio of height to weight, it may be unsuitable for assessing children who are particularly short or tall for their age, or for short, muscular children; there is a potential to obtain misleading results in such children.

Growth charts
There is no standard 'cut-off' used for defining short or tall stature. Diagnosis of abnormal growth is usually based on a child's height measurement outlying recommended percentile points on a growth chart.[48] Growth charts used in current or recent practice have been derived using data collected in large population samples of normal healthy children.[49] Ideally, the sample used to develop such reference charts should be of sufficient size to provide data representative of the population in which it will be applied.[41] Many countries have developed growth reference charts, that are specific to their populations, for use in routine clinical practice.[50]

Prior to the 1990s, the Tanner and Whitehouse charts were widely used to define stature in the UK. These charts were developed in the 1960s using data primarily from children living in London. However, increasing secular trends towards taller children, and concerns regarding the generalisability of the chart to other geographical locations, led to the development of new reference curves.[51] The new charts, known as the UK 1990 charts, were constructed using measurements from over 25,000 children from seven different nationally representative data sets.[51] Using these charts, a height below the 0.4th centile or 2nd centile has been recommended to define short stature and a height above the 99.6th or 98th centile is used to define tall stature in need of further investigation.[1] These charts have demonstrated high validity in comparison with earlier charts[52] and are widely acknowledged as being the most appropriate for use in the UK.[49,51] However, they do not include specific centiles for ethnic communities; this should be considered in the interpretation of the centiles.[39]

Defining childhood obesity
BMI varies from childhood through to adulthood and is different for boys and girls.[47] The definition used for overweight and obesity is determined with comparison with a population reference using appropriate cut-off values.[47] The UK 1990 charts

are recommended for use in the UK and provide age and sex classifications for overweight above the 91st centile and obesity above the 98th centile. More recently, the International Obesity Task Force (IOTF) has developed an international classification system based on data collected from six countries (Brazil, Hong Kong, The Netherlands, Singapore, Great Britain and the USA).[53] The IOTF defined obesity in relation to adult BMI definitions.

Referral issues

The threshold for referral used in any growth monitoring programme needs to have a sensitivity sufficiently high to ensure that an acceptable proportion of children with a growth-related disorder are detected. Conversely, the specificity of the chosen threshold will determine the number of children without a growth-related disorder who are referred for further investigation. There is a clear trade-off between these: higher sensitivity will mean lower specificity (and hence more referrals); higher specificity will mean lower sensitivity (and hence more missed cases). Current knowledge of the performance of referral criteria used for growth monitoring is incomplete and there are variations in the criteria used or recommended.[9,36,54] The performance of different referral criteria for use in growth monitoring has recently been assessed in a study of Dutch children with TS.[36] The study found that referral criteria based on an absolute height measurement did not perform well compared with criteria that adjusted for parental height or the use of height velocity. In the UK, the current recommendations for referral are based on a single measurement of height at school entry.[55] Others have suggested that, in order to minimise the number of unnecessary referrals and provide a more efficient use of resources, the criteria should also consider the crossing of height centiles and parental heights,[9] or the use of a height velocity for some children and a direct referral from a single screen for children with more severe growth retardation.[56,57]

Practical issues

Growth monitoring may appear easy to undertake; the measurements required are relatively simple and quick. However, the practical and resource requirements for delivering a programme consistently, and the need to ensure that appropriate referral criteria and procedures for further investigation are implemented, mean that it is difficult to do well.

The introduction of a routine growth monitoring programme has considerable organisational and resource requirements. The direct costs include those associated with equipment and staff involved in the measurement and recording of stature and weight and the costs associated with subsequent referral for further investigation. Investment in staff and training is also needed to improve the consistency and appropriateness of referral patterns[58] and to ensure that maximum coverage is achieved.[59]

The validity of growth data depends on the accuracy and reproducibility (precision) of the measurements. Variation in measurement technique and inter-observer differences can be minimised by appropriate staff training and monitoring to ensure competence in technique. Where possible, the individual child should be measured by the same observer using the same instrument. Ideally, the observer should not be aware of previous height data. However, approximately 90% of the variation comes from children themselves and is therefore largely unavoidable; consideration should be given to the timing of measurement to allow for diurnal variation.[41]

In addition to staff training and monitoring to improve the accuracy and precision of measurements, routine growth monitoring programmes need to ensure the availability of suitable equipment and facilities. Variations in the type of instrument used to measure height and weight and the frequency of calibration are evident in the published literature and suggest the need to adopt a uniform policy on the equipment used.[60] The growth charts used need to be appropriate for the population of interest and staff need to be trained in the plotting and the interpretation of height measurements. Documented, standard operating procedures are needed to identify children requiring referral for further investigation. The chosen criteria for referral need to have appropriate sensitivity and specificity, to ensure that an acceptable proportion of children with a growth-related disorder are detected, and to minimise the numbers of children without a growth-related disorder who are referred for further investigation.

It can be concluded from the previous sections that monitoring a child's height and weight might be a useful exercise. A number of growth-related conditions might be identified for which treatment and/or counselling can be provided. However, the current role of growth monitoring is unclear, in particular when to measure height and determining the most appropriate referral

mechanism. It can also be noted that growth monitoring is difficult to do well and a number of practical issues have already been identified. The aim of this review was to elucidate the role of growth monitoring including obesity and to examine issues that might impact on the effectiveness and cost-effectiveness of such programmes.

Chapter 2
Research questions

Aim of the project

The aim was to determine the clinical impact and cost-effectiveness of routinely monitoring growth in children between the ages of 4 and 11 years in order to identify growth-related conditions, including obesity.

Objectives

The objectives were to:

- determine the detection rate, age at diagnosis and route to diagnosis of the target growth-related conditions in the UK population

- determine the clinical effectiveness of growth monitoring in terms of the age of diagnosis and management/outcome of children
- determine the diagnostic performance of growth monitoring strategies for the identification of growth-related conditions
- evaluate any evidence of human resource requirement of growth monitoring programmes
- evaluate any evidence of attitudes of children, parents and healthcare professionals to growth monitoring
- determine the likely cost-effectiveness of routine growth monitoring.

11

Chapter 3
Review methods

An advisory panel was established. In addition to providing subject-specific input during the review, members of the panel were invited to offer comment on the protocol and draft report. Details of advisory panel members can be found in Appendix 1. The systematic review was undertaken in accordance with the Centre for Reviews and Dissemination (CRD) guidelines for undertaking systematic reviews and other published guidelines.[61] Details of protocol amendments are given in Appendix 2.

Search strategy

A database of published and unpublished literature was established from systematic searches of electronic sources, reference lists of retrieved studies and extensive consultation with experts and professionals actively involved in the field.

Searches were originally designed to retrieve references on growth screening where specific target conditions and general search terms such as 'growth disorders' were used. However, this focus was found to be over-restrictive, and the search was therefore widened to include all references on child growth screening.

Studies were identified by searching major medical databases such as MEDLINE, EMBASE, BIOSIS, Science Citation Index, LILACS and Pascal from inception to July 2005 (see Appendix 3 for details of the search strategy).

In addition, information on studies in progress, unpublished research and research reported in the grey literature was sought from a range of relevant databases, including Inside Conferences, Systems for Information in Grey Literature (SIGLE), Dissertation Abstracts and the National Technical Information Service (NTIS).

Attempts to identify further studies were made by contacting clinical experts and examining the reference lists of all full publication articles retrieved. Unpublished information on current practice and audit data were sought by directly contacting all Primary Care Trust (PCT) leads in child health/community paediatrics and all

Strategic Health Authority (StHA) leads in child health services in England and Wales.

References on developing countries were excluded, but no other limits of date, language, study design or country of publication were applied. Animal studies were excluded from the search results where possible. The results of the searches were imported into Endnote version six bibliographic management software and deduplicated.

Searches for economic evaluations were undertaken using the search strategy detailed above. This search had no study design filter so economic studies were identified at the same time as the studies were assessed for the reviews. In addition, searches of the NHS Economic Evaluation database (NHS EED) and OHE Health Economic Evaluation Database (HEED) were undertaken, along with a search of the Economics Working Papers archive. To help inform the economic modelling, additional searches were undertaken to retrieve records on quality of life in children with the target conditions assessed in this review and oxandrolone in the treatment of TS, along with statistical sources and other sources of relevant information.

Study selection

Systematic review of effectiveness

Two reviewers screened titles and abstracts for relevance independently; any disagreements were resolved by consensus. Full papers of potentially relevant studies were obtained and assessed for inclusion by one reviewer and checked by a second. Disagreements were resolved by consensus or referral to a third reviewer when necessary. Authors were contacted for additional information required to determine eligibility when necessary.

Selection of target conditions
A list of potential target conditions was established using the paediatric diagnostic clinical decision support tool Isabel.[62] Final target conditions were decided through consultation with clinical experts and by examining the literature to determine those that have:

1. A reasonable likelihood of the primary presentation being short or tall stature and occurring in the relevant age range (i.e. we focused on growth-related conditions that are less likely to be identified by routes other than growth monitoring in school age children).
2. Generally accepted treatment options (including conditions where the management of the patient is likely to be significantly affected by the diagnosis, as in monitoring for treatable complications).
3. Been identified in previous studies of growth monitoring programmes.

In the light of the current concerns regarding the increasing prevalence of obesity in children and young people, the review also assessed obesity as a potential target condition.

Inclusion criteria for the review
Separate inclusion criteria for the five phases of the project derived from a systematic review of the literature are summarised in *Table 3*. Systematic reviews that met a predefined set of quality criteria[63] and were relevant to any of these phases were to be included if they fulfilled all the inclusion criteria. If they did not meet all inclusion criteria, systematic reviews were used as sources of potentially relevant primary studies.

Applying the inclusion criteria
A two-stage process was used to determine the eligibility of full paper publications. First, a general inclusion screen was undertaken to identify primary studies measuring growth in the target population. Studies meeting these criteria were then assessed using a detailed inclusion screen to identify studies meeting the full inclusion criteria detailed in *Table 3*.

Economic evaluations
Studies were included in the review if they met the criteria of being full economic evaluations, namely that they included an explicit analysis of both costs and effects for an intervention and at least one comparator and were considered to be useful in answering the research questions relating to cost-effectiveness.[64]

Data extraction

Systematic review of effectiveness
A database programmer was consulted to develop data extraction forms using Microsoft Access. These were piloted independently on a small selection of studies and adjusted as necessary. Data extraction was performed by one reviewer and checked by a second. Disagreements were resolved by consensus or by referral to a third reviewer when necessary. Foreign language papers were extracted by a single reviewer competent in that language but were not checked by a second reviewer (one Spanish study, one German paper supplementing an English publication).

For each included study, data were extracted on study identifier, objective, geographical details (area and country), location (community or school), selection procedure, number of children approached and number actually measured, age of measurement, gender and ethnicity of study population. In addition, data specific to each phase of the review were extracted as detailed below.

Studies of detection rate of growth monitoring programmes
Data were extracted on who measured the children, the method used to measure height and weight and the method of recording these measurements. In addition, the following details were extracted: the referral threshold used, number of children referred for tall or short stature, number of children diagnosed with any growth-related condition, the specific condition diagnosed and whether these were new or existing cases.

Studies of clinical effectiveness of routine growth monitoring
No studies were identified that met the inclusion criteria for clinical effectiveness of growth monitoring. Therefore, no data were extracted.

Studies of diagnostic performance of growth monitoring programmes
Growth
No studies were identified that met the inclusion criteria of diagnostic performance of growth monitoring programmes for the detection of growth-related conditions. Therefore, no data were extracted.

Obesity
Data were extracted on study design (cohort or case–control), details of the reference standard and index test used, diagnostic threshold and the results to allow for the construction of a 2×2 table.

Studies of human resource requirements of growth monitoring programmes
Data were extracted on staff training programmes, the impact of the growth monitoring programme

TABLE 3 Inclusion criteria for each phase of the review

	Epidemiology of growth-related conditions and comparison of detection rates	Clinical effectiveness of routine growth monitoring	Diagnostic performance of growth monitoring strategies	Human resource requirements of growth monitoring programmes	Attitudes of children, parents and healthcare professionals to growth monitoring
Study design	Studies of any design (case reports will be excluded)	Individual or cluster RCTs or CCTs comparing screened versus unscreened populations or different monitoring strategies	Diagnostic cohort or case–control studies	Studies of any design	
Population	Children between the ages of 4 and 11 years in Western Europe (including Scandinavian countries), North America, or Australia/New Zealand (excluding studies of aboriginal populations). Studies which also include overlapping age groups outside the prespecified range will be included			Healthcare professionals delivering growth monitoring programmes	Children, parents and healthcare professionals involved in a growth monitoring programme
Target conditions	GHD; JH; TS and other syndromes associated with short stature; psychosocial growth failure caused by emotional, physical, sexual abuse and neglect; clinical conditions associated with tall stature (including precocious sexual maturation, Klinefelter and Marfan syndromes) and obesity. Studies that address other target conditions for growth monitoring will be considered for inclusion on an individual basis following discussion within the review team			Not applicable	Not applicable
Intervention (method of growth measurement)		Any single or serial height measure, height/weight ratio, BMI, alternative obesity indices (e.g. waistline, skinfold), as part of a population-level assessment. Growth measurement may be combined with an intervention targeting nutrition/obesity			
Outcome measures	Detection rate of target conditions, age at diagnosis, route to diagnosis	The effects of the growth monitoring programme on age at diagnosis, patient management/outcome, including referrals to appropriate specialists	Data to construct a 2 × 2 table of the method of growth measurement in relation to final diagnosis	Any measure of resource requirement including staff numbers, workload or training requirements. Any subsequent impact on referral to specialists	Any measure of attitude such as take-up rate, effects of 'false positive' results and consequences of a positive finding

CCT, controlled clinical trial; RCT, randomised controlled trial.

on staff workload and effect on performance of growth monitoring programmes of any staffing parameters studied. Any details of the impact on subsequent referrals to specialists were also documented along with any relevant cost data.

Studies of attitudes to growth monitoring programmes
Data were also extracted on the number of children or parents refusing to take part in the growth monitoring programme or referral for further investigation, and any other outcome relating to attitudes towards the programme.

Studies of economic evaluations
Data extraction was performed on the included studies according to the methods adopted by NHS EED.[61] In addition, all resource use and unit cost data were extracted from each study. This additional information was considered to be useful in informing resource use and cost data for the decision analytic modelling phase of the project.

The full structured abstracts for each of the included studies detail the principal methods and results of each study, and also discussions on the generalisability of the results to the UK context.

Quality assessment
Systematic review of effectiveness
Quality assessment was carried out by one reviewer and checked by a second. Disagreements were resolved by consensus or referral to a third reviewer when necessary. Data specific to the type of study were extracted. Quality assessment tools were used to evaluate studies of detection rate, studies of diagnostic accuracy and randomised controlled trials (RCTs).

Studies of detection rate of growth monitoring programmes
The majority of studies described a growth monitoring programme already in place and reported a 'diagnostic yield' based on the numbers of children screened, the number found to be below/above a threshold for height and the number subsequently diagnosed with a growth-related condition. In this type of study, only short or tall children are followed up. Hence only true positives and false positive results can be obtained. Children found to be of normal height are not followed up, so false negatives and true negative results cannot be determined. These studies cannot be treated as studies of diagnostic accuracy and, as such, are inherently flawed if assessed

using a tool specific for diagnostic accuracy [Quality Assessment of Diagnostic Accuracy Studies (QUADAS)]. Therefore, it was considered appropriate to develop a review-specific tool to assess the methodological quality of these studies that focused on determining how reliable these studies are at determining diagnostic yield of monitoring in terms of growth disorders. This tool encompassed concepts of external validity such as the representativeness of the sample, internal validity such as consistency of measurement and reporting issues such as attrition rates. The methodological tool and details on how studies were scored are reported in Appendix 4.

Studies of diagnostic accuracy of obesity
Included diagnostic accuracy studies for obesity were assessed for methodological quality using a modified version of the QUADAS tool. The modified QUADAS tool assessed whether selection criteria for children had been described, whether partial and/or differential verification bias had been avoided (all children received verification using the same reference standard of diagnosis) and whether incorporation bias had been avoided (the index test did not form part of the reference standard). The checklist also addressed the question of whether the reference standard and index tests had been reported in sufficient detail to permit replication, and whether test review bias had been avoided (the results of tests had been interpreted independently of each other). Finally, the studies were checked with regard to the reporting of uninterpretable results, whether all withdrawals had been accounted for and whether cut-off thresholds were pre-defined.

The modified QUADAS tool together with details on how studies were scored is reported in Appendix 5: Modified QUADAS tool to assess the quality of diagnostic accuracy studies.

Randomised controlled trials of human resource interventions
RCTs were assessed using a checklist developed by the Scottish Intercollegiate Guidelines Network[65] and used in National Institute for Health and Clinical Excellence (NICE) guideline development.[66] The checklist comprised 10 questions. These addressed the appropriateness of the research question and the adequacy of randomisation, allocation concealment and blinding methods, the baseline comparability of groups, that the only difference between the groups was the treatment under investigation, the adequacy of outcome assessment, level of loss to follow-up, whether intention-to-treat analysis was

used and comparability of results across study sites. The checklist is reported in Appendix 6.

Studies of economic evaluations

The quality assessment of each included study was undertaken using two methods. First, the quality of economic evaluations was assessed using a modified version of the 35-point checklist developed for authors of economic evaluation submissions to the BMJ,[67] to which an additional item was added (item 36) in order to report whether or not the authors had addressed the issue of the generalisability of the results. Each item in the checklist was given one of four responses: (a) yes, (b) no, (c) not clear and (d) not applicable. The checklists were completed independently by two health economists, with discrepancies being discussed and a final agreement reached. The direction of the result in terms of costs and effects was represented using the hierarchical matrix for economic evaluations.[68]

Second, for each study a critical review (textual) summary was completed following the approach adopted by the NHS EED database.[64] This includes an appraisal of the validity of the choice of comparator(s), the validity of the analysis of effectiveness results, the validity of the benefit measure used in the economic analysis, the validity of the cost results and a variety of other important issues, including whether or not the authors compared their results with those of other (similar) studies, whether generalisability was addressed by the authors, the principal limitations and strengths of the study and finally the implications of the study in terms of clinical practice and future research.

Data synthesis

Systematic review of effectiveness

Results were analysed and synthesised separately for each of the phases of the review as detailed below. The results of the review were then summarised against the UK NSC criteria for screening programmes.

Studies of detection rate of growth monitoring programmes

The number of cases of conditions outlined in *Tables 1* and *2* were tabulated. Detection rates of new cases only were calculated separately for conditions of interest and all new cases of growth related conditions and were presented in graphical format and 95% confidence intervals (CIs) were calculated as follows:

$$p = \frac{\text{number of new cases}}{\text{number of children measured } (N)}$$

$$\text{lower CI} = p - \left[1.96 \sqrt{\frac{p\,(1-p)}{N}} \right]$$

$$\text{upper CI} = p + \left[1.96 \sqrt{\frac{p\,(1-p)}{N}} \right]$$

In instances where the number of new cases was zero, the upper CI was estimated as $3/N$.[69]

The results were presented narratively by age group or age range of children measured, with more emphasis given to the larger study populations that may be considered representative of the total UK population.

Studies of clinical effectiveness of routine growth monitoring

No studies met the inclusion criteria.

Studies of diagnostic performance of growth monitoring programmes

Growth
No studies met the inclusion criteria.

Obesity
Results were analysed according to the method used to identify obesity and, within these groups, methods were further grouped by specific threshold.

For each individual dataset, the sensitivity and specificity with 95% CIs were calculated from the 2×2 tables and were presented. CIs were calculated using the 'Wilson score confidence interval' method.[70] This method is recommended as it prevents values close to 1 from having an upper confidence limit which exceeds 1. Sensitivity and specificity were selected as the main outcome measures because they represent a simple expression of the relationship between two tests and how this varies with threshold.

Heterogeneity precluded statistical pooling, therefore results were presented in a narrative synthesis, grouped by index test with more consideration given to study populations that may be considered representative of the total UK population.

Studies of human resource requirements of growth monitoring programmes

A narrative synthesis was presented.

Studies of attitudes to growth monitoring programmes
A narrative synthesis was presented.

Studies of economic evaluations

The full structured abstracts for each of the economic evaluations were reported, detailing the principal methods and results of each study, and also discussions on the generalisability of the results to the UK context. Summaries were presented in a narrative.

Chapter 4
Results of the literature search

The literature searches identified 33,689 references. These were screened for relevance and 737 were considered to be potentially relevant. Full paper copies of these articles were obtained and were assessed for inclusion in the review. *Figure 1* shows the flow of studies through the review process and the number of studies excluded at each stage.

A total of 31 studies reported in 37 publications met the effectiveness review inclusion criteria (*Tables 4* and *5*). Four studies were reported in one publication. Two of the studies meeting the inclusion criteria were obtained through contact with PCTs and StHAs.[71,72]

Four economic evaluations met the inclusion criteria for the review of cost-effectiveness of growth monitoring.[21,73–75]

Overview of studies included in each phase of the review

Tables 4 and *5* give an overview of the questions addressed by each of the included studies.

Detection rate of growth monitoring programmes

Twelve studies provided data on growth-related conditions detected by growth monitoring programmes.[56,59,71,76,79,81–83,86,87,91,93] Eight studies were conducted in the UK,[56,59,71,76,82,86,91,93] one in Sweden,[79] one in Spain,[81] one in Germany[83] and one in the USA.[87] All studies were of a 'diagnostic yield' design. A change in the protocol as described in Appendix 2 meant that studies monitoring trends in obesity to derive prevalence were excluded from the review.
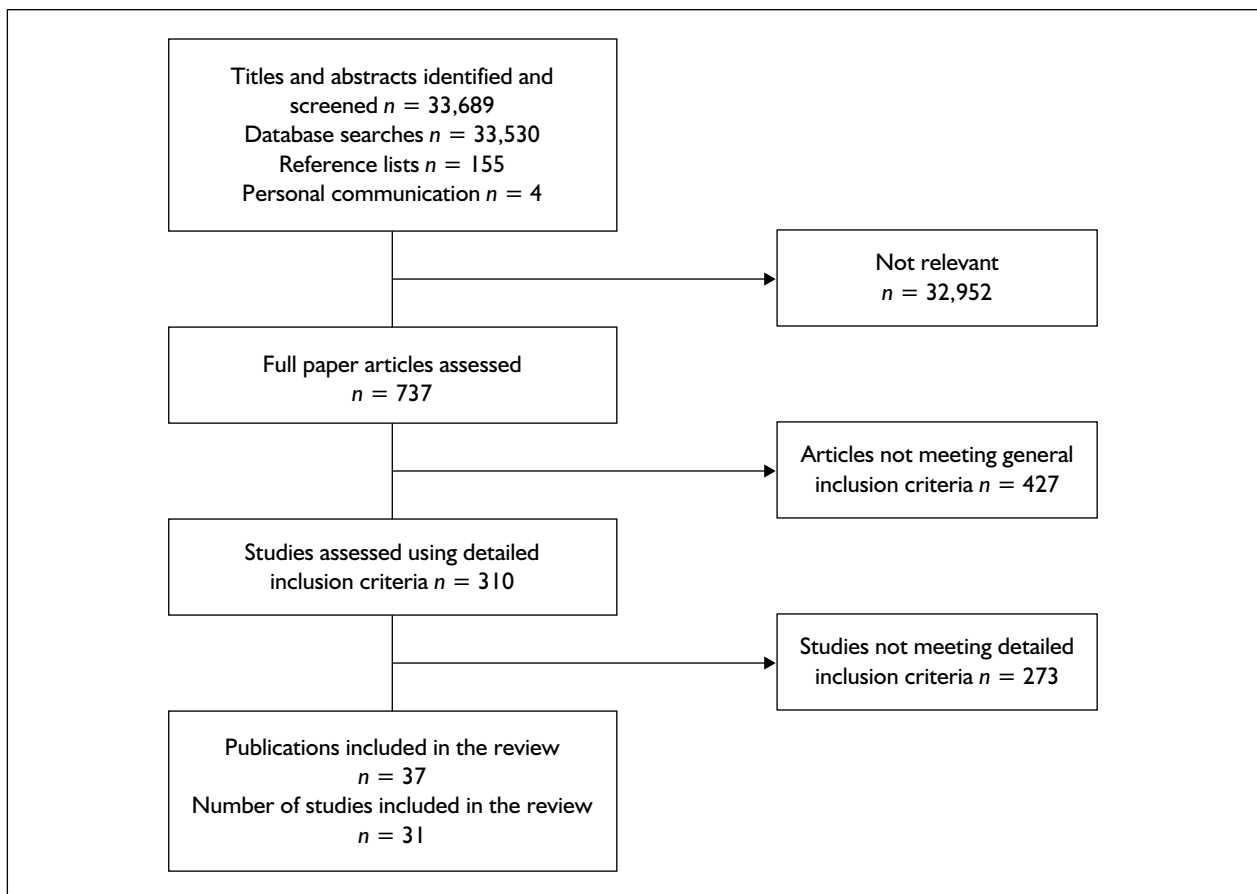
FIGURE 1 Flow chart of studies through the review process

TABLE 4 *Overview of questions addressed by growth studies*

Study and related publications	Addresses review questions				
	Detection rate	**Clinical effectiveness**	**Diagnostic accuracy**	**Human resources**[a]	**Attitudes to growth monitoring**[a]
Agwu (2004)[76]	✓	✗	✗	✓	✓
Ahmed (1995)[56,77]	✓	✗	✗	✓	✓
Aszkenasy (2005)[71]	✓	✗	✗	✓	✓
Banerjee (2003)[59]	✓	✗	✗	✓	✓
Cotterill (1996)[78]	✗	✗	✗	✓	✗
Cernerud (1994)[79]	✓	✗	✗	✓	✓
Cowan (2001)[80]	✗	✗	✗	✓	✗
de la Puente (1999)[81]	✓	✗	✗	✓	✓
Hearn (1995)[82]	✓	✗	✗	✓	✓
Keller (2002)[83–85]	✓	✗	✗	✗	✓
Lacey (1974)[86]	✓	✗	✗	✗	✓
Lindsay (1994)[87]	✓	✗	✗	✓	✓
Lipman (2004)[88]	✗	✗	✗	✓	✗
Mulligan (1998)[89]	✗	✗	✗	✓	✗
van Buuren (2004)[90]	✗	✗	✗	✓	✗
Vimpani (1981)[91,92]	✓	✗	✗	✓	✓
Voss (1992)[41,93,94]	✓	✗	✗	✓	✓
Welch (1982)[95]	✗	✗	✗	✓	✗
White (1995)[96]	✗	✗	✗	✓	✓

[a] Some studies did not directly address the review question. However they provided some limited data. More details are given on page 21.

TABLE 5 *Overview of studies of obesity*

Study	Addresses review questions			
	Clinical effectiveness	**Diagnostic accuracy**	**Human resources**	**Attitudes to growth monitoring**
Bedogni (2003)[97]	✗	✓	✗	✗
Cernerud (1994)[79]	✗	✗	✓	✓
Ebbeling (1999)[98]	✗	✓	✗	✗
Ellis (1999)[99]	✗	✓	✗	✗
Himes (1989)[100]	✗	✓	✗	✗
Lazarus (1996)[101]	✗	✓	✗	✗
Marshall (1991)[102]	✗	✓	✗	✗
Mast (2002)[103]	✗	✓	✗	✗
Reilly (1999)[104]	✗	✓	✗	✗
Reilly (2000)[105]	✗	✓	✗	✗
Routh (2005)[72]	✗	✗	✗	✓
Schaefer (1998)[106]	✗	✓	✗	✗
Welch (1982)[95]	✗	✗	✓	✗
Wickramasinghe (2005)[107]	✗	✓	✗	✗

Clinical effectiveness of growth monitoring programmes

No studies were identified that met the inclusion criteria for evaluating the clinical effectiveness of growth monitoring to detect disorders of stature or obesity.

Diagnostic performance of growth monitoring programmes

Growth

No studies were identified that met the inclusion criteria for evaluating the diagnostic performance of growth monitoring programmes for the identification of disorders of stature.

Obesity

Eleven studies provided data on the diagnostic performance of measures used to identify obesity.[97–107] A variety of index tests and reference standards were evaluated. Six studies evaluated a measure of BMI using different thresholds and classifications.[99,101,103,105–107] One study evaluated BMI and skinfold thickness.[97] Three studies evaluated BMI and skinfold thickness in addition to weight,[100] weight for height[98] or relative weight.[102] One study evaluated BMI and ideal body weight.[104] The majority of included studies used densometrically defined body fat as the reference standard. Two studies were conducted in the UK,[104,105] two in the USA,[98,99] two in Australia,[101,107] two in Germany,[103,106] two in Canada[100,102] and one in Italy.[97]

Human resource requirements of routine growth monitoring programmes

One cluster RCT conducted in the USA,[88] two before-and-after studies (one in the UK and one in the USA)[80,95] and one UK audit[71] specifically evaluated an intervention relating to the human resources requirements of growth monitoring. Three of these were related to the measurement of height[71,80,88] and one to both weight and height.[95] One Dutch and two UK studies determined the impact of growth monitoring guidelines or changes in practice on human resources.[78,89,90]

Three of the included growth monitoring programme studies reported cost data.[56,59,79] Four of the included growth monitoring programme studies reported on human resource-related data as part of the process of evaluating the monitoring programme[56,82,87,96] and eight of the included growth monitoring programmes used methods to ascertain or verify the height of children identified as meeting the predefined referral criteria.[56,59,76,79,81,82,91,93]

Attitudes towards routine growth monitoring programmes

Just one study was identified that specifically addressed attitudes towards growth monitoring. This UK study focused on monitoring for obesity and used questionnaires to survey teachers and school nurses.[72]

One of the growth monitoring programme studies assessed the attitudes of healthcare professionals to growth monitoring in Sweden[79] and two UK studies and one US study provided other relevant additional attitudinal data.[87,91,93]

Attendance at initial measurement and at further investigation were extracted for each of the growth monitoring programme studies and reasons for eligible children not attending were also noted.[56,59,71,76,79,81–83,86,87,91,93,96]

Economic evaluations

Four published studies met the inclusion criteria for full economic evaluations addressing a question related to the cost-effectiveness of post-monitoring interventions for target conditions included in the systematic review.[21,73–75] The cost-effectiveness findings reported in a NICE Technology Appraisal on the use of human GH (somatropin) in children with growth failure[108] are also summarised as the report contains data from company submissions not available in other published sources. No studies were identified that formally addressed the cost-effectiveness of child growth monitoring programmes themselves.

Two UK studies were found dealing with short stature and the cost-effectiveness of GH treatment.[21,73] Two US studies on obesity were also found, one dealing with the cost-effectiveness of a primary prevention programme[74] and one reporting the cost-effectiveness of two treatment interventions for childhood obesity.[75]

Overview of studies excluded from the review

Of 310 papers assessed using detailed screening criteria, 273 were excluded. A total of 114 studies were excluded as no screening or monitoring of height or weight in an unselected population was performed; 107 studies provided details of screening or monitoring of height or weight in an unselected population but were excluded as no data other than epidemiological data on obesity were presented. Thirty-nine studies provided details of screening or monitoring of height or

weight in an unselected population and data other than epidemiological data on obesity but did not include participants with the target conditions and were excluded as no data on human resources or attitudes were reported. Thirteen studies provided details of a screening or monitoring of height or

weight in an unselected population, data other than epidemiological data on obesity and included participants with the target condition of interest but were excluded as no relevant outcome data were presented. Bibliographic details of excluded studies are available from the authors.

Chapter 5
Results of the review of clinical effectiveness

Detection of growth-related conditions and comparison of detection rates

Twelve of the included growth monitoring programme studies provided data on the diagnostic yield of growth related conditions from routine growth monitoring.[56,59,71,76,79,81–83,86,87,91,93] Only two studies[79,83] focused on conditions relating to tall stature and short stature; the remainder aimed to identify conditions of short stature only. A brief overview of the programmes is given in *Table 6*. Appendix 7 provides more detailed descriptions of growth monitoring programmes in the included studies.

Overview of the growth monitoring programmes

The number of children measured in the included programmes ranged from 1592[59] to 114,881.[87] The percentage of eligible children measured ranged from 45%[81] to 90%,[76] where reported. Most of the monitoring programmes were based on a single screen of a child's height to detect growth-related conditions. Some studies measured children of the same age group[59,71,76,86,93] whereas others studied a cross-section of ages using single or serial measurement but providing results only for the entire cohort rather than by age group.[56,79,81–83,87,91] Full details of ages measured are given in *Table 6*.

Table 7 gives details of personnel involved in measuring and the measuring equipment used. Five programmes were delivered by a school nurse,[71,79,82,86,93] one by school nurses and support workers,[76] one by a health visitor,[56] one by a health visitor or classroom assistant,[59] one by a practising paediatrician,[83] one by a trained volunteer[87] and one by an investigator[91] and one study did not report who measured the children.[81] A range of measuring equipment was used across the growth monitoring programmes, as detailed in *Table 7*. Details of reference charts and diagnostic thresholds used are also presented. Briefly, three studies used the UK 1990 charts,[59,71,76] five used the Tanner and Whitehouse charts[56,82,86,91,93] and four used country-or study-specific charts.[79,81,83,87]

In some of the included growth monitoring programme studies, children who met the threshold for referral were re-measured, usually by trained auxologists.[56,59,76,79,81,82,91,93] In some cases this involved referral to a community growth clinic or similar, where preliminary investigations to ascertain the cause of short stature were also undertaken with the aim of reducing the number of short normal children referred to a specialist for further investigation. Preliminary investigations included bone age assessment, determination of mid-parental height, height velocity or karyotype testing. However, it is not possible in most cases to determine whether conditions were diagnosed at the clinic or after subsequent referral to a specialist endocrinologist or paediatrician.

Methodological quality of growth monitoring programme studies

Table 8 gives an overview of the results of the methodological quality assessment. Of the 12 studies assessed, one met nine out of the 10 quality criteria[56] and three clearly met eight of the criteria.[59,81,93] Half of the studies met seven,[71,76,82,83,87,91] one met six[79] and one met only three criteria.[86] Ten studies reported a clearly defined selection procedure, providing details of eligibility for the study, and indicating whether the sample was random or a whole cohort based on age group and/or region.[56,59,71,76,81–83,87,91,93] Eight studies explicitly made attempts to contact all children who were eligible for measurement.[56,59,79,81,83,87,91,93] However, five studies failed to measure more than 80% of their sample.[56,59,81,82,86] Three further studies did not explicitly state the number of eligible children, only the number who were measured; therefore, it was not possible to assess their level of coverage.[79,83,91] A description of a reproducible protocol for taking and interpreting height measurements, detailing equipment used and charts and thresholds for referral, was provided by 10 of the studies.[56,59,71,76,79,81–83,91,93] Nine studies described methods to ensure the competence of the people carrying out the measurement protocol through adequate training in measurement techniques,[56,59,76,81–83,87,91,93] but only six reported checking the competence of these measurers in

24

TABLE 6 Overview of growth monitoring programmes

Study and location	Selection procedure	No. measured (% of eligible)	Gender – no. of males (%) and ethnicity	Measured age(s)
Agwu (2004),[76] UK	All children in one school year (September 1999 to August 2000) in the Sandwell district (87 schools) were eligible	3,474 (90%)	1,729 (50%) Caucasian 75% Other ethnicity 25%	A single measurement of reception class children aged 4–5 years
Ahmed (1995),[56] UK	All Oxfordshire children measured as part of their routine development checks were eligible. The study took place from 1988 to 1994	20,338 (66%)	11,808 (58%)	A single measurement of two groups of children, aged 3 years ($n = 11,603$) and 4.5 years ($n = 11,477$). 2,742 of these children were measured at both ages to determine height velocity
Aszkenasy (2005),[71] UK	Audit of all children in the area measured as per routine school entry height monitoring policy over three school entry years (1999–2001)	9,338 (83%)	NR	A single measurement of three cohorts of children at primary school entry
Banerjee (2003),[59] UK	Audit of all children born between September 1992 and August 1993 and measured in the school year September 1998 to August 1999 in the Rhondda and Taff Ely area	1,592 (68%)	NR	A single measurement of children aged around 6 years (range 5 years 3 months to 6 years 8 months)
Cernerud (1994),[79] Sweden	Random samples of school classes receiving routine health surveillance height monitoring were selected	7,129 (NR)	NR	A single measurement of two groups of children, aged 10 ($n = 3,239$) and 14 years ($n = 3,890$)
de la Puente (1999),[81] Spain	Primary care teams linked to three hospitals in the province of Barcelona were invited to participate. The eight which elected to participate had to screen all children born between 1986 and 1987 under their jurisdiction	2,084 (45%)	1,093 (52%)	A single measurement of a group of children aged between 5 and 8 years: aged 5 years ($n = 10$), 6 years ($n = 811$), 7 years ($n = 1,234$) and 8 years ($n = 26$)
Hearn (1995),[82] UK	All primary and secondary school entrants in Hackney over three school entry years (1990–2) were eligible	9,549 (79%)	NR	A single measurement of two groups of children. Primary school entrants had a mean age of 5 years 3 months ($n = 6,421$). Secondary school entrants group had a mean age of 11 years 8 months ($n = 3,128$)

continued

TABLE 6 Overview of growth monitoring programmes (cont'd)

Study and location	Selection procedure	No. measured (% of eligible)	Gender – no. of males (%) and ethnicity	Measured age(s)
Keller (2002),[83] Germany	Height data submitted over a 2-year period from September 1998 by paediatricians in practice, or in paediatric group practice, or public health school physicians participating in the CrescNet collaborative network were eligible	60,984 (NR)	31,021 (51%)	Data were reported for children and adolescents who had been measured at least once. Males had a mean age of 8.15 years (range 0–18.77 years) and females had a mean age of 8.3 years (range 0–19.28 years)
Lacey (1974),[86] UK	All children born during 1960–2 to mothers living in Newcastle were eligible	2,256 (45%)	NR	A single height measurement at the age of 10 years
Lindsay (1994),[87] USA	Schools in Utah were randomly selected and invited to participate. 251 schools representing 38/40 school districts agreed and were required to measure all children from kindergarten to fifth grade	114,881 (81%)	59,087 (51%)	Serial measurements were taken to assess height and growth velocity. The first measurement was of children aged between 5 and 11 years. 79,495 of these children participated in the second measurement 12 months later
Vimpani (1981), UK[91]	All 2nd and 3rd year (and 4th year in Aberdeen) primary school pupils attending education authorities and some independent schools in Edinburgh, Glasgow and Aberdeen were screened during 5 months in 1975–6	48,221 (NR)	24,670 (51%)	A single height measurement of children aged between 6 and 9 years
Voss (1992), UK[93]	All children in the districts of Winchester and Southampton entering local authority primary schools in two consecutive years (1985/7) were eligible	14,346 (100%)	NR	Single height measurement in children at school entry aged 5 years

NR, not reported.

TABLE 7 *Overview of growth monitoring protocols*

Study	Measurer	Method of measuring height	Method of referencing height	Diagnostic threshold
Agwu (2004)[76]	School Health Nurse, School Health Nurse support worker	Leicester height measure	UK 1990 charts	<0.4th centile
Ahmed (1995)[56]	Health visitor	Microtoise, Minimetre (Raven), Oxford screening wall charts (Castlemead Publications)	Tanner and Whitehouse 1965	1. <–2 SDs (referral to auxologist) 2. <–3 SDs (referral to consultant) 3. SD decrease between measurements (referral to auxologist)
Aszkenasy (2005)[71]	School nurse	Leicester Height Measure	0.4th centile tables based on UK 1990 charts	<0.4th centile
Banerjee (2003)[59]	Classroom assistant, health visitor	Leicester Height Measure	UK 1990 charts	<0.4th centile
Cernerud (1994)[79]	School nurse	Fixed vertical backboard with scale and horizontal blade	Reference data used by schools in the region (unspecified)	Height or height for weight outside 2 SDs, or change of >0.5 SD/year
de la Puente (1999)[81]	NR	Accustat wall-mounted stadiometer	Growth charts for Catalonia	≤3rd centile
Hearn (1995)[82]	School nurse	Minimetre (Raven)	Tanner and Whitehouse 1965	<3rd centile
Keller (2002)[83]	Practising paediatrician	Standiometer (Dr Keller system)	German Synthetic Norm Curve	<3rd centile or >97th centile
Lacey (1974)[86]	School nurse	NR	Tanner and Whitehouse 1965	1st cohort: <3rd centile 2nd cohort: <–3 SDs
Lindsay (1994)[87]	Trained volunteer	Accustat stadiometer	NR	1st phase: <–2 SDs 2nd phase: <3rd centile and growth rate <5 cm/year
Vimpani (1981)[91]	Investigator	Harpenden portable anthropometer	Tanner and Whitehouse 1965	≤–2.5 SDs
Voss (1992)[93]	School nurse	Portable wall-mounted microtoise	Tanner and Whitehouse 1965	<3rd centile

NR, not reported.

TABLE 8 *Quality of detection rate studies*

Study	Were selection criteria clearly described?	Was an attempt described to contact all children who were eligible for measurement?	Were methods described to ensure the competence of those measuring?	Were methods described to check the competence of those measuring?	Was there a reproducible protocol to ensure accuracy and consistency of measurements?	Was >80% of the sample actually measured?	Were details of measurement error provided?	Were all children measured accounted for in measurement results?	Were all children identified as needing follow-up accounted for in terms of diagnosis/false positive/lost to follow-up?	Were sufficient details provided concerning those diagnosed with a growth disorder?
Agwu (2004)[76]	✓	×	✓	×	✓	✓	✓	✓	×	✓
Ahmed (1995)[56]	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
Aszkenasy (2005)[71]	✓	×	×	×	✓	✓	✓	×	×	✓
Banerjee (2003)[59]	✓	✓	✓	×	✓	×	✓	✓	✓	✓
Cernerud (1994)[79]	×	✓	×	✓	✓	?[a]	×	✓	✓	✓
de la Puente (1999)[81]	✓	✓	✓	✓	✓	×	✓	×	✓	✓
Hearn (1995)[82]	✓	×	✓	✓	✓	×	✓	✓	×	✓
Keller (2002)[83]	✓	✓	✓	×	✓	?	×	✓	✓	✓
Lacey (1974)[86]	×	×	×	×	×	×	×	✓	✓	✓
Lindsay (1994)[87]	✓	✓	✓	✓	×	✓	×	✓	✓	×
Vimpani (1981)[91]	✓	✓	✓	✓	✓	?[a]	×	✓	✓	×
Voss (1992)[93]	✓	✓	✓	×	✓	✓	✓	✓	×	✓

[a] This question could not be answered as the number of children eligible for measurement was not explicitly stated.

terms of accuracy and/or reliability.[56,79,81,82,87,91] Seven studies provided information relating to errors in the measurements taken.[56,59,71,76,81,82,93] Only one study failed to account for all measured children in the results.[81] Three studies failed to provide follow-up details for all children referred for further investigation.[76,82,93] Two studies did not provide information on the status (new case versus previously known case) of all children diagnosed with growth disorders.[87,91]

Results of detection rate of growth-related conditions and comparison of detection rates

Appendix 8 details the conditions diagnosed by the studies, where *Table 23* lists all conditions as reported by each study and *Table 24* groups conditions named in the protocol (GHD, TS, JH, psychosocial growth failure and conditions of tall stature) separately from a general group of other conditions. The yields with 95% CIs of new cases for each of the conditions of interest, and for new cases of all conditions, are presented graphically

in *Figures 2* and *3*. The studies are grouped by the age group or age range of children measured, and further detailed descriptions of each study within this same grouping are provided in the following sections.

Results from monitoring at or before primary school entry

Three studies described programmes measuring children at primary school entry only (ages 4–5 years) and provided data on detection rates for this age group.[71,76,93] One study measured children at age 3 and 4.5 years.[56] All four studies were conducted in the UK. Other included studies also measured children at school entry age along with older school age children, but the results were not separated by age.

Voss and colleagues conducted the growth screening programme reported in the Wessex study[93] in order to identify participants for a study of 'normal' short children. Primary school entry cohorts were measured and all children with a
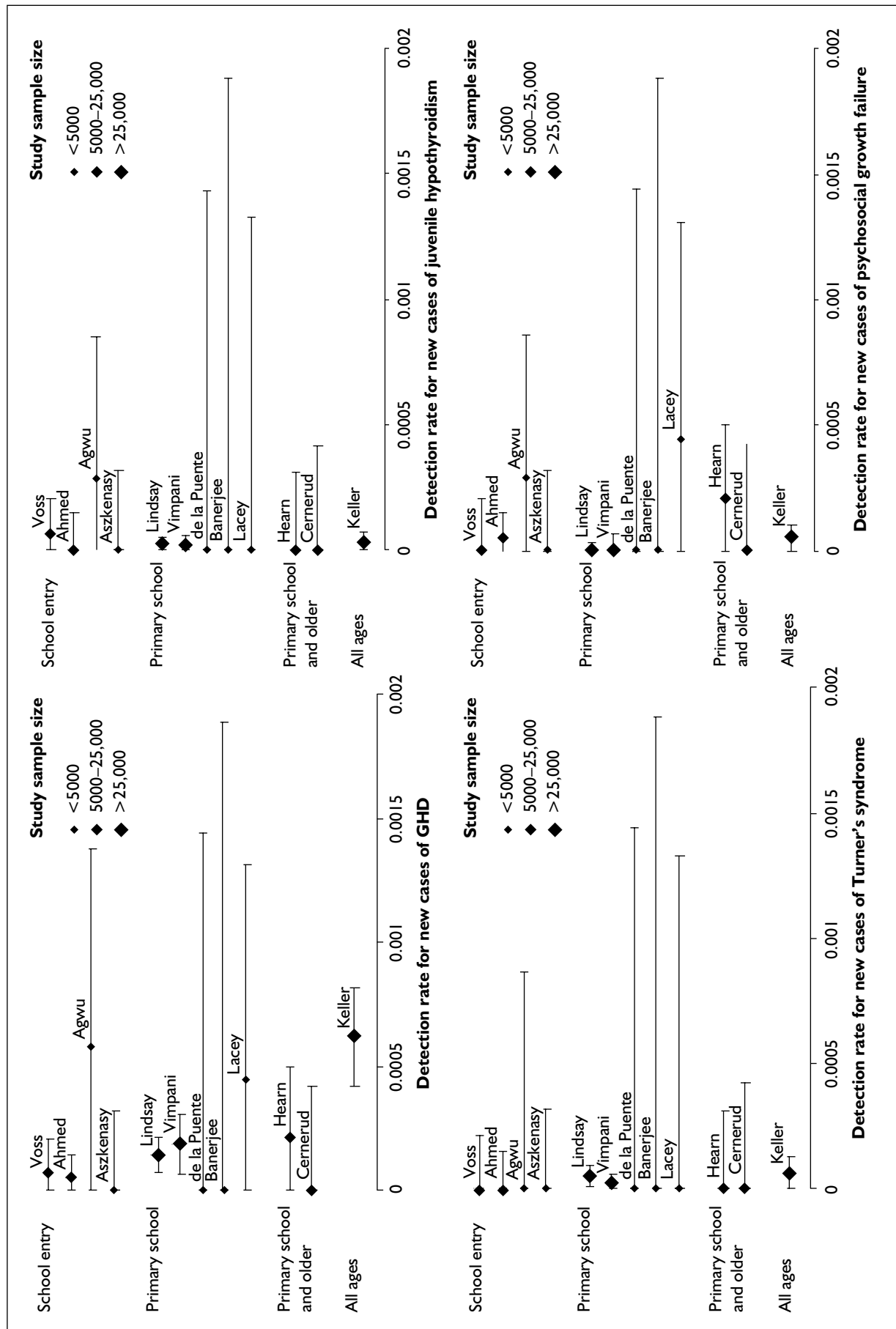
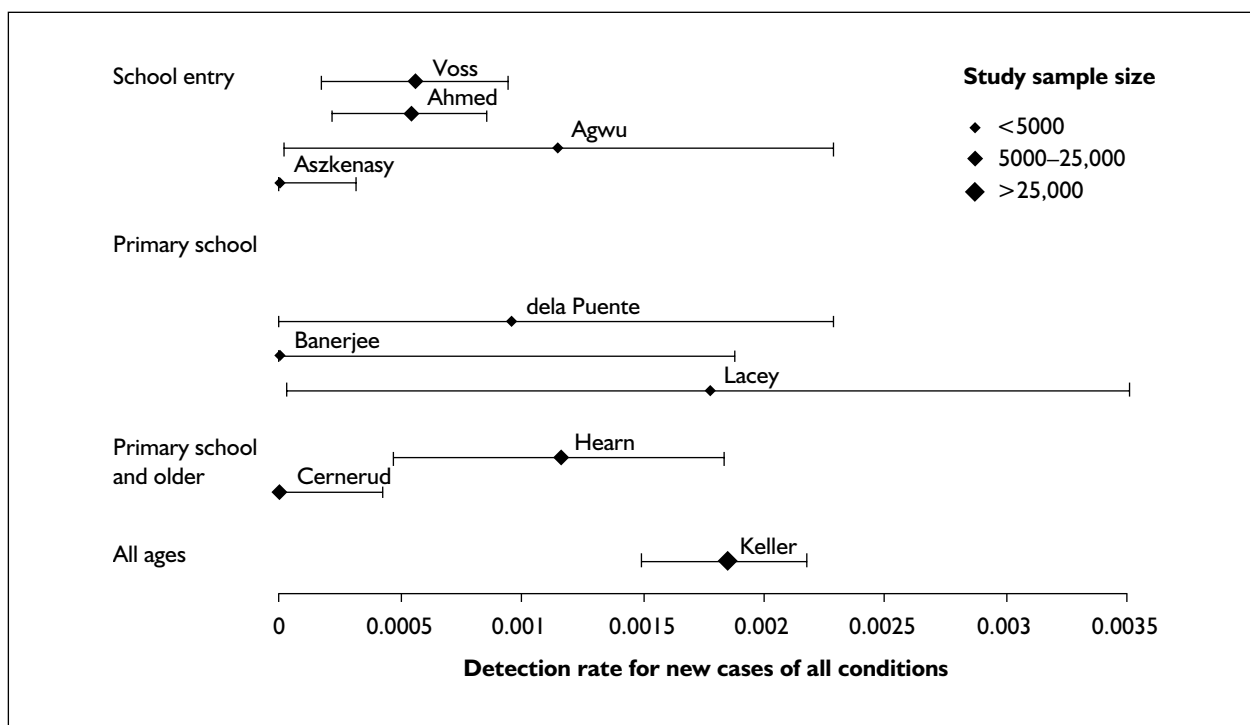*FIGURE 2 Detection rates for new cases of conditions of interest*

FIGURE 3 *Detection rates for new cases of conditions of interest*

height below the 3rd centile based on the Tanner and Whitehouse growth charts were investigated for causes of short stature. A large number of children were screened (n = 14,346), the study met most of the methodological quality criteria and detailed follow-up data were provided for almost all identified children. Among the 180 children meeting referral criteria, two had GHD, one of which was a new case. There was one child with known TS, one new case of JH and six new diagnoses of other medical conditions. The total yield of new cases was therefore 1:1793 children screened. The authors also reported that among the identified normal short children, 23 were asthmatic and 32 were judged to have psychosocial deprivation. It is unclear whether these conditions had been identified prior to the study.

The Oxford study conducted by Ahmed and colleagues[56] was of a similarly high methodological quality to the Wessex study, and also provided results from a large number of children (n = 20,338), although only 66% of those eligible were actually measured, and follow-up data were available for only 66% of referred children. Children were measured at the age of 3 (n = 11,603) or 4.5 years (n = 11,477), and 2742 children were measured at both ages in order to determine height velocity. All children with a height at initial measurement less than 2 SDs below the mean, according to the Tanner and

Whitehouse charts, were referred for further investigation either by an auxologist at a community clinic or, if less than 3 SDs below the mean, by a paediatric endocrinologist. Of the children for whom diagnoses were available, two had GHD, one of whom was a new case, there was one new case of psychosocial growth failure and there were nine new cases of other conditions. The total yield of new cases was therefore similar to that of the Wessex study, 1:1849 children screened.

The studies by Agwu and colleagues[76] and Aszkenasy[71] were smaller and had more methodological limitations than the Oxford and Wessex studies. However, both were conducted after the introduction of the most recent UK recommendations for growth screening[34] involving the use of the 0.4th centile of the UK 1990 centile growth charts as the threshold for referral of short children.

Agwu and colleagues[76] evaluated the Sandwell district growth monitoring programme and published results as a letter. The heights of 3465 children were measured, with followed-up diagnoses of 78% of those referred for short stature. Among these children there were two new cases of GHD (1:1733 children screened) and one new case each of JH and psychosocial growth failure. The yield of new cases was therefore 1:866 children screened.

29

Aszkenasy[71] performed an audit of a growth screening programme in Middlesbrough for the 1999, 2000 and 2001 school entry cohorts and results were available as a presentation. They measured 9338 children, 83% of those eligible. Diagnoses were available for 58% of children meeting the referral threshold, although follow-up of children unknown to the system was only 38%. This study suffered from a high number of refusals and non-attendees at the consultant paediatrician clinic assessment. No new cases of any underlying medical conditions were found among those children who did attend. The authors noted that one child who did not attend for screening, but who would have met referral criteria, was subsequently diagnosed with GHD.

Results from monitoring in children at primary school age only

Five other studies described monitoring programmes involving children of primary school age only.[59,81,86,87,91] Two of the studies measured children at a single age only; one measured at age 6 years[59] and the other at age 10 years.[86] One study measured a cohort of children aged between 5 and 8 years,[81] and another measured a cohort aged between 6 and 9 years.[91] Another study included children aged between 5 and 11 years, most of whom were measured on two occasions, 1 year apart.[87] Three studies took place in the UK,[59,86,91] one in Spain[81] and one in the USA.[87] Two of the studies measured very large numbers of children, one in excess of 100,000[87] and the other around 50,000.[91] The other three studies all had sample sizes of less than 2500.

The Utah study by Lindsay and colleagues was conducted in the early 1990s with the aim of assessing height and growth velocity and determining the prevalence of GHD in American children.[87] This was a very large study, meeting most methodological quality criteria. Schools across Utah were randomly selected to participate in a programme to screen the height of school children from kindergarten to fifth grade (age range 5–11 years), with a follow-up measurement 1 year later. After the first measurement, children with a height more than 2 SDs below the mean for the whole study population were highlighted for further investigation. After the second measurement, those with a height below the 3rd centile of the study population and with a growth rate less than 5 cm per year were referred. In total, 114,881 children were measured at least once. Among the children who were followed up, there were 16 new cases of GHD (1:7180 children screened), six new cases of TS (1:19,147 children

screened) and three new cases of JH (1:38,293 children screened). There were 53 children identified with medical disorders underlying their short stature, but it was unclear how many of these were new cases.

A similar study was conducted in 1975–6 in the UK by Vimpani and colleagues.[91] They attempted to perform a height screen of all second- and third-year primary pupils, along with some fourth-year children (full age range 6–9 years) in three Scottish cities. This large study, involving 48,221 children, was of reasonable methodological quality. However, as with the Utah study, the status (new case versus known case) of many children diagnosed with growth conditions was unclear. Children with a height more than 2.5 SDs below the mean according to the Tanner and Whitehouse charts were referred for further investigation of their short stature. Of the 38 children diagnosed with GHD, four were definitely existing cases and nine were definitely new cases, but the status of the remaining 25 children is unclear. Assuming that there were only nine new cases, the yield of new cases of GHD was 1:5358 children screened. The study picked up one new case each of TS and JH, giving a yield of 1:48,221 children screened for new cases of each condition. There were 140 children with other medical conditions underlying their short stature, but it is unclear how many of these cases had been diagnosed before the screening programme.

De la Puente and colleagues carried out a much smaller study in which they screened just over 2000 children of a similar age range to Vimpani and colleagues (ages 5–8 years) in schools in Spain.[81] Local reference growth charts were used, and children with heights less than the 3rd centile were referred for further investigation. No new cases were found of the target growth conditions, but there were two new cases of other conditions, giving a yield of new cases of 1:1042 children screened.

Another small study was carried out by Banerjee and colleagues in Wales.[59] They aimed to audit the screening of 6-year-olds in the Rhondda and Taff Ely area in the school year 1998–9 in which the heights of 1592 children were recorded. They investigated children with heights less than the 0.4th centile using UK 1990 growth charts and found no new cases of any growth disorders.

In the 1970s, Lacey and Parkin measured the heights of 10-year-olds as part of a larger longitudinal study investigating the development

of children born in Newcastle.[86] The study was both small and methodologically weak. Although the authors measured two cohorts of children, sufficient details on children measured and outcome data were only provided for one group of 2256 children who were born in 1960. Children with heights less than the 3rd centile according to the Tanner and Whitehouse charts were referred for further investigation. Among these children, there was one new case each of GHD and psychosocial growth failure. There were a further two new cases of other conditions, giving a yield of new cases for all conditions of 1:564 children screened.

Results from monitoring in age ranges including children older than primary school age

Two studies included a group of children older than primary school age in addition to those at primary school.[79,82] One UK study measured a group of children at primary school entry and another group at secondary school entry.[82] A Swedish study measured a group of children aged 10 years and a group aged 14 years.[79] Data were not provided separately by age group.

Hearn and colleagues carried out a height screen of 5- and 11-year-old primary and secondary school entrants in the London Borough of Hackney.[82] Just under 10,000 children were measured and the study was of reasonable methodological quality, although diagnoses were reported for only 40% of the children identified as short, and there was no separation of primary school and secondary school entrants in the results. The 3rd centile using Tanner and Whitehouse charts was used as the cut-off for short stature referral. The study identified two new cases of GHD and two new cases of psychosocial growth failure, giving a new case yield of 1:4775 children screened for each. Including seven new cases of other conditions, the yield of all new cases of any condition was 1:868 children screened.

The Swedish study by Cernerud and Edding[79] involved assessment of the regular health surveillance programme by looking at the height and weight measurements of random samples of 10- and 14-year-old children. Altogether 7129 children were measured, and those falling outside 2 SDs of the mean for height, or with a change in growth rate of more than 0.5 SD per year, were highlighted for follow-up by school doctors. The reference was that used by the Stockholm school health service. The school doctors referred children about whom they had concerns for specialist investigation. Among the children

referred to the specialist, there were no new cases of growth disorders.

Results from monitoring children of all ages

The German programme described by Keller and colleagues[83] differed from all the other included studies. This was a computer-based monitoring system, into which routine height measurements taken at over 100 participating paediatric practices throughout a wide area of Germany were input on an ongoing basis. Children with heights above the 97th centile or below the 3rd centile of the German synthetic norm curve were highlighted to the relevant practice, from where children were referred for specialist investigation if this was considered necessary. Results were presented for a 2-year period from 1998 to 2000 when measurements of 60,984 children were taken. Among those referred for specialist follow-up for short stature, there were 38 new cases of GHD (1:1605 children screened), four new cases of TS (1:15,246 children screened), two new cases of JH (1:30,492 children screened) and three new cases of psychosocial growth failure (1:20,328 children screened). There were also six new cases of conditions of tall stature (1:10,164 children screened). Including the 59 new cases of other conditions that were diagnosed, the total yield of new cases was 1:545 children screened.

Diagnostic performance of methods used to identify obesity

Eleven studies of diagnostic accuracy were found relating to the identification of obesity.[97–107]

Overview of diagnostic performance studies

The included diagnostic accuracy studies were, by their nature, based on one-off screening rather than ongoing monitoring. The sample size ranged from 138[107] to 3948,[105] with seven including fewer than 1000 children[97,99–102,104,107] and four including more than 1000 children.[98,103,105,106] An overview of the studies is presented in *Table 9*.

There were two UK studies, both conducted by Reilly and colleagues; these focused on detecting obesity in children aged 8 years[104] and 7 years.[105] The nine non-UK studies all involved children within a range of ages. Two studies only involved children of primary school age, Ebbeling and colleagues[98] with ages 6–9 years and Mast and colleagues[103] with ages 5–7 years. All other studies included children over the age of 11 years: Bedogni and

31

TABLE 9 *Overview of obesity diagnostic accuracy studies*

Study and location	Selection procedure	No. weighed (% of approached)	Gender – no. of males (%) and ethnicity	Weighed age(s) (years)
Bedogni (2003),[97] Italy	A convenience sample of children enrolled in primary and secondary schools in Modena and Parma. Year of data collection not reported	986 (not reported)	486 (49%)	8–12
Ebbeling (1999),[98] USA	Information on children was obtained from data tapes from the NHANES II study which was carried out in 1976–80	1171 (not reported)	585 (50%)	6–9
Ellis (1999),[99] USA	Selection procedure was not explicitly stated; children were living in Houston, TX. Data were collected in 1994–6	979 (not reported)	406 (41%) African American (black) 283 European American (white) 438 Hispanic American (Hispanic) 258	3–18
Himes (1989),[100] USA	Families in Quebec City volunteered to participate in response to media notices. Year of data collection not reported	316 (not reported)	159 (50%)	8–18
Lazarus (1996),[101] Australia	Volunteers recruited from among siblings of outpatients or children of staff and friends at The Children's Hospital, Sydney. Year of data collection not reported	230 (not reported)	119 (52%)	4–20
Marshall (1991),[102] Canada	Children from Alberta who participated in the Canada fitness survey in 1981 were selected using a two-stage stratified sampling frame	540 (90%)	266 (49%)	7–14
Mast (2002),[103] Germany	Random sample (40%) of children in 29/32 schools in Kiel in 1996–9	2286 (not reported)	1146 (50%)	5–7
Reilly (1999),[104] UK	Unclear, though it was stated that they were representative of Scottish children (based in Edinburgh). Year of data collection not reported	240 (not reported)	124 (52%)	8
Reilly (2000),[105] UK	4175 of approximately 14,000 children in the Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC) in 1998. The sample was representative of the birth cohort and UK, although slightly over-representative of wealthier families, and under-representative of ethnic minorities	3948 (95%)	2010 (51%)	7
Schaefer (1998),[106] Germany	Representative sample (19%) of the population in the area of Heidelberg in 1989–90	2554 (not reported)	1276 (50%)	6–19
Wickramasinghe (2005),[107] Australia	Not explicitly reported. The authors stated that subjects were contacted via newsletters and community centres in Brisbane. Year of data collection not reported	138 (not reported)	71 (51%) Australian Sri Lankan 42 Australian white Caucasian 96	5–15

TABLE 10 *Quality of obesity studies*

Study	Were selection criteria clearly described?	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	Did patients receive the same reference standard regardless of the index test result?	Was the reference independent of the index test (i.e. the index test did not form part of the reference standard)?	Was the execution of the index test described in sufficient detail to permit replication of the test?	Was the execution of the reference standard described in sufficient detail to permit its replication?	Were the index test results interpreted without knowledge of the results of the reference standard?	Were the reference standard results interpreted without knowledge of the results of the index test?	Were uninterpretable/intermediate test results reported?	Were withdrawals from the study explained?	Was the cut-off threshold predefined for the index test?	Was the cut-off threshold predefined for the reference standard?
Bedogni (2003)[97]	✗	✓	✓	✓	✗	✓	?	?	?	✗	✗	✗
Ebbeling (1999)[98]	✗	✓	✓	✓	✗	✗	?	?	?	✓	✓	?
Ellis (1999)[99]	✗	✓	✓	✓	✓	✓	?	?	?	✓	?	✗
Himes (1989)[100]	✗	✓	✓	✓	✗	✗	?	?	✗	✗	✓	✓
Lazarus (1996)[101]	✓	✓	✓	✓	✓	✓	?	?	✗	✗	?	✓
Marshall (1991)[102]	✓	✓	✓	✓	✓	✓	?	?	?	✓	✓	✓
Mast (2002)[103]	✓	✓	✓	?	✓	✓	?	?	✓	?	✓	✓
Reilly (1999)[104]	✗	✓	✓	?	✗	✗	?	?	?	?	✓	✓
Reilly (2000)[105]	✓	✓	✓	✓	✓	✓	?	?	?	✓	✗	✗
Schaefer (1998)[106]	✗	✓	✓	✓	✓	✓	?	?	✓	?	✓	✓
Wickramasinghe (2005)[107]	✗	✓	✓	✓	✓	✓	?	?	✓	✓	✓	✓

colleagues[97] age 8–12 years, Ellis and colleagues[99] age 3–18 years, Himes and Bouchard[100] age 8–18 years, Lazarus and colleagues[101] age 4–20 years, Marshall and colleagues[102] age 7–14 years, Schaefer and colleagues[106] age 6–19 years and Wickramasinghe and colleagues[107] age 5–15 years. None of the studies presented data separately for different ages but most grouped data by gender.[98–105,107]

Methodological quality

The results of methodological quality assessment are presented in *Table 10*. None of the 11 studies clearly met all of the 12 QUADAS criteria. The quality assessment was limited by the poor reporting of a number of studies. For two of the criteria, regarding independent interpretation of index test and reference standard, none of the studies reported enough information to assess whether the criteria had been met. In many studies it was not possible to assess a number of other criteria. There were some areas that the studies did address well. All studies described attempts to prevent verification bias by ensuring the whole sample or a random selection of the sample received the reference standard, and that they received the same reference standard irrespective of the index test result. Nine of the 11 studies stated that the index test and reference standard were independent[97,99–102,105–107,109] and eight gave sufficient details of the reference standard so that it could reasonably be replicated.[99,101–103,105–107] Fewer studies provided sufficient information regarding index tests and the methods used to define cut-off thresholds for the index test and the reference standard. The areas most poorly addressed, with fewer than half of the studies meeting the QUADAS criteria, were sample selection procedures, reporting of uninterpretable results and reporting of participants dropping out of the study.

Accuracy results

A number of different index tests and reference standards were used in the studies (*Table 11*). The index tests included BMI, individual skinfold thickness measurements (triceps, subscapular), combined skinfold measures and weight assessed as weight for height, relative weight and ideal

33

TABLE 11 *Index tests and reference standards*

Study	Index test(s)	Reference standard(s)
UK studies		
Reilly (1999)[104]	1. BMI >85th centile (E), SDs >2.00 (I) 2. Ideal body weight >120% (E)	1. %BF (skinfolds) >25% for boys, >32% for girls
Reilly (2000)[105]	1. BMI >a range of centiles from 74th to 98th (E), >IOTF obesity cut-offs, >IOTF overweight cut-offs	1. %BF (BIA) >95th centile (I)
North American studies		
Ebbeling (1999)[98]	1. BMI >85th centile (E) 2. Skinfold (triceps) >85th centile (E) 3. Weight for height >120% of the median (E) 4. Tests 1 and 2 combined	1. %BF (skinfolds) >25% for boys, >30% for girls
Ellis (1999)[99]	1. BMI >85th, >95th centiles (I)	1. %BF (DEXA) >85th, >95th centile (I)
Himes (1989)[100]	1. BMI >85th centile (E) 2. Skinfold (sum of 4) >85th centile (E) 3. Skinfold (subscapular) >85th centile (E) 4. Skinfold (triceps) >85th centile (E) 5. Weight >85th centile (E)	1. %BF (densitometry) >90th centile (E)
Marshall (1991)[102]	1. Skinfolds (sum of 5) >85th centile (E) 2. Relative weight >120% (E) 3. Relative BMI >120% (E) 4. Skinfold (triceps) >85th centile (I)	1. %BF (hydrostatic weighing) >20% for boys, >25% for girls
Other European studies		
Bedogni (2003)[97]	1. BMI >85th, >95th centiles (I) 2. Skinfolds (sum of 4) >85th centile (I)	1. %BF >85th centile (I)
Mast (2002)[103]	1. BMI >90th, >97th centile (E)	1. %BF (BIA) >90th, >97th centiles (I) 2. %BF (skinfolds) >90th, >97th centiles (I) 3. Skinfold (triceps) >90th, >97th centiles (I)
Schaefer (1998)[106]	1. BMI >85th, >90th, >95th centiles (I)	1. %BF (skinfolds) >75th, >85th, >90th, >95th centiles (I)
Australian studies		
Lazarus (1996)[101]	1. BMI >85th, >90th, >95th centiles (I)	1. %BF (DEXA) >85th centile (I)
Wickramasinghe (2005)[107]	1. BMI >95th centile (E), SDs >2.00 (E), >IOTF cut-offs for obesity	1. %BF (isotope dilution) >20% for boys, >30% for girls

(E), external reference data used; (I), derived internally from study population.

body weight. In some cases the obesity and overweight thresholds used were internally derived and in others they were taken from published reference data. Five studies investigated more than one index test[97,98,100,102,104] and six concentrated solely on BMI.[99,101,103,105–107] A variety of reference standards were used by the studies. All studies used reference standards based on %BF but different methods were used to derive this measure. These included calculations using skinfold measurements, BIA, DEXA, densitometry, hydrostatic weighing and isotope dilution methods. One study additionally used triceps skinfold thickness as a reference standard.[103] As with the index tests, a variety of diagnostic thresholds were used. Seven studies used one or more centile cut-offs with centiles either internally derived or taken from published reference data.[97,99–101,103,105,106] In four studies, specific %BF thresholds were used, and these were different for boys and girls.[98,102,104,107] The main results will be summarised in separate sections for index tests based on BMI, weight and skinfolds. Appendix 9 presents the full results of the studies of diagnostic accuracy.

BMI
BMI was used as an index test by all but one of the studies, but the thresholds used to define overweight and obesity varied. These included centile and SD cut-offs derived either internally from the study population or from national or

other published reference data and the IOTF definitions of obesity and overweight using BMI. Eight studies explored more than one overweight/obesity threshold for BMI.[97,99,101,103–107]

There were two UK studies, both conducted by Reilly and colleagues.[104,105] Reilly and colleagues (2000)[105] investigated BMI as a screening test for obesity by looking at a range of centile cut-offs, using UK 1990 reference data, along with the IOTF cut-offs for overweight and obesity. The reference standard used was %BF estimated using BIA with a diagnostic threshold of >95th centile derived from the study population of 3498 children. The optimum cut-off for BMI was found to be the 92nd centile, giving both a high sensitivity of 92% (95% CI: 0.87 to 0.95) and high specificity (95% CI: 0.91 to 0.92). For BMI >95th centile, a cut-off often used to define obesity, specificity was higher at 0.94 (95% CI: 0.88 to 0.94) but sensitivity was lower at 0.88 (95% CI 0.83 to 0.92). The IOTF BMI definitions of overweight gave comparable sensitivity and specificity to the 95th centile cut-off, but using the IOTF BMI definitions for obesity gave poorer sensitivity of 0.72 (95% CI: 0.63 to 0.80) for girls and only 0.46 (95% CI: 0.37 to 0.56) for boys. The study by Reilly and colleagues (1999)[104] was smaller, meeting fewer methodological quality criteria. The study looked at two BMI-based definitions of overweight/obesity in children compared with reference standards of >25% body fat (derived from skinfold measurements) in boys and >32% body fat in girls. The obesity definition BMI SDs >2.00, derived from the study population of 240 children, was found to have poor sensitivity of 0.36 (95% CI: 0.15 to 0.65) in boys and 0.60 (95% CI: 0.23 to 0.88) in girls, but specificity was high at 0.98 (95% CI: 0.94 to 1.00) and 0.99 (95% CI: 0.95 to 1.00), respectively. BMI >85th centile (BMI SDs >1.04) using UK 1990 reference data gave a higher sensitivity of 0.82 (95% CI: 0.52 to 0.95) in boys and 1.00 (95% CI: 0.57 to 1.00) in girls, but lower specificity at 0.88 (95% CI: 0.81 to 0.93) and 0.89 (95% CI: 0.82 to 0.94), respectively.

Wickramsinghe and colleagues[107] conducted a small Australian study to investigate three BMI-based obesity thresholds, two of which, those defined by the IOTF and the British Growth Standards, used the same reference data as the Reilly (2000) study,[105] and the other used US data and a threshold defined by the Centers for Disease Control and Prevention (CDC)/National Center for Health Statistics (NCHS). The reference standard was %BF estimated using an isotope dilution method, and the diagnostic threshold for obesity

was >20% body fat for boys and >30% for girls. Using the IOTF definitions of obesity, no cases of obesity were detected, so sensitivity was 0 and specificity 1. The other definitions did pick up some cases, but sensitivity was poor. For example, for white Caucasian girls, using a threshold of BMI SDs >2.00 according to British Growth Standards, sensitivity was 0.05 (95% CI: 0.01 to 0.25) and using the CDC/NCHS definition of BMI >95th centile it was 0.21 (95% CI: 0.09 to 0.43). Specificity was 1.00 (95% CI: 0.90 to 1.00) in each case.

Three other studies also used US reference data for BMI overweight/obesity thresholds.[98,100,102] A US study conducted by Ebbeling and colleagues[98] used BMI >85th centile (based on the NHANES I US study) to identify obesity in a sample of 1171 6–9-year-olds. Sensitivity was 0.84 (95% CI: 0.72 to 0.92) and specificity was 0.88 (95% CI: 0.85 to 0.91) for boys and 1.00 (95% CI: 0.92 to 1.00) and 0.86 (95% CI: 0.83 to 0.89) for girls, using %BF (skinfolds) with gender-specific cut-offs as a reference standard. Himes and Bouchard[100] conducted a study in Canada and also used BMI >85th centile, but based on a study of Cincinnati youths, to identify obesity in a sample of 316 8–18-year-olds. Sensitivity was low for both boys at 0.29 (95% CI: 0.14 to 0.50) and girls at 0.23 (95% CI: 0.12 to 0.41). Specificity was 0.99 (95% CI: 0.96 to 1.00) and 1.00 (95% CI: 0.97 to 1.00), respectively. In another Canadian study, Marshall and colleagues[102] considered relative BMI >120% of the median (based on NCHS data) as an index test, with gender-specific %BF (hydrostatic weighing) obesity cut-offs as the reference standard, in a sample of 540 7–14-year-olds. Sensitivity was similar in boys at 0.69 (95% CI: 0.54 to 0.80) and girls at 0.74 (95% CI: 0.57 to 0.86). Specificity was 0.93 (95% CI: 0.89 to 0.95), and 0.91 (95% CI: 0.87 to 0.94), respectively.

In a German study, Mast and colleagues[103] assessed BMI >90th and >97th centiles, using German reference data, compared with a range of reference standards, at the same centiles but derived from the study population, to identify overweight and obesity in a sample of 2286 5–7-year-olds. Specificity was similar, and over 0.90, in each comparison made. For both boys and girls, BMI for identifying obesity (threshold >97th centile) had the greatest sensitivity when compared with a reference standard of %BF estimated using BIA: boys 0.85 (95% CI: 0.70 to 0.94) and girls 0.79 (95% CI: 0.63 to 0.90). Similarly with the overweight threshold (>90th

centile), sensitivity of BMI with this reference standard for boys was 0.82 (95% CI: 0.73 to 0.88) and for girls 0.78 (95% CI: 0.69 to 0.85).

Four studies used the study population to define BMI threshold values for overweight/obesity internally.[97,99,101,106] In a US study, Ellis and colleagues[99] investigated BMI >85th and >95th centiles to define overweight and obesity, respectively, compared with %BF (DEXA) as a reference standard with the same centile cut-offs, also derived internally, in a sample of 979 3–18-year-olds. For defining overweight, BMI >85th centile had similar sensitivity in boys at 0.90 (95% CI: 0.80 to 0.95) and girls at 0.94 (95% CI: 0.87 to 0.97), and specificities of 0.83 (95% CI: 0.79 to 0.87) and 0.83 (95% CI: 0.79 to 0.86), respectively. The >95th centile cut-off however, was more sensitive for girls [0.90 (95% CI: 0.74 to 0.96)] than boys [0.71 (95% CI: 0.50 to 0.86)]. Specificity was increased to 0.92 (95% CI: 0.90 to 0.94) and 0.91 (95% CI: 0.88 to 0.93), respectively. In an Australian study, Lazarus and colleagues[101] used BMI thresholds of >85th, >90th and >95th centiles to identify overweight/obesity in a sample of 230 4–20-year-olds, with a reference standard of %BF (DEXA) >85th centile. The highest sensitivity was achieved with the 85th centile cut-off: 0.71 (95% CI: 0.54 to 0.83) for the whole sample, with a specificity of 0.95 (95% CI: 0.91 to 0.97). Using BMI >95th centile, sensitivity was reduced to 0.29 (95% CI: 0.17 to 0.46), and specificity was 0.99 (95% CI: 0.96 to 1.00). Results split by gender were similar. In a German study, Schaefer and colleagues[106] used the same BMI centile thresholds as the Lazarus study, and compared these with %BF (skinfolds) with several centile thresholds also internally derived. At BMI >85th centile, comparison with %BF >75th centile gave a sensitivity of 0.49 (95% CI: 0.45 to 0.53) and a specificity of 0.96 (95% CI: 0.95 to 0.97), and these changed to 0.81 (95% CI: 0.74 to 0.87) and 0.88 (95% CI: 0.87 to 0.89) when comparison was with the %BF >95th centile. Sensitivity was lower using the 95th centile cut-off for BMI: 0.20 (95% CI: 0.17 to 0.23) when compared with %BF >75th centile rising to 0.55 (95% CI: 0.46 to 0.63) when compared with %BF >95th centile. Specificity was 0.97 (95% CI: 0.96 to 0.98) and 0.99 (95% CI: 0.98 to 0.99), respectively. In an Italian study, Bedogni and colleagues[97] also investigated BMI >85th and >95th centile to identify obesity in a sample of 986 8–12-year-olds. The reference standard was %BF (BIA) >85th centile, internally derived. BMI >85th centile gave a sensitivity of 0.65 (95% CI: 0.57 to 0.72) and a specificity of 0.95 (95% CI:

0.93 to 0.96). At BMI >95th centile, these measures changed to 0.39 (95% CI: 0.32 to 0.48) and 0.99 (95% CI: 0.98 to 0.99), respectively.

Weight
In one of the UK studies, Reilly and colleagues[104] also considered a weight-based index test for obesity, >120% of ideal weight based on WHO reference values. Sensitivity of this measure, compared with gender-specific %BF (skinfolds) thresholds for obesity, was 0.64 (95% CI: 0.35 to 0.85) in boys and 1.00 (95% CI: 0.57 to 1.00) in girls, with specificities of 0.84 (95% CI: 0.76 to 0.90 and 0.89 (95% CI: 0.82 to 0.94), respectively.

Three North American studies also investigated weight-based index tests.[98,100,102] Marshall and colleagues[102] used relative weight >120% of the median using NCHS reference data. Compared with gender-specific %BF (hydrostatic weighing) obesity thresholds, sensitivity was 0.49 (95% CI: 0.35 to 0.63) for boys and 0.58 (95% CI: 0.41 to 0.74) for girls, with respective specificities of 0.95 (95% CI: 0.91 to 0.97) and 0.95 (95% CI: 0.92 to 0.97). Ebbeling and colleagues[98] looked at weight for height >120% of the median, again using NCHS data, alone and in combination with triceps skinfold (TSF) measurement, as measures of obesity. Compared to gender-specific %BF (skinfolds) obesity cut-offs, sensitivity for weight for height alone was 0.79 (95% CI: 0.66 to 0.88) for boys and 0.93 (95% CI: 0.77 to 0.98) for girls, with specificities of 0.97 (95% CI: 0.95 to 0.98) and 0.95 (95% CI: 0.93 to 0.97), respectively. Combining weight for height with TSF >85th percentile did not significantly alter these results. Himes and Bouchard[100] used weight >85th centile (data from a US study) as an index test compared to a reference standard of %BF (densitometry) >90th centile. Sensitivity of this measure was low in both boys at 0.43 (95% CI: 0.24 to 0.63) and girls at 0.17 (95% CI: 0.07 to 0.34). Specificity was 0.95 (95% CI: 0.90 to 0.98) and 0.98 (95% CI: 0.93 to 0.99) respectively.

Skinfold thickness
Four non-UK studies used index tests based on skinfold measurements.[97,98,100,102] Three North American studies investigated TSF measurements >85th centile as an index test for overweight/ obesity.[98,100,102] Ebbeling and colleagues[98] used TSF centile data from the NHANES I US study. Compared with gender-specific %BF obesity thresholds, also estimated from skinfold measurements, sensitivity was 0.94 (95% CI: 0.84 to 0.98) for boys and 0.98 (95% CI: 0.88 to 1.00) for girls, with respective specificities of 0.90 (95%

CI: 0.87 to 0.92) and 0.88 (95% CI: 0.85 to 0.90). Himes and Bouchard[100] also used TSF reference data but from a different US source. Compared with %BF (densitometry) ≥90th centile sensitivity was low for both boys at 0.24 (95% CI: 0.11 to 0.45) and girls at 0.23 (95% CI: 0.12 to 0.41) and specificities were 1.00 (95% CI: 0.97 to 1.00) and 0.97 (95% CI: 0.92 to 0.99), respectively. Marshall and colleagues,[102] derived the TSF 85th centile internally. Sensitivity compared with gender-specific %BF (hydrostatic weighing) reference standards was 0.64 (95% CI: 0.50 to 0.77) for boys and 0.68 (95% CI: 0.50 to 0.81) for girls. Specificity was 0.95 (95% CI: 0.91 to 0.97) and 0.93 (95% CI 0.90 to 0.96), respectively.

Himes and Bouchard[100] also considered subscapular skinfold >85th centile using US reference data. Sensitivity of this measure compared with a reference standard of %BF (densitometry) ≥90th centile was 0.38 (95% CI: 0.21 to 0.59) for boys and 0.30 (95% CI: 0.17 to 0.48) for girls, with respective specificities of 0.99 (95% CI: 0.96 to 1.00) and 0.99 (95% CI: 0.96 to 1.00).

Three studies investigated measures of obesity derived from several skinfold measurements combined.[97,100,102] Marshall and colleagues[102] used a measure based on a sum of five skinfolds and a threshold of >85th centile using the Canada Fitness Survey (1985) reference data. Compared with gender-specific %BF (hydrostatic weighing) reference standards, this measure gave a sensitivity of 0.80 (95% CI: 0.66 to 0.89) for boys and 0.91 (95% CI: 0.86 to 0.94) for girls, with specificities of 0.90 (95% CI: 0.85 to 0.93) and 0.97 (95% CI: 0.84 to 0.99), respectively. Himes and Bouchard[100] used a measure based on four skinfolds also with a threshold of >85th centile, but based on data from a US study. Sensitivity, using %BF (densitometry) ≥90th centile as a reference standard, was 0.57 (95% CI: 0.37 to 0.76) for boys and 0.80 (95% CI: 0.63 to 0.90) for girls with respective specificities of 0.85 (95% CI: 0.78 to 0.90) and 0.82 (95% CI: 0.74 to 0.88). Bedogni and colleagues[97] also used a sum of four skinfolds, but the >85th centile cut-off was derived internally. Sensitivity of this measure, compared with %BF (BIA) >85th centile, was 0.75 (95% CI: 0.67 to 0.81) and specificity was 0.94 (95% CI: 0.92 to 0.95).

Human resource requirements of growth monitoring programmes

Four studies reported on interventions related to the human resource issues of growth monitoring.[71,80,88,95] Three of these were related to the measurement of height[71,80,88] and one to both weight and height.[95] Three studies used data to determine the impact of recommended guidelines on measurement, charts or thresholds on subsequent referral.[78,89,90] Full details of these studies are presented in Appendix 10.

Three of the included growth monitoring programme studies provided cost data associated with delivery.[56,59,79] Nine of the included growth monitoring programmes used methods to ascertain or verify the height of children identified as meeting the predefined referral criteria.[56,59,71,76,79,81,82,91,93] Four of the growth monitoring programme studies reported on the human resource related data as part of the process of evaluating the monitoring programme.[56,82,87,96] Full details of the monitoring programmes are detailed in Appendix 7.

Costs

Three studies provided cost data associated with growth monitoring.[59,56,79] Banerjee and colleagues assessed the cost implications of implementation of monitoring and identifying children with a height below the 0.4th centile for referral in the Rhondda and Taff Ely area of Wales.[59] The approximate overall cost to the NHS Trust of delivering the programme was £14,550. This was based on measuring 2354 children aged 5 years in the school year of 1998–9 and a possible 10 sessions of growth monitoring per week. The cost per session was £50 for the health visitor and £25 for the nursery nurse; training costs were £1200 per academic session. Additional administration costs were £250 and miscellaneous costs were £2000.

Ahmed and colleagues conducted an earlier study of growth monitoring in the Oxford district of the UK and reported that the approximate annual cost was £10,000.[56] This was based on measuring the height of 20,338 children aged 3 and 4.5 years from 1988 to 1992. The cost of a part-time study coordinator, mileage, ongoing computer costs, stationery and postage, telephone usage, Minimeters, data entry, patient travel expenses, printouts and photocopying were considered. The costs were revised to incorporate the use of a triage system by the inclusion of an estimate of the auxologist's time and an initial consultant appointment. The costs were £9630 for 180 children found to be growing slowly. The cost of direct referral of 180 children to the paediatric endocrinologist at £90 a visit with a minimum of three visits to determine growth rate gave a total

of £48,600, suggesting that the use of a triage system is cost-effective.

Cernerud and Edding reported on the costs of a growth monitoring programme in Sweden.[79] The costs of measuring equipment used to measure the height and weight of more than 3000 children was low, being less than US$15 per instrument per year (1994 prices). The authors stated that personnel costs for the measurements represented 'a very small part of the salary for a nurse', although precise cost data were not given.

Accuracy of measurement

Eight of the included growth monitoring programmes reported on a process of re-measuring the height of children to confirm initial observations.[56,59,76,79,81,82,91,93] Three of these programmes did not report on the number of children with a height outside the referral threshold.[76,79,91] Five programmes reported the number of children with heights above the referral threshold,[56,59,79,81,82,93] which ranged from 5%[81] to 22% of children supposedly meeting the referral criteria.[93] However, this high value may in some cases have been due to the school nurses being instructed to refer children who were on or near the referral threshold.[41]

One of the included growth monitoring programme studies directly compared the accuracy of measurement taken by community staff to those of an auxologist.[96] The SD of the difference was 0.5 cm, meaning that 95% of the community staff's measurements fell within 1 cm of the auxologist's measurements. In addition, the Oxford and Hackney growth monitoring programmes undertook assessment prior to commencing the programme to determine the accuracy of those involved in the measurement and the measurement techniques used.[110,111] In the Utah growth monitoring programme, a project coordinator was employed to supervise 27% of the schools to ensure efficiency of the procedure, and unacceptable measurements resulted in more training.[87]

Training

Three studies provided information relating to training.[77,80,88] Lipman and colleagues conducted a multicentre RCT in the USA to determine whether training healthcare providers on the correct measurement technique and the use of equipment for measuring linear growth improves the accuracy of growth measurement.[88] The educational intervention consisted of being given appropriate measuring equipment and training sessions on how to use and install equipment and

on growth disorders. At baseline, approximately 30% of children were measured using an appropriate technique. Following the educational intervention, significantly more children were measured using the correct technique in the intervention group compared with the control group at 3 months (54% versus 23%; $p < 0.0005$) and at 6 months (74% versus 26%; $p < 0.0005$). Similarly, significantly more children were measured accurately in the intervention group compared with the control group at 3 months (55% versus 37%; $p = 0.003$) and at 6 months (70% versus 34%; $p < 0.005$). The accuracy of linear measurement (defined as within 0.5 cm of the value obtained by the study coordinators) also improved from 1.2 to 0.5 cm following the intervention.

Cowan and Gregory conducted a study in the UK to determine the effect of an intensive training programme on referral rates of children with abnormal growth from community-based staff.[80] Health visitors, school nurses and community paediatricians attended a 1-day training course in measurement technique and the importance of measuring and recording growth. Lectures were also given on the medical conditions associated with growth abnormalities. An audit of the clinic database before and after the study found that the number of referrals doubled during the 18-month period following the training intervention, although 66% of these were made in the first 6 months, suggesting that the impact on referral was not sustained.

The Oxford growth monitoring programme reported that despite providing health visitors with training in the use of Microtoises to measure height, 9.6% of children aged 3 years and 4.3% of those aged 4.5 years were measured with a wall chart and 23% of children aged 3 years and 5% of those aged 4.5 years were measured with a tape measure.[77]

Delivery of growth monitoring programmes

Three studies provided information on the delivery of growth monitoring programmes.[56,71,95] Aszkenasy introduced a new computerised form with feedback that did not require a graphical plot to determine whether the height of a child was below the predefined referral criteria for further investigation.[71] In the year before the form was introduced, only 43% of children with a height meeting the referral criteria were identified by the school nurse. This increased to 70% in the second year and 83% in the third year.

Welch and colleagues compared the effectiveness of a preschool assessment of height, weight, vision, hearing, blood pressure and dental caries by physicians with a school screening programme delivered by trained volunteers and public health nurses in the identification of abnormalities.[95] In the school screening programme, height and weight measurements were performed by physical education teachers and values were plotted by public health nurses. No significant differences were found in the number of height or weight abnormalities identified by the physicians' preschool programme and the school programme.

The Oxford growth monitoring programme reported that the use of a triage system consisting of community-based growth clinics minimised the number of false positives and over-investigation and reduced travel and anxiety for families. The programme found that 51% of children referred to the auxologist were found to be growing normally.[56] The potential benefit of this type of system was also commented on in the Hackney growth monitoring programme as allowing for initial assessment of short children prior to referral and leading to a higher rate of identification of cases of abnormal growth in those attending referral clinics.[82] Full details of these studies are presented in Appendix 7.

Impact of growth monitoring threshold on referral

Three studies considered the impact of growth monitoring thresholds on referral.[78,89,90] Van Buuren and colleagues used longitudinal data on the heights of children in The Netherlands to determine the number of referrals when adhering to proposed guidelines of the Dutch Institute for Health Care Improvement.[90] The proposed guidelines introduced six screening rules based on the height standard deviation score (HSDS). Boys younger than 10 years and girls younger than 9 years were referred if one or more of the following criteria were met:

1. HSDS is lower than –2.5 (absolute height SD).
2. HSDS is lower than –1.3 and HSDS is 1.3 lower than the target height SD (parental height corrected).
3. The growth curve deflects by more than 0.25 SD per year (deflection).
4. HSDS decreases by more than 1 SD over several years (slow SD loss).
5. In children born small for dates (birth length SD lower than –1.88), HSDS is lower than –1.88 after the age of two (no catch-up).

6. HSDS is lower than –1.3 and the child has disproportion or dysmorphic features (clinical features).

The results suggested that strict adherence to the proposed guidelines would lead to a large number of false positives, impair regular practice and create avoidable anxiety. Specifically, the percentage of referrals for the different criteria were: absolute height SD 6.2%, parental height corrected 5.9%, deflection 31.5%, slow SD loss 5.5%, no catch-up 1.4%, clinical symptoms no data available. Combining rules 1–4 would result in 38.2% of children measured being referred, corresponding to 77,000 children in The Netherlands each year. The authors stated that this figure was 30 times greater than expected.

Hearn and colleagues used data from the Hackney growth monitoring programme study[82] to determine the effect of the introduction of UK 1990 charts on the number of children with heights meeting the referral threshold.[78] The study also determined the height at diagnosis of children attending a local growth clinic for treatment for GHD. For children aged 5 years, the Tanner and Whitehouse charts identified 1% of children with a height below the 3rd centile and the UK 1990 charts identified 3.3% of children with a height below the 3rd centile and 0.47% below the 0.4th centile. For children aged 11 years, the Tanner and Whitehouse charts identified 2.1% and the UK 1990 charts identified 2.93% of children with a height below the 3rd centile and 0.34% below the 0.4th centiles. The authors therefore suggest that the use of the UK 1990 charts with a cut-off for referral for short stature below the 3rd centile would lead to an increase in workload of 2–3-fold, and the use of 0.4th centile would reduce workload by 50%. The latter may exclude a number of children with abnormalities from further assessment. Of 69 children with a diagnosis of GHD receiving treatment and height below the 2nd centile, 28 had a height at diagnosis between the 0.4th and 2nd centiles.

In a second UK-based study, Mulligan and colleagues used longitudinal data on 486 children to determine the impact on referrals of the introduction of guidelines recommending that all children have height measurements at the age of 5 years and again at between the age of 7 and 9 years.[89] The number of children who would require referral following the proposed guidelines was seven at school entry age (four children were <0.4th centile and three were >99.6th centile),

and at age 8 years there were no new referrals. Eleven children (2.3%) were considered to be 'slow growing' and nine (1.9%) children had an increase in height of more than one centile. Comparison of the community-based results with those obtained in an ideal research setting suggested that the variance of the change of height score was less when measurements were undertaken in the ideal research conditions. The authors concluded that adhering to the recommendations would not lead to an excessive number of inappropriate referrals. These findings highlighted the issue of inter-observer error.

Attitudes of children, parents and healthcare professionals to growth monitoring

Only one of the included studies specifically assessed the attitudes of children, parents or healthcare professionals to growth monitoring and that focused on monitoring for obesity.[72] One of the growth monitoring programme studies included an assessment of the attitudes of healthcare professionals to growth monitoring as part of its evaluation,[79] and three others provided additional attitudinal data.[87,91,93] Limited relevant attitudinal data were reported in the individual growth monitoring programme studies, therefore attendance at initial height measurement and referral were considered as appropriate surrogate outcomes and any reasons given for non-attendance were reported.[56,59,71,76,79,81–83,86,87,91,93,96] Details of the studies reporting attitude data are given in Appendix 11.

Attitudes to growth monitoring

One study specifically aimed to assess attitudes to growth monitoring.[72] Routh and colleagues gathered baseline data on the BMI of a sample of children in order to identify a low-cost method of detecting obesity that would be acceptable to schools. A total of 252 children aged 9–10 years participated in the one-off screen and three children (1.2%) refused to participate. However, parental consent was obtained on an opt-out basis by informing parents of the project and explaining reasons behind it. To avoid stigmatisation, no emphasis was given to weight and the lesson was designed around graphs and numbers rather than health, with weighing being one of many activities in the lesson. Each child was measured separately and readings remained private; 20% were found to be overweight and 7% were obese. A questionnaire was sent to teachers and school nurses to evaluate

the method. Four of the seven nurses delivering the programme and five of eight teachers whose classes participated returned the questionnaire. The results suggested a general satisfaction with the method of obtaining data on obesity. The weighing of children was said to have been conducted sensitively with no stigmatising of overweight children, and overall the lesson in which the monitoring took place was found to be enjoyable. It must be borne in mind, however, that the schools electing to take part in the programme may have differed from those not taking part in terms of attitudes to monitoring obesity. Furthermore, this is a small study, the results of which would need replicating with a larger sample, and the views of children were not directly assessed.

Attendance at initial measurement

Ten of the included growth monitoring programmes provided data on the number of children eligible for measurement and the actual number measured.[56,59,71,76,81,82,86,87,93,96]

Three monitoring programmes reported parental refusal as a reason for non-measurement.[56,59,96] In one of these studies, the authors stated that the main reason for refusal was that the parent was aware that their child was overweight and did not want attention drawn to this condition.[96] Other reasons for non-measurement included relocation of the child,[56,86] the child being absent from school on the day of measurement[87,91] or not attending or had been transferred to a different practice.[56] In some studies, the measurement was not recorded appropriately or data were incomplete[56,76,87,96] One study reported problems with school attendance and shortage of school nurses as a possible reason for low coverage[59] and one stated that there were problems with encouraging and reminding some of the health visitors to measure children.[56]

Three of the included growth monitoring programmes did not provide explicit details on the coverage or number of children eligible for measurement.[79,83,91] Full details on attendance at initial measurement are presented in Appendix 11.

Attendance at referral

Eleven of the included growth monitoring programme studies gave details of the number of children attending referral and the number meeting the criteria for referral.[56,59,71,76,79,81,82,86,87,91,93] Six of the included programmes stated that referral for further investigation was declined in children

meeting the referral criteria.[56,71,86,87,91,93] The numbers declining ranged from approximately 4%[87] to 17%.[91] Five programmes reported the number of children meeting the referral criteria who did not attend for further investigation,[56,59,71,81,82] ranging from approximately 4%[56] to 50%.[71] One study stated that reasons for non-attendance at further investigation needed to be explored.[71] In five of the growth monitoring programme studies, relocation was cited as the reason for not attending at referral[56,71,82,87,91]

Other attitude-related outcomes

Four further studies, not specifically designed to assess attitudes to growth monitoring, nevertheless provided further information.[79,87,91,93] In the Swedish programme described by Cernerud and Edding,[79] the benefits of growth surveillance were assessed by an expert panel consisting of experienced school nurses and senior paediatricians working as school doctors. The most important benefits were found to be the opportunity to meet children under structured and uncontroversial conditions, conditions such as malnutrition and drug abuse may be revealed during measurement, growth data can be used in discussions with children and their parents, and data are provided for future medical consultations and can be used for public health research.

In the Utah study,[87] the authors reported that 88% of children diagnosed with GHD and 60% of those diagnosed with TS had seen a physician on at least one occasion prior to the study and physicians had raised concerns regarding the heights of 60% of children with GHD and 67% of children with TS. Referral to an endocrinologist had been suggested for 50% of children with GHD and 17% of those with TS, but only 25% of those with GHD and no children with TS had been assessed by an endocrinologist. The authors stated that the reason why some children at risk of having a growth disorder may not be examined is lack of parental concern or refusal to follow up.

Vimpani and colleagues[91] noted that the parents of five of the six children with a new diagnosis of GHD, in whom genetic cause of disease could be excluded, had previously consulted their GP but none had been referred for further investigation.

In the Wessex study[93] described by Voss and colleagues, it was reported that some independent schools were not willing to participate as height measurements were not routinely made or it was perceived that height 'problems' were not their concern and should be dealt with by the schools' GP.[41]

Chapter 6
Economic evaluations and modelling

Any decision regarding the monitoring of child growth in schools ultimately needs to take into account its cost-effectiveness. Consequently, economic evaluations of short stature and obesity were reviewed and models were developed using the available information to evaluate the cost-effectiveness of child growth monitoring and obesity prevention programmes. GB and MR, the expert panel on the review team, provided information on health state utilities for the stature and obesity model, respectively. The reviewed economic evaluations are presented first.

Economic evaluations of short stature

Bryant and colleagues[21]

This study examined the use of GH treatment in children with GHD, TS, chronic renal failure (CRF), PWS and idiopathic short stature (ISS). The aim was to determine the effectiveness and cost-effectiveness of GH in increasing the final height of children in comparison with no GH treatment or placebo. Effectiveness data came from a systematic review of studies that examined biosynthetic human GH (somatropin). Outcomes included in the review were final height and short-term growth responses to treatment such as HSDS and height velocity. The review sought evidence from RCTs or systematic reviews of RCTs that assessed GH effects. Where data were not available from these study designs, evidence from the most reliable of the lower quality studies was used. Studies reporting quality of life data were also included in the review. Economic evaluations were included if they examined costs and consequences (health outcomes) for the intervention and the relevant comparator (no intervention or placebo). Data synthesis of the review and economic data was undertaken using modelling (in Excel) and was based on the best available evidence.

The study was conducted in Southampton, UK, with resource use data being derived from an expert panel, and unit costs from the NHS Reference Cost data set or Personal and Social Services Resource Unit (PSSRU) data.

No published economic evaluations or no studies reporting utility data suitable for modelling were found.

The findings of the systematic review indicated that both short-term growth and final height can be improved with GH treatment. In terms of short-term growth velocity at 1 year, gains ranged from zero to 1 SD above normal growth velocity for children of the same age. In terms of final height results, the results ranged from 2 to 11 cm (GHD, 8–11 cm; TS, 5 cm; CRF, 3–9 cm; PWS, 10–11 cm; ISS, 2–7 cm).

The cost data indicated that treatment with GH is expensive. The lifetime incremental cost (versus monitoring only) ranges from £43,100–53,400 for GHD to £55,500–83,000 for PWS. When applied to children aged 8–15 years for England and Wales, it would result in discounted costs of £904 million for complete treatment. The costs of treating only the four licensed conditions would be approximately £180 million.

The review presented synthesised results in the form of cost per centimetre gained. The ratio was £6000 for GHD, £16,000–17,400 for TS, £7400–24,100 for CRF, £13,500–27,200 for ISS and possibly in the region of £7030 for PWS (using 2000 prices).

In a range of sensitivity analyses, the variables included length of treatment (1–13 years), final height effect (10–300% of base case from trials), GH dose (varying by indication), GH cost (£15–25/mg), annual range of discounting of benefit (0–6%) and costs (0–12%). Results were most sensitive to effectiveness, GH dose and costs due to length of treatment.

The major limitations of this model were that the quality of studies providing final height data was low and data on the effectiveness of treatment were derived from studies in which children were treated relatively late and for a short period (5–8 years). Additionally, for comparison with other healthcare interventions, and in informing the present modelling study on child growth, the lack of a generic health outcome using quality-adjusted life-years (QALYs) is a limiting factor;

other conditions not targeted by a child growth monitoring programme may have varying clinical effectiveness outcomes such as final height reduction (for tall stature children) and reduction in co-morbidities (for nutritional obesity and managed conditions such as Marfan's syndrome).

The conclusions of the Bryant study were that GH is already prescribed in the UK but a full course of treatment is expensive. As only a minority of children with licensed conditions are currently receiving GH, the impact on the NHS budget of increased prescribing would be substantial. The most significant increase would be caused by prescribing GH to children with ISS, who, it should be noted, do not have any known underlying pathology but represent the subgroup of children with short stature.

Anthony and Stevens[73]

This study examined the cost–utility of growth hormone treatment in children with GHD, TS, CRF and ISS. The economic perspective was that of the UK NHS. The study population came from the south and west regions of the UK and comprised primarily pre-pubescent boys and girls (age range 1–13 years) with GHD, CRF, TS or ISS.

Effectiveness data were derived from a systematic review of the literature. Twenty-eight studies were included in the review; eight RCTs, one controlled trial and 19 case series designs.

Effectiveness was measured primarily in terms of height gain. Children treated for GHD achieved a mean of 2.6 SDs over an average of 5 years. Children treated for TS achieved a mean gain of 8.1 cm (3.2 inches) over original projections over a 5-year treatment period. Children treated for CRF achieved a mean gain of 1.48 SDs when compared with no treatment over a 2-year period.

The initial height of children was –2.94 SDs (control group –2.82 SDs). The final height of the treated group was –1.55 SDs (control group –2.91 SDs).

Treated children with ISS gained a mean of 1.1 SDs over a 3-year period. Their initial height was –2.7 SDs and final height was –1.6 SDs.

No significant differences were found in terms of psychological benefits when a treated group (of children with the entire range of disorders) was compared with a control group of ISS patients. Adverse events were not viewed to be a substantial consequence of the treatment.

The measure of benefit used in the economic analysis was QALYs. The authors used the Index of Health-related Quality of Life (IHQL) as the method of utility valuation for children with the target disorders. The authors stated that the greatest QALY gain from treatment (within the treatment period) would be approximately 0.1 and the least gain would be approximately –0.1. Given that treatment is given at an average age of 10 years, for approximately 4–6 years, the authors proposed that children would gain 0.5 QALYs in the best scenario and would lose 0.5 QALYs in the worst scenario.

The projected annual cost of GH treatment for a 9-year-old GHD child was calculated from a published source, using a resource use dosage rate of 15 units/m^2/week at 1995 prices. Children with TS, CRF and ISS were reported to require double this dose and all patients would require increased dosages at the onset of puberty.

The authors reported that the best scenario in terms of benefit was 1.5 more QALYs (assuming 15 years of benefit, 5 years of which were in the treatment period) and 5.5 more QALYS (assuming 55 years of benefit, 5 years of which were in the treatment period).

The annual average cost of treating a 9-year-old GHD child was reported to be approximately £7000, based on the lower dose regimen. The approximate cost for a 9-year-old child with TS, CRF or ISS would be £14,000.

The (best scenario) cost per QALY for GHD children was reported to be between £5700 and £20,800. The cost per QALY for children with TS, CRF and ISS was reported to be between £11,400 and £41,700 (assuming between 15 and 55 years of benefit).

The authors concluded that GH treatment should be recommended for children with short stature associated with GHD, TS and CRF. There is currently insufficient evidence to support the use of this treatment in the children with ISS. The few studies examining psychological benefits of the treatment presented conflicting results.

In terms of the limitations of this study, the measure of benefit used for the utility analysis was based on an estimate, and discounting was not applied.

The results, however, are clearly generalisable to the UK NHS and relevant to the present study.

The authors suggested that (given the high cost of GH treatment) future good-quality controlled trials are needed with longer follow-up periods in order to determine reliably the benefits of GH treatment. In addition, more research is required to address the motivation for, and expected benefits arising from the use of GH treatment.

Economic evaluations of obesity interventions

Wang and colleagues[74]

This study assessed the cost–utility and cost–benefit of 'Planet Health', a preventative school-based intervention designed to reduce the prevalence of obesity in youth of middle-school age (evaluated at age 14 years following a 2-year intervention). Intervention material was infused into four major subject areas (language arts, mathematics, science and social studies) and into physical education. Sessions focused on decreasing television viewing, decreasing consumption of high-fat foods, increasing fruit and vegetable intake and increasing moderate and vigorous physical activity. Comparisons were made with a no-intervention alternative, whereby students received the usual curricula and physical education classes. A societal perspective was adopted in the economic analysis.

The study population comprised male and female students of middle-school age but, as a *post hoc* subgroup analysis suggested that only female students were found to benefit from the programme (which was delivered to both males and females), the economic evaluation was restricted to females but included the costs of all students. Effectiveness and required epidemiological data for the modelling were derived from a number of sources: a randomised controlled study (five intervention schools and five control schools involving 1203 students) reporting efficacy data on the Planet Health intervention, a study predicting obesity in young adulthood from childhood and parental obesity and the National Health and Nutrition Examination Study Epidemiological Follow-up Study (NHANES I EFS).

A decision model was created to calculate the cost-effectiveness of the health intervention over 25 years. A two-stage overweight progression model was used to determine the expected number of adult overweight cases by age 40 years among the 310 female students in the intervention group, compared with the same 310 students in a hypothetical no-intervention group. Overweight was defined as a BMI of at least 25 kg/m^2.

The trial found that, during the 2-year intervention, the prevalence of obesity among girls declined from 23.6 to 20.4% in the intervention schools (n = 310), but increased from 21.5 to 23.7% in the control schools.

After controlling for baseline covariates, the prevalence of obesity among girls in the intervention schools was reduced significantly compared with girls in the control schools (odds ratio 0.47, 95% CI: 0.24 to 0.93; p = 0.03).

No significant differences were found among boys.

The measure of benefit used was QALYs gained. Direct costs included the costs of the intervention and medical costs of being overweight, which included the direct healthcare and medication costs associated with women who were currently 40 years of age and who maintained an overweight status through to age 65 years. The medical costs estimated were those associated with events of fatal and non-fatal coronary heart disease, hypertension, diabetes, symptomatic gallstones and osteoarthritis. Indirect costs due to lost productivity consisted of the costs associated with lost or impaired ability to work or to engage in leisure activities because of morbidity and lost economic productivity because of death. However, indirect costs were not included in the cost-effectiveness analysis (CEA).

Sensitivity analyses were undertaken by varying 10 parameters (e.g. the conditional probabilities of being overweight, the years of healthy life scores, the expected number of years of life and the annual number of work days lost).

The results of the modelling (extrapolation of the change in obesity prevalence into future health benefits) suggested that the number of QALYs saved due to the Planet Health intervention would be 4.13, and the intervention costs for Planet Health were US$33,677. The medical care costs that would be averted because of Planet Health were estimated at $15,887. The costs of lost productivity that would be averted because of Planet Health were $25,104. Hence the authors estimated that Planet Health would be associated with savings, from a societal perspective, of $7,313.

The incremental cost per QALY gained was $4,305 when Planet Health was compared with no intervention. Sensitivity analyses suggested that the results remained cost-saving to society under

most scenarios. The results of the Monte Carlo simulation resulted in 95% CIs between $1612 and $9010 per QALY saved.

The authors concluded that the Planet Health programme was cost-effective and cost-saving. The authors also concluded that school-based prevention programmes of this type were likely to be cost-effective uses of public funds. However, the conclusions of this study are clearly limited as the trial showed no overall benefit, and the modelling was based on the subgroup analysis which assumed a real benefit but only in female children.

Goldfield and colleagues[75]

The aim of this study was to determine the cost-effectiveness of two protocols for the delivery of family-based behavioural treatment for childhood obesity. The study was a mixed treatment comparison incorporating both group and individual treatment approaches. The comparator was group treatment only. The common components of the treatments were: a 13-session programme on diet, activity, behavioural change techniques, parenting and coping with psychosocial problems; the Traffic Light Diet; reinforcement for physical activity; self-monitoring; and stimulus control.

The economic perspective was that of the US health service and the setting was community. Effectiveness data were derived from a randomised controlled study involving 31 families, who were seen at 6 and 12 months after the treatment was started. The primary health outcomes used were the reduction in standardised BMI (Z-BMI) and the percentage overweight.

The benefit measures used in the economic analysis were the reductions in Z-BMI and percentage overweight.

At 12 months, a decrease of 0.005 percentage overweight units per US dollar was observed for the mixed group, compared with a decrease of 0.014 percentage overweight units per dollar with the group treatment ($p < 0.01$). At 12 months, a decrease of 0.0004 Z-BMI units per dollar was observed for the mixed group, compared with a decrease of 0.001 Z-BMI units per dollar with the group treatment.

The authors concluded that family-based behavioural treatment for childhood obesity was more cost-effective when provided in a group format than when provided using a combined group and individual approach. The cost-effectiveness of the treatment extended to parents.

The study was limited by its sample size and no power calculations were reported. Also, the follow-up period was short and therefore long-term outcomes were not assessed. The interventions were not assessed relative to a 'do-nothing' alternative, which limits the use of the results in modelling growth monitoring programmes. The authors stated that further research is needed to determine if the current results are generalisable to more obese children, since the population studied was mildly to moderately obese.

Quality assessment of included studies

Critical textual summaries are shown for each of the four studies in the commentary sections of each structured abstract in Appendix 12. The checklist, summary score and hierarchical matrix results for each included study are shown in *Table 12*.

TABLE 12 *Details of the included studies*

Question/study	Bryant (2002)[21]	Wang (2003)[74]	Goldfield (2001)[75]	Anthony (1996)[73]
1. The research question is stated	Y	Y	Y	Y
2. The economic importance of the research question is stated	Y	Y	Y	Y
3. The viewpoint(s) of the analysis are stated	Y	Y	Y	P
4. The rationale for choosing the alternative programmes or interventions are stated	Y	Y	Y	Y
5. The alternatives being compared are clearly described	Y	Y	Y	Y
6. The form of economic evaluation is stated	Y	Y	Y	Y
7. The choice of form of economic evaluation is justified in relation to the questions addressed	Y	Y	N	Y

continued

TABLE 12 *Details of the included studies (cont'd)*

Question/study	Bryant (2002)[21]	Wang (2003)[74]	Goldfield (2001)[75]	Anthony (1996)[73]
8. The source(s) of effectiveness estimates are stated	Y	Y	Y	Y
9. Details of the design and results of effectiveness study are given (if based on a single study)	N/A	Y	Y	N/A
10. Details of methods of synthesis or meta-analysis of estimates are given (if based on an overview of a number of effectiveness studies)	Y	N	N/A	Y
11. The primary outcome measure(s) for the economic evaluation are clearly stated	Y	Y	Y	Y
12. Methods to value health states and other benefits are stated	N/A	Y	N/A	Y
13. Details from the subjects from whom valuations are obtained are given	N/A	Y	N/A	P
14. Productivity changes (if included) are reported separately	N/A	Y	N/A	N/A
15. The relevance of productivity changes to the study question is discussed	N/A	Y	N/A	N
16. Quantities of resources are reported separately from their unit costs	Y	P	Y	Y
17. Methods for the estimation of quantities and unit costs are described	Y	P	N	Y
18. Currency and price data are recorded	Y	Y	P	Y
19. Details of currency of price adjustments for inflation or currency conversion are given	N/A	N/A	N/A	N/A
20. Details of any model used are given	Y	Y	N/A	N
21. The choice of model used and key parameters on which it is based are justified	Y	Y	N/A	N
Analysis and interpretation of results				
22. The horizon of costs and benefits is stated	Y	Y	Y	Y
23. The discount rate is stated	Y	Y	N/A	N
24. The choice of rate is justified	Y	Y	N/A	N
25. An explanation is given if costs or benefits are not discounted	N/A	N/A	N/A	N
26. Details of statistical test and confidence intervals are given for stochastic data	Y	N/A	Y	N/A
27. The approach to sensitivity analysis is given	Y	Y	N	Y
28. The choice of variables for sensitivity analysis is justified	Y	Y	N/A	Y
29. The ranges over which the variables are varied is stated	Y	Y	N/A	Y
30. Relevant alternatives are compared	Y	Y	N/A	P
31. Incremental analysis is reported	Y	Y	N	Y
32. Major outcomes are reported in both a disaggregated and an aggregated form	Y	Y	Y	Y
33. The answer to the study question is given	Y	Y	Y	Y
34. Conclusions followed from the data reported	Y	Y	Y	Y
35. Conclusions are accompanied by the appropriate caveats	Y	Y	P	Y
36. Generalisability issues are addressed	Y	P	N	N
Total Y + P	**28**	**30.5**	**16**	**22.5**
Total N/A	**7**	**3**	**14**	**4**
Percentage of applicable items (Y/36 − N/A) × 100	**96%**	**92.4%**	**73%**	**70%**

Hierarchical decision matrix showing direction of result (costs and effects)

Cost		Effect	Cost		Effect	Cost		Effect	Cost		Effect
+	R_1	–	+	R_1	–	+	R_1	–	+	R_1	–
0	R_2	–	0	R_2	–	0	R_2	–	0	R_2	–
+	R_3	0	+	R_3	0	+	R_3	0	+	R_3	0
–	A_1	–	–	A_1	–	–	A_1	–	–	A_1	–
0	A_2	0	0	A_2	0	0	A_2	0	0	A_2	0
+	A_3	+	+	A_3	+	+	A_3	+	+	A_3	+
–	G_1	0	–	G_1	0	–	G_1	0	–	G_1	0
0	G_2	+	0	G_2	+	0	G_2	+	0	G_2	+
–	G_3	+	–	G_3	+	–	G_3	+	–	G_3	+

Y, yes (score = 1); N, no (score = 0); N/A, not applicable; P, partial (score = 0.5).

The summary score for the Bryant study[21] was 96%, indicating that most items were adequately addressed. The study also has good generalisability to the UK NHS and reported data and results in a transparent way, making it useful in informing the modelling of growth monitoring strategies. The hierarchical decision matrix indicates that the intervention studied (GH treatment) would be associated with additional benefits (effects) and additional costs (cell A_3).

The Wang study[74] adequately addressed 92.4% of applicable points, and the hierarchical matrix result was cell A_3. This indicates that the intervention is (moderately) more costly but associated with health benefits when only the cost–utility analysis (CUA) is considered. When a societal perspective is considered, it is both cost saving and associated with health benefits, and the matrix cell is G_3. The major limitation of this study is that the results may not be generalisable to the UK due to differing cost structures and prevalence rates of obesity found in the USA. The authors highlighted the fact that more research is needed on the relationship between overweight status in children and obesity in adults and the QALYs and costs due to lost productivity of overweight and non-overweight adults. However, the model used may be adaptable for the UK setting and provide a means to assess provisionally the impact of growth monitoring when associated with primary prevention programmes to address obesity in school children.

The Goldfield study[75] adequately addressed 73% of the checklist items, and the decision matrix shows the interventions being equal in effectiveness with one intervention (group treatment alone) being less costly. This study would have been more useful in informing the present growth monitoring study if a 'do-nothing' control group had been included, the effectiveness being determined from baseline values and subsequent changes due to the interventions. The study has some limitations in terms of its generalisability to the UK context, but could potentially be adapted for the 'screen, find and treat' elements of growth monitoring to manage obesity, if UK-specific data were used. The hierarchical decision matrix for this study is cell G_1, indicating equal effectiveness and lower costs (cost minimisation).

Finally, Anthony and Stevens' study[73] adequately addressed 70% of the checklist items. Although clearly relevant to the UK context, major limitations were inadequate reporting of the model that was used to generate the results and that costs and benefits were not discounted.

However, this study did generate useful utility scores for short stature and these could be adopted for the growth monitoring model. The hierarchical decision matrix for this study is cell A_3, indicating increased benefits at higher costs.

NICE guidelines on GH treatment for children

Although not a formal economic evaluation according to our inclusion criteria, a NICE guideline on the use of growth hormone (somatropin) in children with growth failure[108] is summarised as it reports cost-effectiveness data from company submissions and UK data on prevalence and costs of treating children with GHD, TS, chronic renal impairment (CRI) and PWS.

Four models were submitted by manufacturers. One estimated an incremental cost-effectiveness ratio (ICER) of £3600 per centimetre gained in boys with GHD, £4264 in girls with GHD and between £6395 and £9215 in girls with TS. Another model estimated an ICER of £2118–2156 per normalised height year for GHD, £3560–4025 for TS and £1238–2339 for CRI. Normalised height was defined as less than 2 SDs lower than the relevant population mean.

Two manufacturers used the methods employed by the Wessex DEC study[73] to generate utility measures. The estimates from the manufacturers were between £5500 and £9000 per QALY gained in GHD, between £10,500 and £18,000 for TS and between £5000 and £11,000 for CRI. These models also adopted estimates for utility gains that were not based on primary data and therefore need to be treated with caution.

Of particular interest is the consideration of the NICE Committee: although they had reservations about the estimated utility gains due to treatment, the underlying assumptions were thought to be reasonable. The NICE Committee stated that "the utility gain from height gain in treatment of GH deficiency, TS, CRI or PWS was a worthwhile gain for the resources, given its lifelong value and the psychological importance to the child".

Overall diagnostic algorithm for monitoring and subsequent investigations

In order to scope potential modelling approaches for the current study, it was necessary to map the

overall healthcare problem in terms of monitoring growth and diagnosing and treating underlying causes of 'abnormal stature'. This process was undertaken with consideration of simultaneous growth monitoring for stature abnormalities and for obesity, according to the review protocol. The overall algorithm reveals a complex decision problem, as shown in *Figure 4*.

Four realistic strategies are considered:

1. *Growth monitoring for height and weight: find and treat short stature conditions AND obesity*. This strategy would utilise the most cost-effective method to monitor height and weight and refer children with a stature outside specified thresholds (short or tall) to a paediatrician or endocrinologist for diagnostic tests and treatment where relevant. The most common conditions associated with short stature are shown at the top of the algorithm [GHD, TS, ISS, PWS, CRF, abuse (psychological, physical or sexual), JH, coeliac disease, inflammatory bowel disease (Crohn's disease) cranial tumours]. 'Other' conditions completes the likely findings. The most common conditions associated with tall stature are precocious sexual maturation, Marfan syndrome and Klinefelter syndrome. Obese children, identified during monitoring, would be referred for an appropriate and cost-effective weight reduction intervention. Those found to be of normal height and not obese would not be referred for further investigations or treatment.
2. *Monitor short stature only AND provide primary prevention for obesity (all children)*. This strategy would be identical with (1) for stature monitoring and referral, but adopt, in parallel, a primary prevention programme that was administered to all children – similar to the Planet Health programme.[74] Cases of short stature and obesity may be expected to be investigated randomly over time due to parents' or GPs' concerns or observations.
3. *No growth monitoring and primary prevention for obesity*. This strategy would not monitor for either stature or obesity, but provide a primary prevention programme for obesity to all school children. This strategy may be justifiable if evidence shows limited new yield of cases related to short stature. Cases of short stature and obesity may be expected to be investigated randomly over time due to parents' or GPs' concerns or observations.
4. *No growth monitoring, no primary prevention for obesity*. This strategy is the comparator for the above and can be considered to be current UK

practice. Cases of short stature and obesity may be expected to be investigated randomly over time due to parents' or GPs' concerns or observations.

The choice of modelling questions

To inform the above strategies, the approach to modelling was considered in the light of two distinct scenarios and time frames. The first scenario would be based on a short-term cost per case detected analysis using standard diagnostic accuracy parameters (sensitivity, specificity), costs of tests and prevalence data for each condition. However, the very limited and incomplete diagnostic accuracy data derived from the systematic review did not permit this approach.

Long-term modelling based on probabilities of events, according to diagnostic yield from studies included in the systematic review, would be a second approach. As the data from the review permitted this approach, data were combined with effectiveness and cost data for treatments of the underlying conditions that were identified by the systematic review. It was therefore necessary to undertake additional searches to populate the model with long-term effectiveness and utility data.

As the monitoring timeframes, underlying causes and long-term health consequences of obesity and short stature conditions are dissimilar, along with the likely comparators necessary to assess the cost-effectiveness of growth monitoring and subsequent interventions, it was decided that two models would be built to address the following questions:

1. What is the most cost-effective approach to monitor stature and diagnose/treat underlying causes? This model is referred to as the 'stature model'.
2. What is the most cost-effective approach to monitor and treat, or prevent, childhood obesity in relation to avoiding long-term diseases associated with obesity? This model is referred to as the 'obesity model'.

Methods – long-term modelling

The long-term modelling incorporates diagnostic yield, treatment effectiveness, monitoring costs, treatment costs and QALYs for those underlying causes of short stature and obesity for which data are available. Because the modelling, in effect,
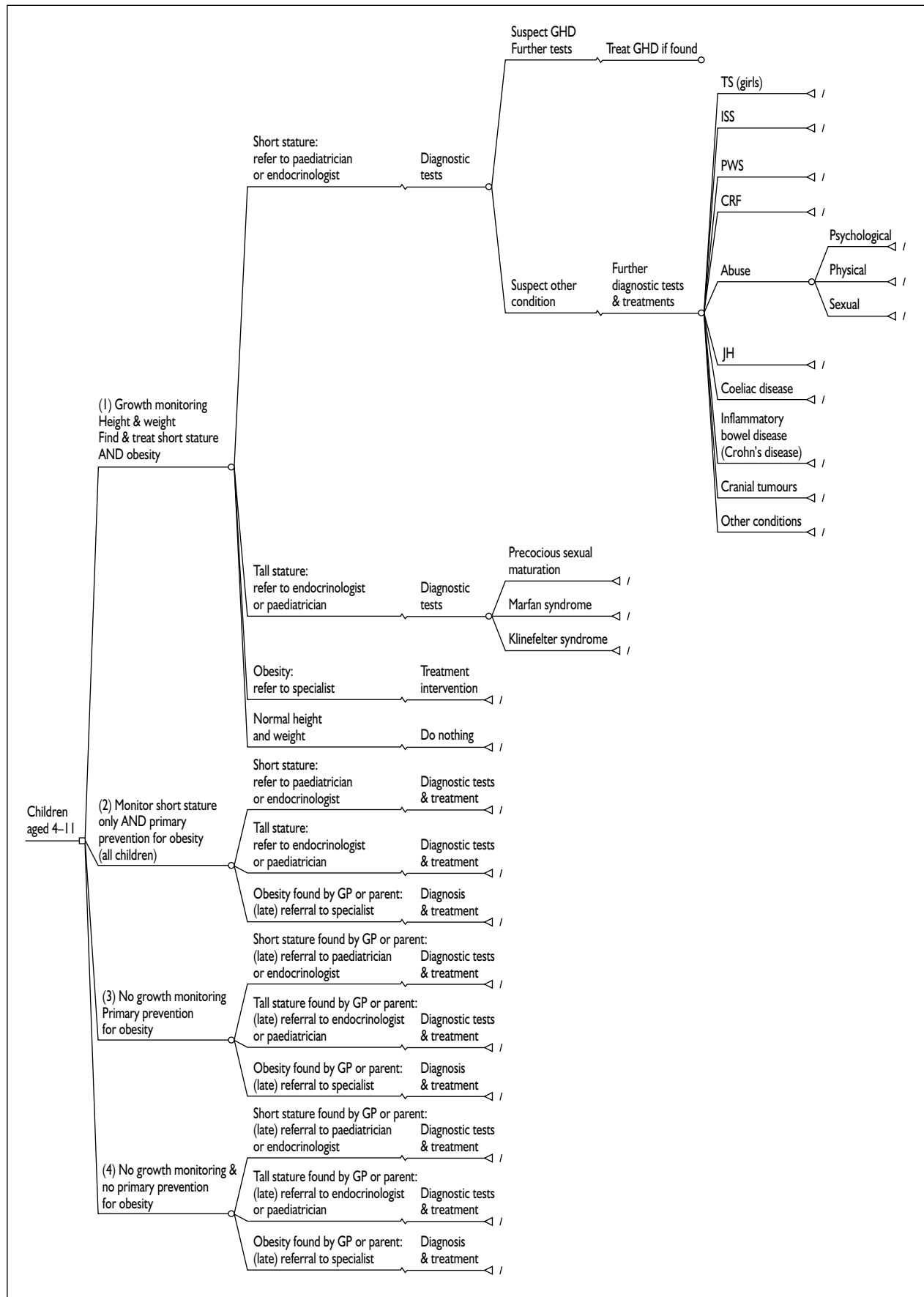
FIGURE 4 *Overall diagnostic and treatment algorithm for monitoring*

involves embedded decisions regarding the most appropriate or cost-effective treatment for included conditions, the most appropriate treatment and associated resource use data were used.

Following, as closely as feasible, the NICE guidelines for economic modelling,[112] the principal aim of the long-term modelling is to provide an answer, however tentative, to the question of whether or not the introduction of a growth monitoring programme offers good value for money for the NHS. This can only be measured in terms of the relative costs and health benefits associated with the identification of growth abnormalities, diagnosis and treatment of target conditions, in comparison with the status quo (no monitoring programme).

Consistent with the review protocol, primary prevention interventions for obesity were considered as part of growth monitoring strategies such that all plausible alternatives could be evaluated. In this regard, it should be noted that, as the prevalence of obesity is known to be increasing, a policy decision may be made to implement primary prevention interventions similar to the Planet Health[74] programme included in the review of economic evaluations. Indeed, many initiatives are under way in the UK to prevent childhood obesity. Therefore, a growth monitoring programme would have value in determining whether or not the prevalence of obesity was decreasing and to assess the effectiveness of programmes to either prevent or identify and treat obesity.

Due to the wide scope of the modelling it was decided to adopt the following approach:

1. Use existing models available in the literature to update and extend, where feasible. The models used were Wang (primary prevention of obesity),[74] Goldfield (monitor for and treat obesity)[75] and Bryant (diagnose and treat short stature with growth hormone drugs).[21]
2. There should be an effective intervention that can justifiably be used in the modelling (i.e. one intervention can be selected where there are multiple treatment options for a particular condition).
3. The prevalence of the condition should be high enough to warrant its inclusion in the modelling (i.e. very rare conditions will have very little impact on the overall cost-effectiveness of a growth monitoring programme and can therefore be excluded).

4. Obtain estimates of health state utilities from the literature, or estimates by the clinical experts, for all identified conditions (treated and untreated). The Wang study[74] was used, which included QALYs as the health outcome. The Goldfield study[75] was combined with branches of the Wang model to determine QALYs for the monitor and treat obesity strategies. QALY estimates for the stature model were based on the other included economic evaluation.[73]

The analysis took the form of a CUA from the NHS perspective, covering lifetime costs and QALYs associated with each of the models' strategies. Discounting of costs and benefits was undertaken, when relevant, using the rate recommended by NICE (3.5%).[113] The models were run using a hypothetical cohort of 594,000, which represents the number of 5-year-old children in England and Wales[114] and therefore the results reflect the likely impact on the NHS in terms of benefits and costs. All modelling was undertaken using a combination of Microsoft Excel and TreeAge Professional 2005 (TreeAge Software, Williamstown, MA, USA).

Sensitivity analysis

Probabilistic sensitivity analyses[115] using Monte Carlo simulations (1000 iterations) were conducted by fitting appropriate distributions to variables where this was feasible. For all other variables, point estimates were used. This enabled ranges to be generated for all evaluated strategies, with uncertainty being graphically presented using cost-effectiveness acceptability curves (CEACs)[116] and scatter plots.

Conditions included in the modelling

The included conditions were determined by examining those that were found in growth monitoring studies. Conditions that were found very rarely or not at all were excluded. The conditions included, therefore, are:

1. three conditions associated with short stature that can be treated with GH therapy, namely:
   (a) GHD
   (b) TS in girls
   (c) ISS (treatment, however, is controversial)
2. short stature caused by JH
3. short stature caused by psychosocial factors (one generic category)
4. other conditions (relevant or otherwise) found as part of a stature monitoring programme
5. childhood obesity and associated long-term health outcomes.

TABLE 13 *Lifetime QALY gains for included short stature conditions*

Condition	Early treatment	Source	Late treatment	Source
GHD	5	73, 117	2.5	Expert panel
TS	5	73, 117	2.5	Expert panel
ISS	2.5	73, 117	1.25	Expert panel
JH	2.5	Expert panel	1.25	Expert panel
Psychosocial	5	Expert panel	2.5	Expert panel

As can be seen from the heterogeneity in diagnostic and treatment pathways, effectiveness measures, health benefits and costs, producing a credible long-term modelling solution represents a major challenge.

Quality-adjusted life-years for the stature model

In order to populate the stature model with utility values a supplementary review of the literature was undertaken (see Appendix 3 for details of the search strategy). The findings were that many studies had assessed quality of life (QoL) in children with short stature, but only two provided data that could be employed in the stature model.

The study by Anthony[73] utilised an estimate of 5.5 additional QALYs in the most optimistic scenario for children treated with GH in achieving a more normal final height. The literature that was reviewed mostly found that children who are 'normal shorts' do not experience any significant fall in QoL, whereas children with TS and CRF were willing to trade off time to gain additional QALYs due to final height gain.[117] Empirical evidence from the latter study showed that women with TS, individuals with renal failure and those with ISS expressed a wish to be taller, with an estimated reduction in QoL of 2–4%. The general finding was that those who are 'normal shorts' did not exhibit any signs of reduced QoL, and as such the authors 'falsify' the assumption of a direct relationship between short stature and QoL. The reduction in QoL for ISS is probably due to unsuccessful coping strategies, whereas for patients with TS and CRF reductions in QoL are due to co-morbidities such as infertility in TS and kidney problems (renal failure after transplant or renal failure during dialysis) in CRF.

In order to analyse the impact of the model in relation to early detection and treatment through monitoring, it was necessary to estimate a reduced number of QALYs for the comparator (no monitoring). In this case, it is assumed that children will be referred at a later stage and receive appropriate treatment for a shorter period with a reduction in effectiveness and health gain. In the most optimistic scenario for growth monitoring, children may never be referred and treated, in which case their QALY gain would be zero. However, it is more likely that they will be referred at some later stage compared with monitoring, and therefore a 50% reduction in lifetime QALY gain was estimated by the expert panel advising the review (GB for conditions of stature and MR for obesity).

Based on the above findings and considerations, the estimates in *Table 13* were used to populate the model for children found to be short who were treated early with GH drugs, treatment for JH and interventions for children suffering from psychosocial causes of short stature.

Stature model

The structure for the short stature model is shown in *Figure 5*.

The model comprises two options:

1. Monitor for short stature at entry to school (age 5 years) and refer to paediatrician or endocrinologist all pupils below the chosen threshold. Diagnose and treat patients according to their underlying condition.
2. No monitoring. Short stature children will be referred in an *ad hoc* manner by either a GP in normal clinical practice or by concerned parents. This will involve a delay in diagnosis of underlying conditions and reduced costs (due to shorter treatment) and reduced QALY gains.

Input parameters for stature model
Pooled probability estimates from review
Table 14 details how the probabilities of all outcomes for the model, as derived from the raw data from the systematic review were calculated [see the section 'Detection of growth-related conditions and comparison of detection rates' (p. 23) and
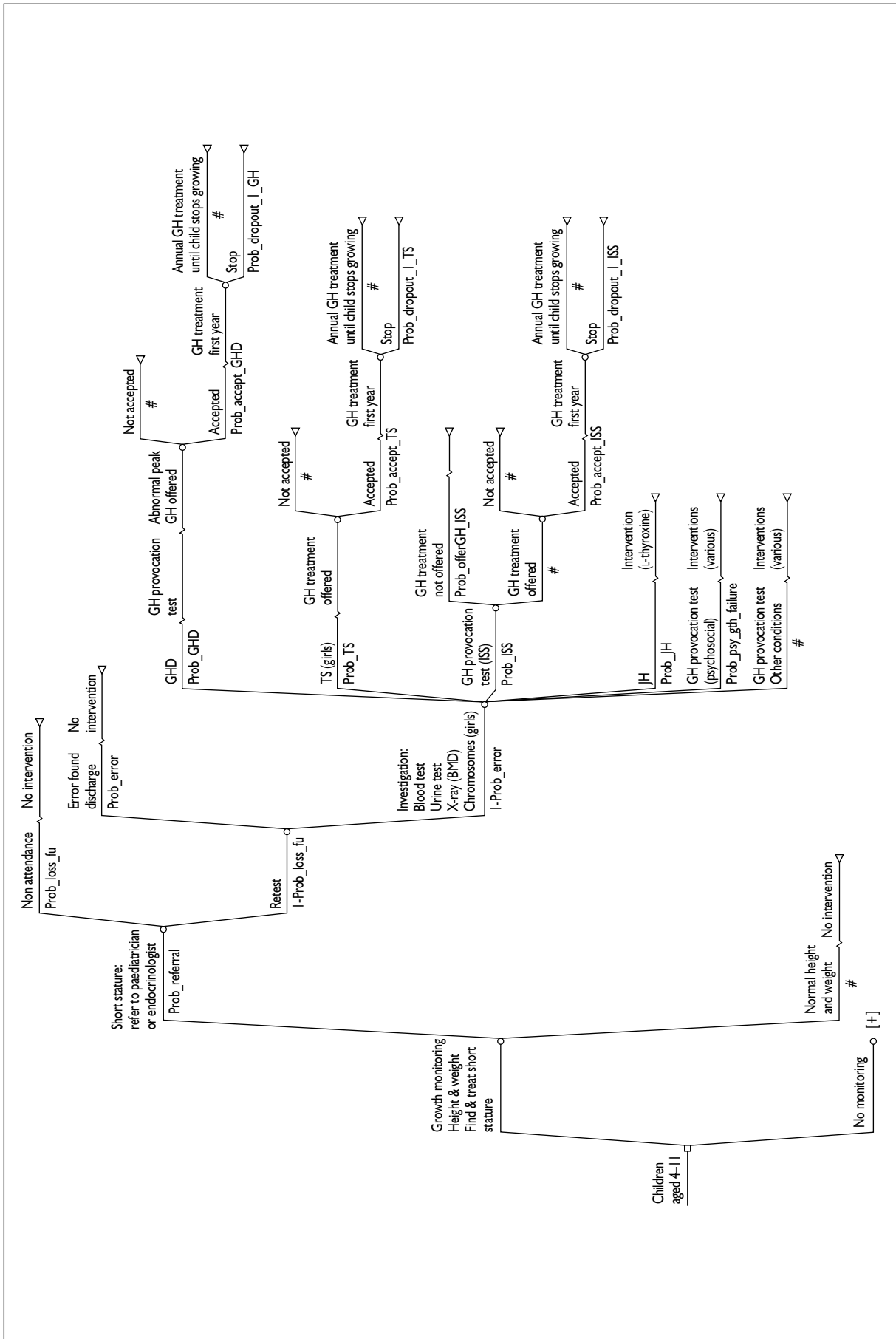
FIGURE 5 *Structure of stature model*

Appendix 8 for further details of the results of the systematic review]. The pooled totals for each condition were calculated and these data formed the raw values for beta and Dirichlet distributions used in the stature model. Pooled means were calculated using all included studies and the following equation:

$$\text{pooled probability} = \text{SUM(events)}/\text{SUM(sample sizes)}$$

In order to generate distributions for the probabilistic sensitivity analyses, beta distributions were chosen for the probabilities of referral, loss to follow-up and measurement error. Taking the probability of referral as an illustrative example (Prob_referral in *Table 15*), the beta distribution (which is bound by 1 and therefore appropriate for probabilities) requires two variables: the event rate, *r*, and the sample size, *n*. These values for the probability of referral are 2184 and 290,507, respectively.

At the chance node in the stature model where underlying causes (GHD, TS, ISS, JH, psycho-social, other) are distributed, it was necessary to fit the Dirichlet distribution[118] as this generates random distributions for more than two branches. For each branch, raw data for all conditions (branches) are used. To illustrate, the parameters used to determine the probability of the underlying causes of short stature listed above, after relevant diagnostic tests are (70;11;957;7;8;711), where included parameters sum to the relevant sample size appearing at the chance node (in this case 1764 as shown in *Table 14*, with the numbers of ISS and psycho-social diagnoses scaled up accordingly). The probabilities at all other branches are calculated in a similar manner.

Point estimates for a number of probabilities for the stature were taken from Bryant and colleagues' study,[119] as shown in *Table 15*.

Cost estimates
Table 16 gives an overview of cost data used in the model. The cost data for the stature model were derived from a number of sources and, where relevant, reflated to 2004 values using the Consumer Price Index available from the PSSRU data set.[120]

In order to ensure that the model reflected costs from an NHS perspective, a hypothetical cohort was used that reflected the birth cohort for 2001, based on the Census for that year. The figure used was 594,000.[114]

For GH-related treatment, cost data (including hospital admissions) were taken from the Bryant study.[21] A chromosome test (blood karyotype) was also included for girls, and therefore applied to half of the model cohort, to reflect the fact that all girls are tested in this manner for TS. The cost for this test was taken from the Department of Health Reference Cost source using the closest category.[121]

For children found to have psychosocial problems a generic cost estimate was used, based on fostering a child for either 6 or 12 years, depending on the point at which they were detected (early or late). Estimates were based on the best available data, which covered the allowances received by foster parents from the UK Social Services.[122]

The cost of the monitoring programme was based on two UK studies and the Department of Health Reference Cost data set.[59,121,123] Due to the wide range in estimates, based on differences in personnel used in the original studies and the inclusion of administration and training for health professionals carrying out the monitoring, a uniform distribution was used in the sensitivity analysis based on ranges from the included studies.

The cost of treating 'other' conditions detected by the monitoring programme (such as asthma) were assigned a cost value of zero in the model as the expert panel assumed that these conditions would be treated (or were being treated) as part of standard clinical practice.

Discounting was applied to relevant cost data where it extended beyond 1 year, according to NICE guidelines.

Results for stature model
The mean and incremental baseline results using probabilistic sensitivity analysis, along with ranges of costs and effects, are shown in *Table 17*.

The uncertainty in the ICER result (baseline = £9500 per QALY) is reflected in the incremental cost-effectiveness scatter plot in *Figure 6*.

The uncertainty in the cost and effectiveness outcomes for each strategy is reflected in the cost-effectiveness scatterplot in *Figure 6*.

The top, right-hand cloud shows the distribution of costs and QALYs for the monitoring programme and the bottom, left-hand distribution is the no-monitoring option. As there is almost complete horizontal and vertical separation

TABLE 14 Probability of outcomes derived from the review

Study	Referral No.	Referral Total	Loss to follow-up No.	Loss to follow-up Total	Measurement error No.	Measurement error Total	GHD No.	GHD Total	TS No.	TS Total	ISS No.	ISS Total	JH No.	JH Total	Psychosocial No.	Psychosocial Total
Cernerud (1994)[79]	20	7,129	0	20			0	20	0	20	0	20	0	20	0	20
Agwu (2004)[76]	18	3,465	4	18			2	14	0	14	5	14	–	14	–	14
Ahmed (1995)[56]	260	20,338	76	260	35	260	1	149	0	149	69	149	0	149	–	149
Azkensay (2005)[71]	24	9,338	15	24	10	24	0	9	0	9	9	9	0	9	–	9
Lacey (1974)[86]	111	2,256	13	111			1	98	0	98			–	98	0	98
Voss (1992)[93]	180	14,346					1	156	0	156			0	156	0	156
Hearn (1995)[82]	146	9,549	2	100	25	100	2	59	0	59	48	59	–	59	2	59
Vimpani (1981)[91]	449	48,221	91	449			9	358	1	358			1	358	0	358
Lindsay (1994)[87]	630	114,881	75	630			16	555	6	555			3	555		
Keller (2002)[83]	346	60,984					38	346	4	346	193	346	2	346	0	346
Pooled totals	**2,184**	**290,507**	**276**	**1,612**	**70**	**384**	**70**	**1,764**	**11**	**1,764**	**324**	**597**	**8**	**1,764**	**5**	**1,209**
Pooled mean	**0.008**		**0.171**		**0.18**		**0.04**		**0.006**		**0.02**		**0.0045**		**0.004**	

TABLE 15 *Effectiveness and benefit data for the stature model*

Variable name	Description	Estimate	Distribution	Source
Prob_accept_GHD	Probability of accepting GH treatment for patients diagnosed with GHD	0.91	Point	21
Prob_accept_ISS	Probability of accepting GH treatment for patients diagnosed with ISS	0.71	Point	21
Prob_accept_TS	Probability of accepting GH treatment for patients diagnosed with TS	0.83	Point	21
Prob_dropout_1_GH	Probability of a GHD patient dropping out of GH treatment after 1 year	0.093	Point	21
Prob_dropout_1_ISS	Probability of an ISS patient dropping out of GH treatment after 1 year	0.29	Point	21
Prob_dropout_1_TS	Probability of a TS patient dropping out of GH treatment after 1 year	0.17	Point	21
Prob_error	Probability of a measurement error at initial monitoring	0.182	Beta, $n = 384$, $r = 70$	See *Table 14*
Prob_GHD	Probability of having GHD after diagnostic tests	0.040	Dirichlet, List(70; 11;957;7;8;711)	See *Table 14*
Prob_ISS	Probability of having ISS after diagnostic tests	0.543	Dirichlet, List(957; 70;11;7;8;711)	See *Table 14*
Prob_JH	Probability of having JH after initial diagnostic tests	0.0045	Dirichlet, List(8;70; 11;957;7;711)	See *Table 14*
Prob_loss_fu	Probability of being lost to follow-up after referral from growth monitoring	0.171	Beta, $n = 1612$, $r = 276$	See *Table 14*
Prob_offerGH_ISS	Probability of being offered GH treatment after a diagnosis of ISS	0.0238	Point	21
Prob_other	Probability of having other conditions following a monitoring programme and diagnostic tests	0.403	Complement based on Dirichlet probabilities	Systematic review (*Table 14*)
Prob_psy_gth_failure	Probability of having psychosocial conditions after initial diagnostic tests	0.0040	Dirichlet, List(7; 70;11;957;8;711)	See *Table 14*
Prob_referral	Probability of being referred after a monitoring programme	0.0075	Beta; $n = 290,507$, $r = 2184$	See *Table 14*
Prob_TS	Probability of having TS after diagnostic tests	0.0062	Dirichlet, List(11; 70;957;7;8;711)	See *Table 14*
QALY_max	Maximum lifetime QALY gain for early treatment of TS and GHD	3.75	Uniform, 2.5–5	See *Table 13*
QALY_max_0_5	Half the maximum lifetime QALY gain for late treatment of TS and GHD	1.875	Uniform, 1.25–2.5	See *Table 13*
QALY_min	Lifetime QALY gain for late treatment of TS and GHD	0.938	Uniform, 0.625–1.25	See *Table 13*
QALY_min_0_5	Half the QALY gain for late treatment of ISS, psychosocial causes and JH	0.469	Uniform, 0.312–0.625	See *Table 13*

between the two strategies, it is possible to conclude that the monitoring programme is always more effective and always more costly.

Figure 7 shows the distribution of the samples of incremental cost-effectiveness comparing between the two strategies. The uncertainty in the model may be reflected in the percentage of time that monitoring is cost-effective given a decision-maker's willingness to pay (WTP) threshold; at a WTP of £30,000 per QALY, for example (generally taken to be the UK upper value), the monitoring programme is cost-effective over 100% of the distribution in *Figure 7*.

Figure 8 shows the CEAC for the model. It can be interpreted as follows: at a WTP threshold below £9500, the no-monitoring option is more cost-effective, but as the WTP increases the monitoring programme becomes the optimal choice. At a WTP of £20,000 per QALY, the probability of monitoring being cost-effective is equal to one.

TABLE 16 *Cost data for the stature model in 2004 prices*

Variable name	Description	Value (£)	Distribution	Source
Pop_Eng_Wales	The population size of 5-year-olds in England and Wales	594,000	Point	114
Cost_blood_test	Cost for a blood test (full blood count, chemical profile, thyroid, IGF) at an outpatient visit	22.6	Point	21
Cost_chromosome_test	Cost of a chromosome test (blood karyotype) for all girls who are referred	185		121
Cost_day_admission	Cost of a hospital day admission (for GH provocation test)	142.7	Point	21
Cost_district_nurse	Cost of a district nurse visit	37.4	Point	21
Cost_fostering_long	Average discounted cost of fostering for 12 years	83,189	Point	122
Cost_fostering_short	Average discounted cost of fostering for 6 years	45,872	Point	122
Cost_GHD_long	Cost of treating a child, detected early, with GHD (8 years)	39,523	Point	21
Cost_GHD_short	Cost of treating a child, detected late, with GHD (4 years)	18,109	Point	21
Cost_GHprov_test	Cost of a GH provocation test	330.7	Point	21
Cost_GH_GHD_annual	Average annual cost of treating a GHD patient with GH drugs	6,463	Point	21
Cost_investigation_pit	Cost of a pituitary function test (performed on 10% of GHD patients annually)	390.7	Point	21
Cost_ISS_annual	Average annual cost of treating an ISS patient with GH drugs	10,485	Point	21
Cost_ISS_long	Average discounted cost of treating a child, detected early, with ISS (8 years)	66,357	Point	21
Cost_ISS_short	Average discounted cost of treating a child, detected late, with ISS (4 years)	30,404	Point	21
Cost_JH_long	Average discounted cost of treating a child, detected early, with JH (8 years)	2,833	Point	121,124
Cost_JH_short	Average discounted cost of treating a child, detected late, with JH (4 years)	1,562	Point	121,124
Cost_mon_programme	Cost per child of a growth monitoring programme	6.75	Uniform, 3.5–10	59,121,123
Cost_MRI	Cost of an MRI scan in the diagnosis of GHD	142.7	Point	21
Cost_other	Cost of treating other cases found as part of a monitoring programme	0	Point	Assumption
Cost_outpatient_visit	Cost of an outpatient visit for referred children	109.8	Point	21
Cost_specialistNurse_visit	Cost of a specialist nurse home visit for patients commencing GH treatment	74.8	Point	21
Cost_TS_annual	Average annual cost of treating a patient with GH who has TS	10,724	Point	21
Cost_TS_long	Average discounted cost of treating a child, detected early, with TS (8 years)	67,872	Point	21
Cost_TS_short	Average discounted cost of treating a child, detected late, with TS (4 years)	31,098	Point	21
Cost_urine_test	Cost of a urine test	4.5	Point	21
Cost_Xray_hand	Cost of a hand X-ray	13.6	Point	21
Discount_rate_cost	Discount rate for costs	0.035	Point	113

TABLE 17 *Summary of baseline cost-effectiveness results*

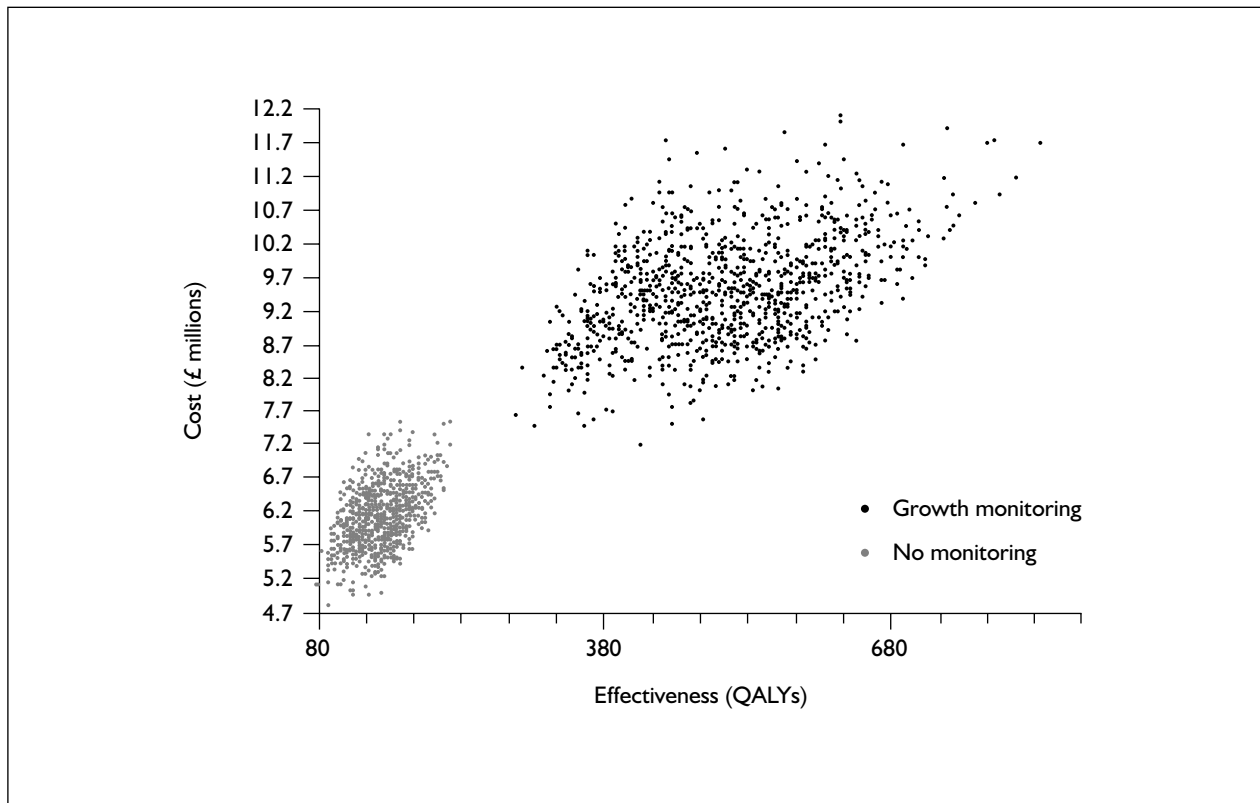|	Monitoring		No monitoring	
	Cost (£ million)	QALY	Cost (£ million)	QALY
Lower CI	8.1	329	5.3	104
Upper CI	11.1	690	7.1	197
Mean and incremental cost-effectiveness				
Mean	9.5	501	6.1	148
Incremental	3.4	352		
Mean ICER	£9500 per QALY			

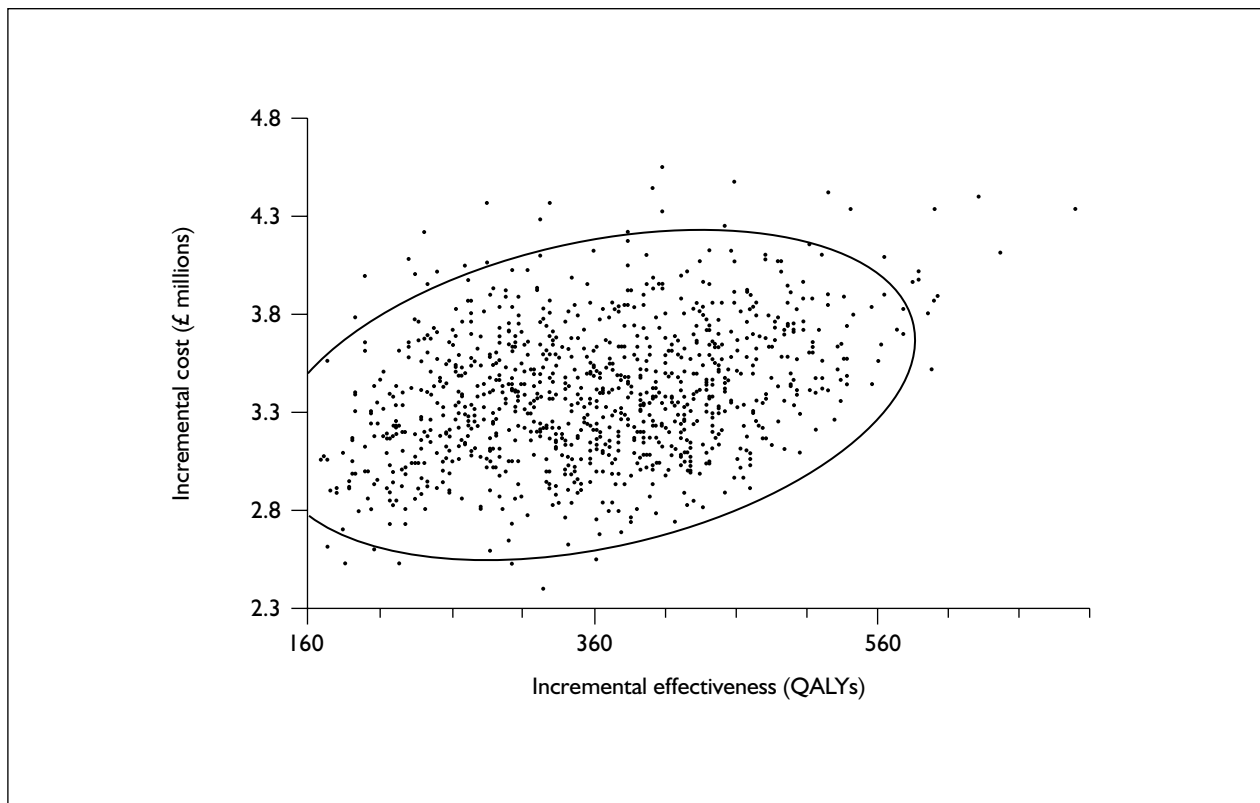**FIGURE 6** *Distribution of costs and effects*

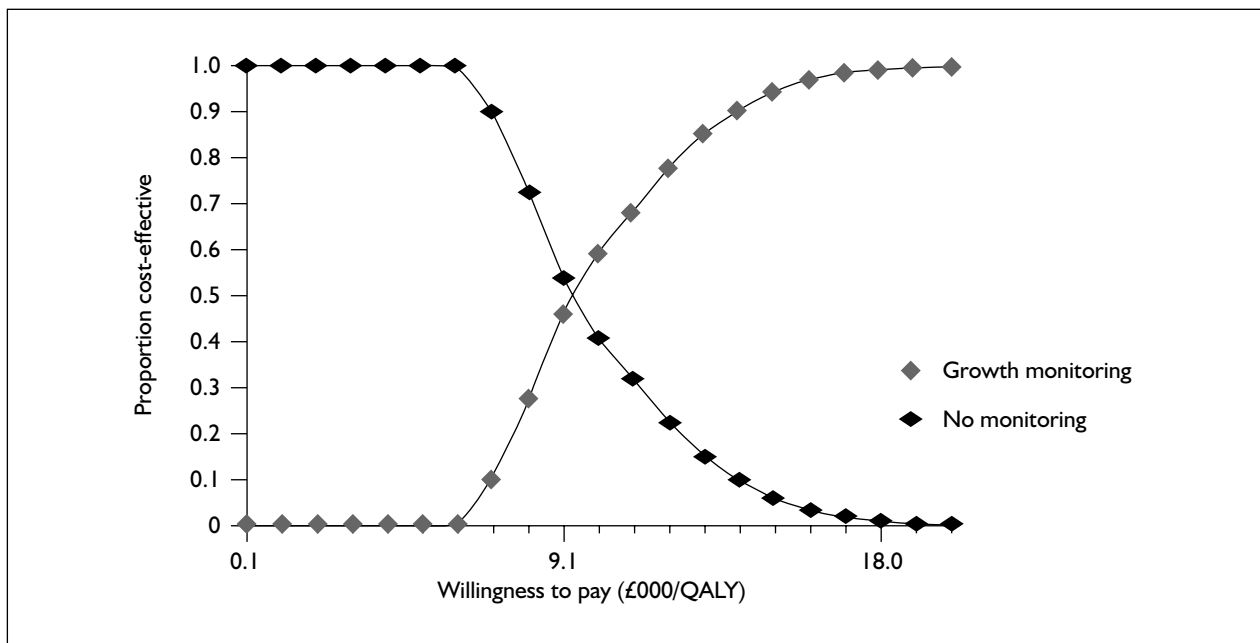**FIGURE 7** *Scatterplot of the ICER*

FIGURE 8 *CEAC for stature model*

TABLE 18 *Summary of intermediate outcomes (population of 5-year-olds in England and Wales: 594,000)*

Probability	Value	Number	Treated
Referral	0.008	4,752	
Non-attendance	0.171	813	
Retest	0.818	3,887	
Measurement error	0.182	707	
Initial investigation	0.829	3,222	
Find new GHD case among tested group	0.04	129	117
Find new TS case among tested group	0.006	19	16
Find new ISS case among tested group	0.543	1,750	26
Find new JH case among tested group	0.005	16	16
Find new psychosocial case among tested group	0.004	13	13
Find 'other' cases among tested group	0.403	1,299	Unknown
Total treated cases	216		
Incremental cost (£) per case detected	**18,222**		

The intermediate outcomes of the growth monitoring programme, based on a hypothetical cohort of all 5-year-old children in England and Wales, are shown in *Table 18*. The table shows the probability and number that can be expected to be referred, not attend, retested, measurement error, initial investigation, new cases of GHD, TS, ISS, JH, psychosocial, other cases and total cases treated. The incremental cost of the programme is £4 million and 216 cases are detected, making the incremental cost per case detected £18,222.

The results of the stature model can be summarised as follows:

- The solution indicates that monitoring results in the detection of new cases and is associated with health improvements at an additional cost.
- The baseline incremental cost per QALY of slightly under £10,000 is well within accepted WTP thresholds for the UK (usually £20,000–30,000 per additional QALY).
- The results of the probabilistic sensitivity analyses indicate that the monitoring programme is cost-effective 100% of the time over the given distributions for a WTP threshold of £30,000 per QALY.
- The main area of uncertainty is the QALY gains for early treatment (versus late detection and

treatment). These are based on estimates from a previous study that were broadly acceptable to NICE in its guidelines on GH treatment in children. Similar values were assigned to patients with JH and psychosocial causes of short stature, based on expert opinion. Uniform distributions were used, with late detection and treatment being assigned half of the early detection and treat estimates.

- Disutilities due to the monitoring programme (associated with being labelled as 'short' or the harms of treatment) have not been considered in the analysis.
- Cost estimates are based on UK studies and sources, reflated to 2004 prices. Apart from the cost of monitoring, which has a uniform distribution based on the minimum and maximum value in the literature, all cost estimates are regarded as point estimates as they are associated with much less uncertainty and obtaining credible distributions was not feasible.

Obesity model

The structure of the obesity model is shown in *Figure 9* and comprises four strategies, as follows:

1. *Primary prevention based on the intervention arm of the Planet Health model*. The programme would be given to all school children aged 11 years (according to the original study[74]), which would reduce the prevalence of obesity at age 14 years, based on the criterion of >85th percentile. This would reduce the incidence of overweight in adults (age groups 21–29 and 40–65 years) and therefore chronic conditions such as heart disease, diabetes and stroke. The medical costs associated with treating obesity-related morbidities would be reduced through avoided cases of obesity. The costs and outcomes (QALYs) are based on the original US study with costs being converted to UK pounds and reflated to 2004 prices. Because the original study demonstrated benefits only for girls, this strategy has two branches to reflect the differences in costs and benefits for both boys and girls.
2. *Primary prevention based on the intervention arm of the Planet Health model plus growth monitoring*. This strategy is identical with (1) but includes the cost of monitoring all school children at one point in time (ages 4–11 years or other). The benefit of monitoring would be to observe trends in the prevalence of obesity and could be justified if monitoring for stature is also cost-effective.

3. *Monitor for obesity and treat obese children*. This strategy is based on the economic evaluation by Goldfield and colleagues,[75] which demonstrated the cost-effectiveness of reducing weight in individuals by means of a group approach (see structured abstracts in Appendix 12 for full details). As this study involved only a short-term analysis, the intervention arms of the Wang model have been used to calculate long-term costs and QALYs. It is assumed that the intervention reduces the prevalence of obesity at age 14 years, consistent with the Wang study. The inclusion criterion for the original study was >91st percentile, which is higher than the Planet Health criterion, and Goldfield and colleagues used percentage reduction in overweight and standardised BMI as the clinical outcome. For this reason, the effectiveness of the Goldfield study was adjusted, as described in the next section.
4. *No monitoring, no primary prevention and no treatment*. This is the do-nothing comparator, which enables the relative costs and benefits to be calculated. The branch used is the no-intervention arm of the Wang model.

Input parameters for the obesity model
Table 19 lists the input parameters, sources and values (either point estimates or distributions) used for the obesity model.

The first variable (A) represents the excess medical costs of obesity. In order to calculate a value for the NHS, the original US value was converted from US dollars to UK pounds using Purchasing Power Parities (PPP) data[125] and reflated to 2004.[120] A similar approach was used to convert the cost of the Goldfield treatment programme (c_intervention)[75] to a UK value.

To adapt the cost of the Planet Health programme itself, the resources comprising the programme were taken from the original study and equivalent resources for the UK found. Appropriate costs for each resource were then derived (for example, the original teachers' salaries were replaced by UK salaries for similar grades of teacher), to produce a reasonable estimate for the UK. In some cases, conversion and reflation exercises were undertaken, as described above, for items such as the Planet Health book and fitness funds. Finally, a cost per child of the programme was calculated by dividing the cost of the Planet Health programme by the number of pupils who participated in the original study.
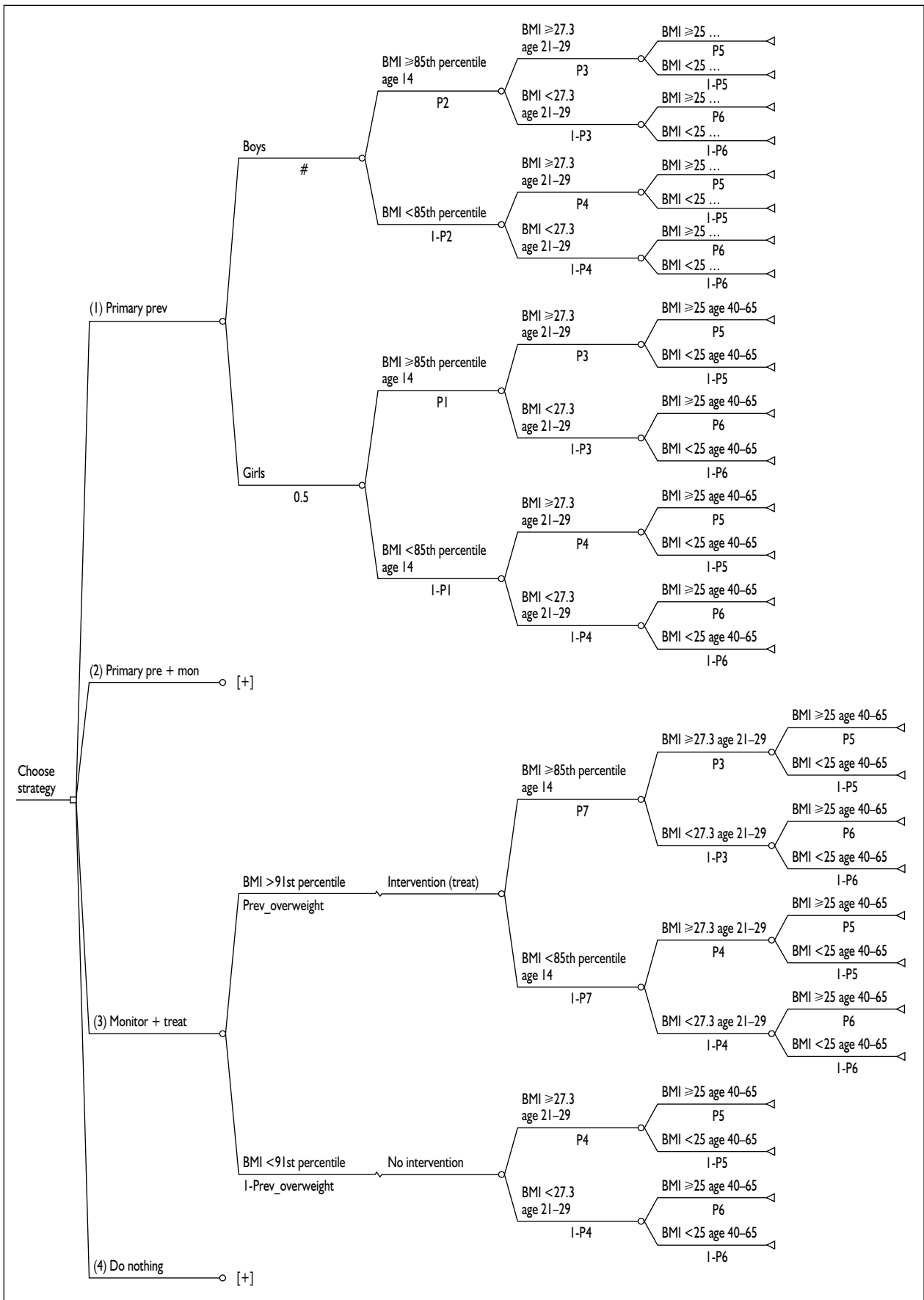
FIGURE 9 *Structure of obesity model*

TABLE 19 *Obesity model input parameters (cost data are in 2004 prices)*

Variable name	Description	Value (£)	Distribution	Source
A	Excess medical costs per overweight woman	3,360	Point	74
c_BMItest	Cost per child of monitoring programme	6.75	Uniform, 3.5–10	59, 121, 123
c_intervention	Cost of group intervention (Goldfield)	344.54	Point	75
c_planet_health	Total per child cost of Planet Health programme	28.99	Point	74
H	Intervention sample size UK cohort of 5-year-olds	594,000	Point	114
L_n	Expected number of years of life after age 40 years among non-overweight women who die during the 25-year period	16.5	Point	74
L_o	Expected number of years of life after age 40 years among overweight women who die during the 25-year period	16.05	Point	74
Mn	Probability of dying during the 25-year period for a non-overweight woman	0.117	Point	74
Mo	Probability of dying during the 25-year period for an overweight woman	0.152	Point	74
P1	Prevalence of students who are obese at age 14 years in the intervention scenario [27.7 × 0.5 × (20.3 + 25.8)]/25.8	0.249	Point	74
P2	A female student obese at age 14 in the no-intervention scenario (UK data)	0.277	Point	2
P3	An overweight female at age 14 years who becomes overweight by age 21–29 years	0.75	Triangular, 0.644, 0.754, 0.862	74
P4	A non-overweight female at age 14 years who becomes overweight by age 21–29 years	0.098	Triangular, 0.075, 0.098, 0.121	74
P5	An overweight woman at age 21–29 years who becomes overweight by age 40 years	0.912	Triangular, 0.849, 0.912, 0.975	74
P6	A non-overweight woman at age 21–29 years who becomes overweight by age 40 years	0.393	Triangular, 0.335, 0.393, 0.451	74
P7	Probability of remaining overweight (>85th percentile) after group intervention (treatment) for obesity	0.92	Triangular, 0.92, 0.94, 0.98	75
Prev_overweight	Prevalence of overweight in children in the UK, based on >85th percentile	0.277	Point	2
QALY	QALYs saved per case of adult overweight prevented	0.588	Point	74 (adjusted using UK discount rate)
r	Annual discount rate	0.035	Point	
Sn	Years of healthy life year per non-overweight woman age 40–65 years	0.83	Triangular, 0.827, 0.835, 0.842	74
So	Years of healthy life per overweight woman age 40–65 years	0.75	Triangular, 0.743, 0.753, 0.764	74

The other primary prevention variables were obtained directly from Wang and colleagues study.[74]

In order to harmonise the treatment strategy (3) with the primary prevention strategies (1 and 2), it was necessary to make some simplifying assumptions, as follows. The Goldfield study found that the group intervention reduced overweight and standardised BMI by approximately 8%. It was therefore assumed that the intervention would have a similar impact on the prevalence of the children treated, consistent with the Wang study. However, because of the difference in inclusion criteria (85th versus 91st percentiles), the figure of 8% was reduced to 6% and tested in the model over a range of 2–8% in the sensitivity analyses.

TABLE 20 *Summary of baseline cost-effectiveness results*

Strategy	Cost (£ million)	Incremental cost (£000)	QALY (000)	Incremental QALYs	ICER (£ per QALY)
(4) Do nothing	1,080 (1,001–1,155		162.6 149.4–176.5		
(1) Primary prevention	1,081 1,001–1,155	664.7	166.1 (153.1–180.1	3,515	189
(2) Primary prevention + monitor	1,084 (1,006–1,159)	398.6	166.1 (153.1–180.1	0	Dominated
(3) Monitor + treat	1,129	487.4	164.6 (151.3–178.8)	–1,539	Dominated
Strategy 3 versus 4		**49,399**	**1,977**		**24,991**
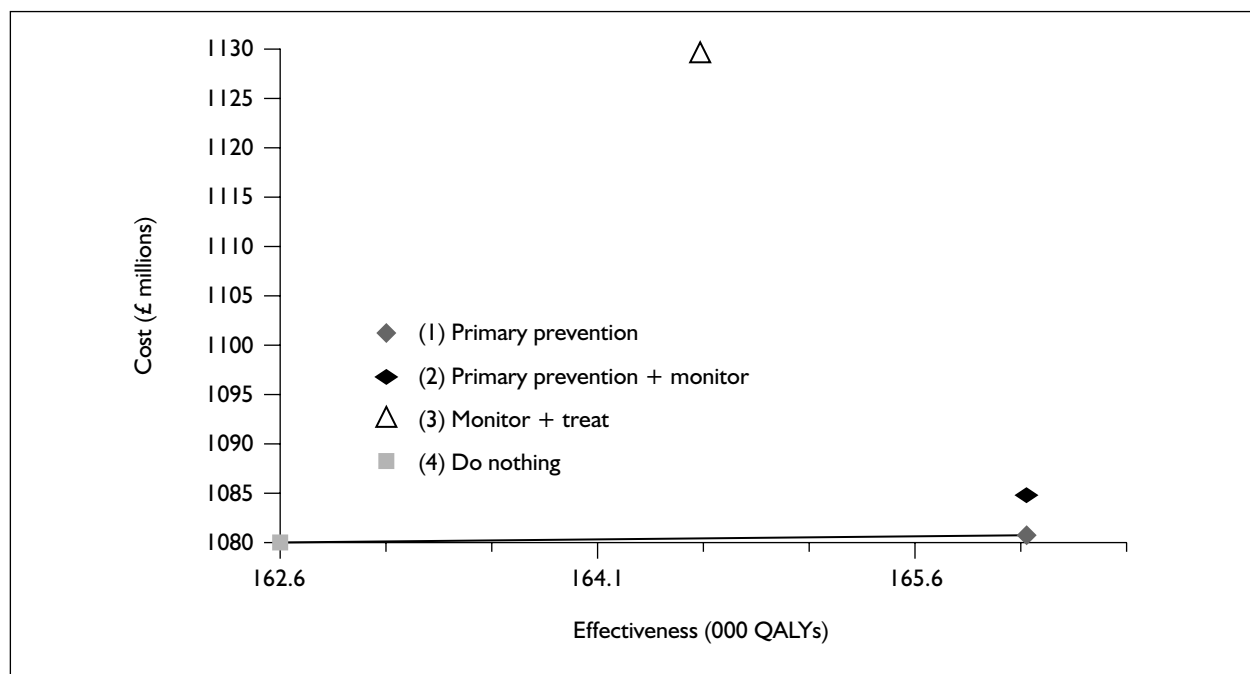
FIGURE 10 *Cost-effectiveness plane for obesity model, from Monte Carlo simulation*

The cost of the monitoring programme (c_BMItest) was calculated according to the method used in the stature model, and the same cohort of children, 594,000, was used to generate cost results relevant to the UK. These are described in the section 'Cost estimates' (p. 54).

Baseline results for obesity model

The baseline results for the obesity model are shown in *Table 20* and graphically in the cost-effectiveness plane of *Figure 10*. The findings are that, consistent with the original Wang study,[74] primary prevention without monitoring is a very cost-effective strategy compared with the do-nothing alternative. The baseline ICER is very low at £189 per additional QALY. Strategy 2 (primary

prevention with monitoring) produced the same number of QALYs as would be expected, at a small additional cost, and would produce a very similar ICER compared with the do-nothing option.

If one considers only the relative cost-effectiveness of the monitor and treat strategy, with respect to the do-nothing option, as might be the case if one only wishes to consider strategies that involve monitoring, the baseline ICER is £24,991 per QALY gained.

The impact on the results of uncertainty in the model can be assessed by examining the results of the probabilistic sensitivity analyses. First, if one examines the ranges associated with each strategy,
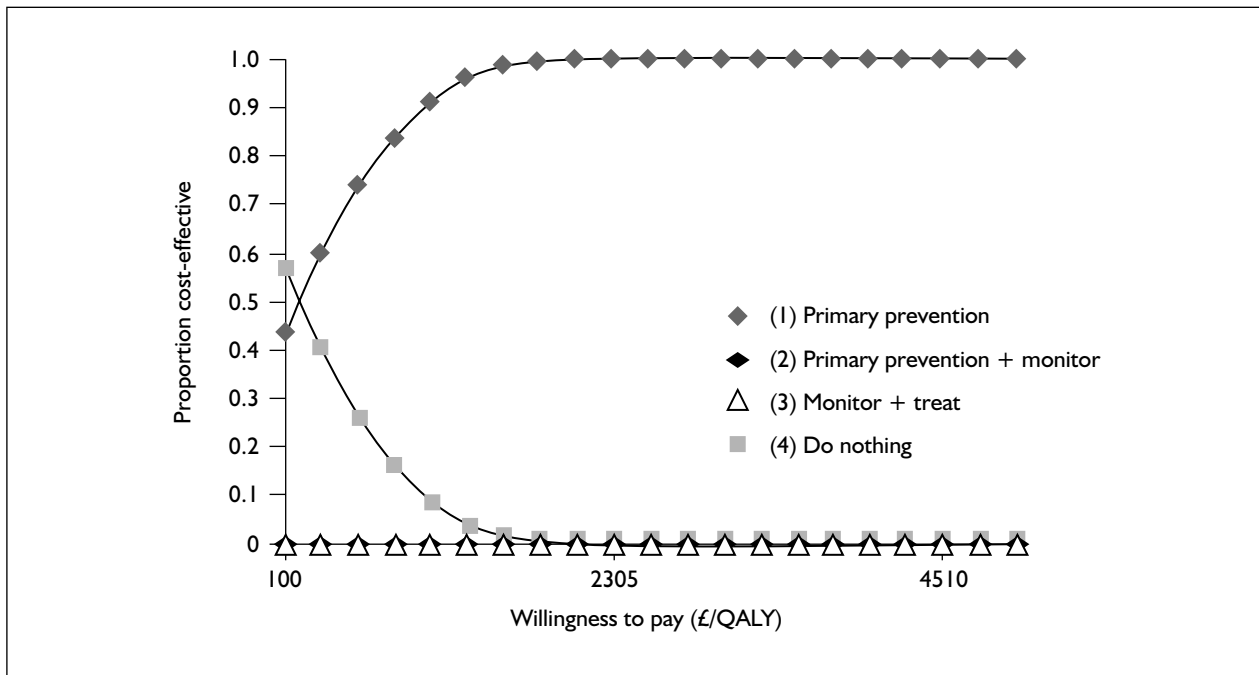
FIGURE 11 *CEAC for obesity model (all strategies)*
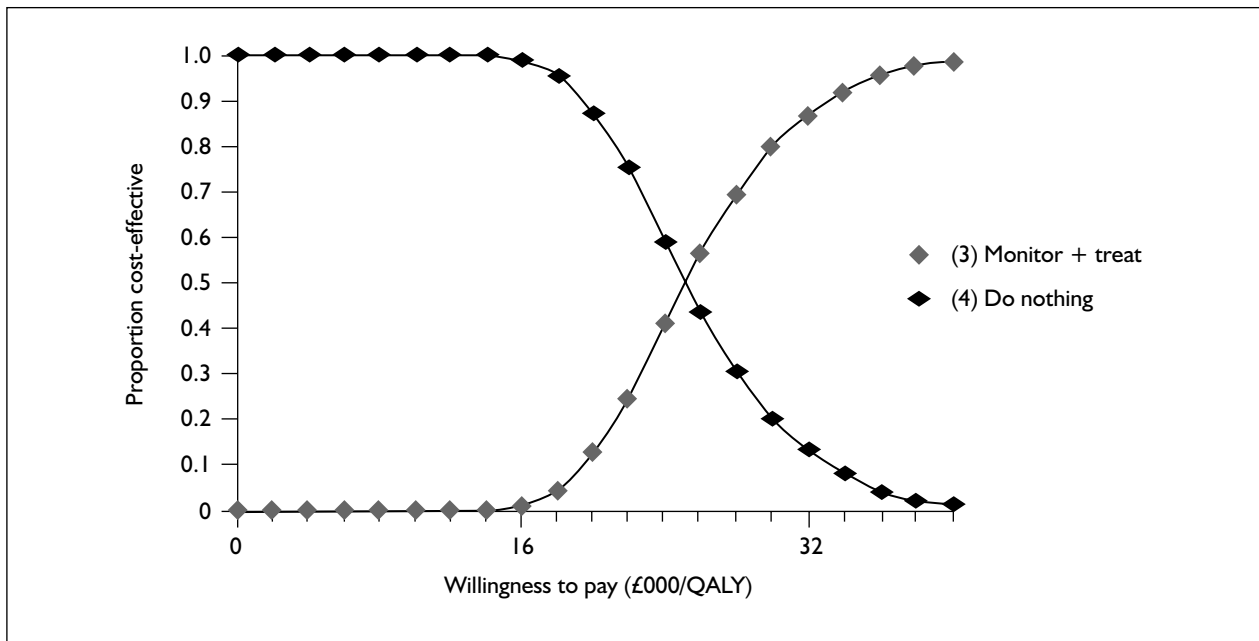
FIGURE 12 *CEAC for obesity model (strategies 3 and 4 only)*

it can be seen that there is considerable overlap in both costs and effects for all four strategies, which makes the results extremely unreliable.

The uncertainty in the model can also be assessed using CEACs. *Figure 11* shows the CEAC for the model when all strategies are included. It shows that primary prevention has a probability of one of being cost-effective at a WTP of approximately

£2000. This remains the case for all further WTP thresholds up to, and beyond, £30,000.

If the primary prevention strategies are ignored, the uncertainty in comparing strategy 3 with strategy 4 can be seen *Figure 12*. This shows that up to a WTP threshold of approximately £26,000 per QALY, the do-nothing option would be preferred. After this figure, the monitor and treat

strategy starts to become cost-effective, and reaches a probability of one of being cost-effective at a WTP of approximately £40,000 per QALY.

The sensitivity analysis also showed that the ICER between strategies 3 and 4 could be as high as £100,000 per QALY under the worst case scenario.

The present findings, therefore, suggest that monitor and treat, using the data available at present, would not be cost-effective.

In summary, the results of the obesity model suggest that:

- Primary prevention is the most cost-effective alternative over a WTP range of £500–30,000 per QALY.
- Monitoring and treating obese children, using the best available estimates, would not be a cost-effective option. The WTP threshold would need to reach £40,000 per QALY for this strategy to be cost-effective at a probability approaching one.
- The findings of the Wang study were that primary prevention is cost-effective only in girls. The results have been confirmed in the new model and adapted for the UK using prevalence and cost data appropriate for the UK. However, structural changes were made to the original Wang model to take account of boys in addition to girls.
- The large number of structural simplifications, assumptions regarding parameter estimates and distributions in the sensitivity analyses mean that the results of the obesity model need to be treated with considerable caution.

Chapter 7
Evidence addressing the National Screening Committee criteria

Does screening for growth-related conditions, including obesity, meet the NSC criteria?

This chapter applies the UK NSC criteria for evaluating a screening programme to screening for growth disorders including obesity. The authors' view on the extent to which each criterion is met concludes each section. It should be noted that not all of the criteria were addressed within the remit of our systematic review.

The condition

1. The condition should be an important health problem
Stature
Short and tall stature may not, in themselves, be regarded as health problems but may be indicators of underlying pathologies. The target conditions for stature monitoring are relatively common: GHD 1:4000, TS 1:2500 live female births and JH 1:1450. The prevalence of psychosocial short stature is unknown, but studies included in this review identified between 1:2256 and 1:20,338 children. If untreated, these conditions can lead to short stature in adulthood and other morbidities. Conditions of tall stature are relatively uncommon with the exception of Marfan syndrome (1:3000–5000) and Klinefelter syndrome (1:500–1000 live male births), although it is unclear how many might be identified through a screening programme.

- *Authors' summary opinion:* Fully satisfied.

Obesity
Obesity is an important health problem, with the UK prevalence of childhood overweight and obesity now standing at 27.7%.[2] Although obesity is not of itself a disease, co-morbidities in childhood may include type 2 diabetes, hypertension, dyslipidaemia, emotional and behavioural problems, asthma and sleep apnoea.[28–30] Obesity in childhood is highly likely to persist into adulthood and the effects of childhood obesity have been linked with morbidity and mortality in adulthood.[30]

- *Authors' summary opinion:* Fully satisfied.

2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor or disease marker and a latent period or early symptomatic stage
Stature
The epidemiology and natural history of GHD, TS, JH and conditions of tall stature are understood. Data on growth as a marker for disease are limited.

- *Authors' summary opinion:* Partially satisfied.

Obesity
The development of childhood nutritional obesity is considered to be linked to societal changes relating to nutrition and children's increased sedentary lifestyle.[28] As stated above, short-term morbidity associated with obesity in childhood is documented. Although there is evidence of persistence into adulthood, current knowledge of the long-term medical consequences of childhood obesity and appropriate indicator measures is limited. Evidence on thresholds of BMI that link with later morbidity is limited.

- *Authors' summary opinion:* Partially satisfied.

3. All cost-effective primary prevention interventions should have been implemented as far as practicably possible
Stature
This criterion does not apply to conditions of short and tall stature, with the possible exception of psychosocial short stature, about which little is known.

- *Authors' summary opinion:* Not applicable.

Obesity

The current review found that primary prevention for obesity has promising cost-effectiveness but is currently unproven, and larger trials are urgently needed. It is therefore not possible to determine whether all cost-effective primary prevention strategies have been implemented.

- *Authors' summary opinion:* Not satisfied – insufficient data.

The test

4. There should be a simple, safe, precise and validated screening test
Stature

Although measurement of height is a safe, non-invasive test, errors can be introduced as a result of inappropriate techniques, instruments being poorly installed and calibrated and variation within the child.[126] Training can be given to improve the precision of those measuring and recording height data.[110,111,126] Height as a screening test for disorders of stature has not been validated in studies of diagnostic accuracy.

- *Authors' summary opinion:* Partially satisfied.

Obesity

Ascertainment of weight and calculation of BMI is safe and simple. However, the review found that BMI is often a poor predictor of obesity when compared with densometrically defined measures of fat mass. A recent systematic review found that single BMI measures track reasonably well from childhood and adolescence into young adulthood.[30]

- *Authors' summary opinion:* Partially satisfied.

5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed
Stature

The UK 1990 charts are a cross-sectional reference for the UK population. There are no specific charts relating to ethnic minority groups. Research comparing the performance of different thresholds for the detection of growth disorders is lacking. Using the consensus cut-off of 0.4th centile at age 5 years, it has been estimated that 80% of those with undiagnosed GHD and 50% of those with undiagnosed TS might be identified.[15] The performance of this threshold for other conditions of short or tall stature is not clear.

- *Authors' summary opinion:* Partially satisfied.

Obesity

Thresholds for referral for obesity are based on population references and are arbitrarily defined. More research is needed to determine a threshold for referral based on morbidity.

- *Authors' summary opinion:* Not satisfied – insufficient data.

6. The test should be acceptable to the population
Stature

Although the review found some evidence that refusal rates for participation in height screening appear to be relatively low, consideration would need to be given to any potential variation in acceptability according to age group screened, ethnicity and gender.

- *Authors' summary opinion:* Not satisfied – insufficient data.

Obesity

Although there is some evidence that weight can be measured in a sensitive and acceptable way when combined with other activities,[72] careful consideration would need to be given to any screening of weight to identify obese children. Tests would need to be conducted in such a way as to avoid stigmatisation of children. As with stature there may be variation in refusal rates and acceptability according to age, ethnicity and gender.

- *Authors' summary opinion:* Not satisfied – insufficient data.

7. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals
Stature

Consideration needs to be given as to whether all children outside the given thresholds merit direct referral to a specialist or whether intermediate referral might prevent unnecessary specialist referrals and reduce parental and child anxiety.[56,82]

- *Authors' summary opinion:* Not satisfied – insufficient data.

Obesity

Any policy would need to take into account the fact that thresholds for referral for obesity are based on population references and are arbitrarily defined.

- *Authors' summary opinion:* Not satisfied.

The treatment

8. There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment
Stature
Identification of children with conditions of short stature such as TS and GHD leads to treatment to increase height and/or treatment and counselling for the underlying disorder. Early treatment is advantageous.[5–8] Treatments for the complications associated with tall stature conditions of Marfan's syndrome and Klinefelter's syndrome are available.

- *Authors' summary opinion:* Fully satisfied.

Obesity
There is currently a lack of proven long-term effective treatment for obesity, although some programmes have demonstrated success in the short term.[31]

- *Authors' summary opinion:* Not satisfied.

9. There should be agreed evidence-based policies covering which individuals should be offered treatment and the appropriate treatment to be offered
Stature
There is an evidence base for growth-related conditions such as GHD and TS.[108]

- *Authors' summary opinion:* Fully satisfied.

Obesity
As there is no consensus on when to refer or which treatments might be most beneficial, the appropriate treatment is unclear.

- *Authors' summary opinion:* Not satisfied.

10. Clinical management of the condition and patient outcomes should be optimised by all healthcare providers prior to participation in a screening programme
Stature
The primary–secondary–tertiary care management pathway for general and specialist health problems in children, including those presenting with abnormal stature/growth, is well established in the UK, as evidenced by the fact that growth monitoring identified only 25–36% of growth disorders, the remainder of children with these conditions being already known to secondary and tertiary care paediatricians (Shah SA, University of Manchester: personal communication, 2000).[9] Nutritional advice and surveillance of severe psychosocial deprivation is carried out in the preschool years. It is unclear if this is optimal.

- *Authors' summary opinion:* Partially satisfied.

Obesity
No clear clinical management pathways exist for children identified as being overweight or obese.

- *Authors' summary opinion:* Not satisfied.

The screening programme

11. There must be evidence from high-quality RCTs that the screening programme is effective in reducing mortality and morbidity
Stature
The review did not locate any RCTs comparing screening with no screening for growth disorders.

- *Authors' summary opinion:* Not satisfied – insufficient data.

Obesity
The review did not locate any RCTs comparing screening with not screening for obesity.

- *Authors' summary opinion:* Not satisfied – insufficient data.

12. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/intervention) is clinically, socially and ethically acceptable to health professionals and the public
Stature
There is no strong evidence from the studies reviewed on the attitudes of health professionals, children and parents to screening for growth disorders. Although there is some evidence of enthusiasm for growth screening amongst community paediatricians and paediatric endocrinologists,[35] the views of school nurses, health visitors, children, parents and others involved in measurement of growth need to be determined.

- *Authors' summary opinion:* Not satisfied – insufficient data.

Obesity

There is no strong evidence from the studies reviewed on the attitudes of health professionals, children and parents to screening to detect obesity.

- *Authors' summary opinion:* Not satisfied – insufficient data.

13. The benefit from the screening programme should outweigh the physical and psychological harm

Stature

Studies included in the review indicate that growth screening detects a number of children with growth disorders who have not been identified through other routes. The benefits of earlier identification of growth disorders have already been discussed. The included studies provided no evidence on the harms of screening, for example being labelled short or the anxiety associated with attending for further investigation, or on the impact of unnecessary tests in false positives. In determining any harms of the screening programme, further research might clarify suitable referral procedures given potentially high rates of non-attendance at follow-up appointments.

- *Authors' summary opinion:* Not satisfied – insufficient data.

Obesity

The review did not locate any studies describing the relative benefits/harms of screening for obesity. Taken in conjunction with a measure (BMI) that is likely to generate false positives, the balance of benefits and harms is unknown. The potential harms of labelling children as obese should be set against the current lack of effective long-term treatments.

- *Authors' summary opinion:* Not satisfied – insufficient data.

14. The opportunity cost of the screening programme should be economically balanced in relation to expenditure on medical care as a whole

Stature

Within the caveats and limitations regarding data quality and reliability, the results of the stature model suggest that monitoring and treating stature-related conditions are likely to be cost-effective within commonly accepted WTP criteria for the NHS.

- *Authors' summary opinion:* Partially satisfied.

Obesity

The results of the obesity model indicate that primary prevention of obesity would be a cost-effective approach and well within WTP criteria for the NHS. However, monitoring and treating obesity are not cost-effective. The model is subject to a great deal of uncertainty regarding monitoring and treating obesity, and the results can only be regarded as tentative.

- *Authors' summary opinion:* Not satisfied – insufficient data.

15. There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards

Stature

A plan and associated standards would need to be developed.

- *Authors' summary opinion:* Not satisfied.

Obesity

A plan and associated standards would need to be developed.

- *Authors' summary opinion:* Not satisfied.

16. Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be made available prior to the commencement of the screening programme

Stature

There is a need to ensure the competence of those measuring. The reviewed studies support previous work on the importance of training to ensure that children are measured accurately and that recording of measurements and subsequent referral are properly conducted. Facilities for measurement may not be routinely available.[35] Staffing and facilities required would depend on the frequency of measurements and thresholds selected.

- *Authors' summary opinion:* Not satisfied.

Obesity

Given uncertainties on the definition of obesity and appropriate treatments, the requirements in terms of staff and facilities are unclear.

- *Authors' summary opinion:* Not satisfied.

17. All other options for managing the condition should have been considered (e.g. improving treatment, providing other services) to ensure that no more cost-effective intervention could be introduced or current interventions increased within the resources available

Stature

- *Authors' summary opinion:* Not applicable.

Obesity

Further research on methods of preventing obesity and cost-effectiveness of various methods remains to be investigated.

- *Authors' summary opinion:* Not satisfied.

18. Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice

Stature

The effect of providing information about monitoring for short or tall stature has not been investigated in the reviewed studies. The impact on refusal rates is unclear.

- *Authors' summary opinion:* Not satisfied.

Obesity

The effect of providing information about monitoring for obesity has not been investigated in the reviewed studies. The advantages of an 'opt in' versus 'opt out' approach to monitoring could be further investigated.

- *Authors' summary opinion:* Not satisfied.

19. Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public

Stature

As there is no strong evidence comparing different screening intervals and thresholds, it may be difficult to justify particular parameters.

- *Authors' summary opinion:* Not satisfied.

Obesity

As there is no strong evidence comparing different screening intervals and there is uncertainty concerning thresholds, it may be difficult to justify particular parameters.

- *Authors' summary opinion:* Not satisfied.

TABLE 21 *Summary of NSC criteria*

Criterion	Authors' opinion	
	Growth disorders	**Obesity**
Disease		
1	Fully satisfied	Fully satisfied
2	Partially satisfied	Partially satisfied
3	Not applicable	Not satisfied – insufficient data
Test		
4	Partially satisfied	Partially satisfied
5	Partially satisfied	Not satisfied – insufficient data
6	Not satisfied – insufficient data	Not satisfied – insufficient data
7	Not satisfied – insufficient data	Not satisfied
Treatment		
8	Fully satisfied	Not satisfied
9	Fully satisfied	Not satisfied
10	Partially satisfied	Not satisfied
Screening programme		
11	Not satisfied – insufficient data	Not satisfied – insufficient data
12	Not satisfied – insufficient data	Not satisfied – insufficient data
13	Not satisfied – insufficient data	Not satisfied – insufficient data
14	Partially satisfied	Not satisfied – insufficient data
15	Not satisfied	Not satisfied
16	Not satisfied	Not satisfied
17	Not applicable	Not satisfied
18	Not satisfied	Not satisfied
19	Not satisfied	Not satisfied

Table 21 summarises the performance of growth monitoring against the NSC criteria. It can be seen that, based on current evidence, screening for growth disorders including obesity does not meet all of the NSC criteria. Although growth-related disorders are important conditions and effective treatments exist for at least some of them, certain criteria regarding the screening programme have not been met.

The performance of height measurement as a screening test has not been evaluated in studies of diagnostic accuracy. There are uncertainties around the appropriate ages to screen and thresholds to adopt. There is a lack of data on attitudes to growth monitoring, so the acceptability of the programme is unclear. Although benefits have been observed in terms of identifying children with growth disorders, potential harms have not been fully investigated. Although several human resource issues have been identified and modelling included in this report indicated that growth monitoring may be cost-effective, the introduction of national growth screening would require plans to be in place for managing and monitoring quality standards and ensuring adequate staffing and facilities. Turning to obesity, although an important condition with a high prevalence, uncertainties concerning the definition and appropriate treatment of obesity mean that most of the screening criteria have not been met. Data on the relative benefits and harms are lacking. The relative merits of screening and treating for obesity compared with alternative prevention methods are unclear.

Chapter 8
Discussion

Weaknesses of the evidence base

A systematic review was conducted to determine the clinical impact and cost-effectiveness of routinely monitoring growth in children between the ages of 4 and 11 years in order to identify growth-related conditions, including obesity. Following extensive searching, 31 studies were included.[56,59,71,72,76,78–83,86–91,93,95–107] Despite a relatively large and international literature, definitive answers to the review questions detailed in the methods section could not be given.

Although all included studies provided data to address one or more review questions, the aim of the studies and hence their design were not always the most appropriate to answer the questions directly. First, and most significantly, no controlled studies were found evaluating growth monitoring versus no growth monitoring for the detection of growth-related conditions including obesity. Hence the clinical effectiveness of growth monitoring programmes cannot currently be established. Second, no studies were found reporting on the diagnostic performance of growth monitoring programmes for the identification of growth-related conditions. The evidence reported in this review was based on studies of growth monitoring programmes already in place reporting a diagnostic yield of new growth-related conditions detected or all growth-related conditions detected (including existing cases). In this type of study, only children found to be too short or too tall according to the designated threshold are followed up. Hence only information about true positive and false positive rates can be obtained. Children found to be of normal height are not followed up, so false negative and true negative rates cannot be determined. These studies are, therefore, inherently flawed in their assessment of the diagnostic performance of growth monitoring. However, they do provide an estimate of the additional benefit of growth monitoring (over no intervention) in terms of number of additional cases detected, and also an estimate of the burden of unnecessary referral and further investigation. Taking a pragmatic approach, in the absence of comparative studies, we developed a tool to assess the methodological quality of these studies to determine the reliability of their estimates of

detection rate of growth disorders. Although a number of studies met most of the methodological criteria assessed, the majority of the studies had some methodological limitations which might impact on the reliability of the diagnostic yield estimates. Almost half of the studies failed to measure >80% of the study sample. A further significant limitation was that several studies did not have a sufficiently large sample size, given the estimated prevalence of the growth-related conditions under consideration, to detect one case of a target condition.[59,76,81,86]

In terms of obesity, 11 studies provided data on the diagnostic performance of measures to identify obesity used in growth monitoring programmes, most commonly BMI using a variety of thresholds.[97–107] These studies had a number of methodological limitations, in particular derivation of the diagnostic threshold internally from the study population was common and limited information was provided on selection procedures and any withdrawals from the study. Perhaps more importantly, the included studies are more informative on the limitations of BMI as a measure of obesity in comparison with direct measures of body fat than on whether screening for obesity using BMI (or any alternative surrogate measure of body fat) is a useful exercise when compared with no screening.

Limited data on the human resource implications of growth monitoring were found both in studies of the detection rate of growth disorders and in studies specifically designed to assess resourcing issues. Of those studies focusing on human resource issues, just one was an RCT.[88] Although a number of important issues concerning training, costs and referral were identified, where data are derived from uncontrolled trials the influence of extrinsic factors cannot be discounted. No studies identified human resource issues specific to obesity screening.

Data on attitudes to growth monitoring were extremely sparse. Just one study was specifically designed to assess attitudes to growth monitoring and this was solely in relation to obesity.[72] This study was expressly designed to minimise negative impact. The remainder of the data were gleaned

from studies evaluating growth monitoring programmes or assessing their human resource requirements. These included the rates of refusal to participate in growth monitoring and reasons for non-attendance at follow-up. As none of the data were derived from controlled trials, the influence of extrinsic factors cannot be determined. For example, it is unclear whether refusal to attend a follow-up appointment varies for those identified through a growth monitoring programme as opposed to those referred through other routes. The data are inadequate to assess any potential harms of growth monitoring in terms of being labelled short, tall or obese, raising anxiety concerning a potential growth disorder or any potential stigmatisation.

The review questions specified that all studies had to be broadly representative of the UK population and target countries were identified as part of the inclusion criteria. However, it was not possible to assess how similar many of the study populations were to the UK as limited demographic details were provided. Although studies generally reported details of the selection procedure and described attempts to contact all those eligible for measurement, the number approached and the percentage of those who were measured were not always reported in the growth monitoring programme studies, which may impact on the yield. Similarly, few of the obesity studies reported on the number approached. This implies that there may be the potential for selection bias, that is, differences between those approached who did not take part and those who did participate.

The majority of studies were based on one-off screening for growth disorders. Overall the data were inadequate to answer policy-level questions on the relative benefits and harms of growth monitoring to detect growth disorders including obesity. Assuming that a case can be made for growth monitoring in general, no studies were found comparing different growth monitoring strategies. Hence the age at which to screen, whether to use single or serial measurements and which thresholds for referral are most appropriate could not be determined. Without comparative studies, it is impossible to isolate the performance of a growth monitoring programme from other factors such as the success in identification of growth disorders through existing routes. Finally, due to study heterogeneity it was not possible to compare the diagnostic yields derived from different criteria for the detection of growth disorders.

Summary of the findings

From the previous chapter, it can be clearly seen that monitoring for growth-related conditions including obesity does not currently meet the NSC criteria. For stature-related disorders this was most apparent in terms of a lack of evidence on test performance and screening programme effectiveness. Although obesity is an important condition with a high prevalence, uncertainties concerning its definition and appropriate treatment mean that most of the screening criteria have not been met. Data on the relative benefits and harms of screening for obesity compared with alternative prevention methods are lacking.

There remains a fundamental question as to whether or not growth monitoring should really be considered a screening exercise and evaluated in such terms. Despite failing to meet some of the NSC criteria, the evidence from this review does indicate that monitoring growth can be a worthwhile exercise in terms of identifying significant pathologies in children that have been missed by other routes. Whether, if improvements were made in terms of more frequent measurement of children at routine healthcare encounters, a specifically targeted screening programme might be less useful cannot be determined. What the results of the review do indicate is that one-off height measurement picks up a small number of children with important growth-related problems including GHD and TS. Where detected, the yield for GHD ranged from 1:1605[83] to 1:20,338.[56] Where reported, for TS the yield was between 1:15,246 and 1:48,221. As a secondary gain, growth monitoring appears to identify children with other potentially treatable conditions. The range of detection rates in the included studies was from 1:545[83] to 1:4595[87] for any condition. Three studies did not detect any new cases of any conditions,[59,71,79] but all of these had important methodological limitations such as small sample size or high attrition rates at follow-up. Due to limitations in the design of the included studies, the age or ages at which monitoring might be most beneficial are unclear. However, the larger, most robust studies of school entry-aged children indicate that a single measure at this age might identify between 1:1849[56] and 1:1793[93] children with any new condition. It is unclear from the studies that include older age groups what might be the incremental gain from further monitoring. Two of the larger studies of primary school age children had an uncertain diagnostic yield as it was not completely clear which cases of growth-related disorder were new and which were pre-existing.[87,91]

Cost-effectiveness modelling conducted in addition to the review indicated that growth monitoring was cost-effective according to accepted WTP thresholds in the UK of £20,000–30,000 per QALY. The mean cost per additional QALY was estimated at £9500. The probabilistic sensitivity analyses showed that monitoring is always cost-effective at UK values of WTP. Furthermore, most of the costs are follow-on costs from referral and treatment. The findings of the economic modelling suggest that monitoring for stature is justifiable although the analysis was not able to capture the harmful effects of monitoring associated with side-effects of further referral, treatment and labelling. There is also considerable uncertainty associated with the estimates used to derive QALYs, as these were based on previous estimates and not empirical evidence, although the GH-related values were given cautious acceptance by NICE in the compilation of guidelines for GH treatment in children. The included studies also highlighted that intermediate stage referral to confirm short stature may increase the cost-effectiveness of growth monitoring programmes.

The validity of the stature model could be improved by the provision of accuracy data from long-term studies that follow up children who are not considered short at monitoring. The present model is based on probabilities derived from *post hoc* diagnostic yield, which meant that the model could not be built using conventional Bayesian analyses in diagnostic applications. More rigorous modelling would also employ diagnostic accuracy data for all subsequent tests once a child is found to meet the criteria for further investigations. In the present model, typical tests for each diagnostic outcome, derived from the systematic review findings based on diagnostic yield, were assigned. Again, this is due to the *post hoc* methods employed.

Within the above caveats for the stature modelling, the findings appear to be credible and reflect the experiences and considerations of clinical experts working in this field.

Effectiveness and cost-effectiveness of growth monitoring may depend on appropriate training, measurement precision and referral systems. The potential benefits and cost-effectiveness of screening for growth disorders should be considered alongside the lack of evidence on attitudes of children, parents and healthcare providers to growth monitoring programmes. The limited data available suggest that, based on low refusal rates, children (and their parents) are willing to participate in a growth monitoring programme. Potential problems with attendance at referral need to be further investigated. The harms of screening for growth disorders are unknown. However, there is evidence from outside this review that the harms of being labelled short do not appear to be significant.[41] There are no data, however, on the impact of screening on those subsequently found to be normal. One included study from Sweden detailed the benefits of screening as perceived by an expert panel of experienced school nurses and senior paediatricians also working as school doctors.[79] Evidence from outside the review indicates that healthcare professionals are supportive of growth monitoring.[35]

The cost-effectiveness model for obesity is based on two published models that addressed primary prevention and treatment. These were adapted to include monitoring as part of primary prevention or monitoring to locate obese children before referring them for treatment. The findings suggest primary prevention is a very cost-effective approach, but there is a great deal of uncertainty in the results.

In order to harmonise the find and treat approach with the primary prevention approach within one model, costs and QALYs were based on both girls and boys. The original Wang model[74] only assessed effectiveness in girls, but in order to harmonise this with the treatment alternative a number of assumptions were made. Assumptions were also made in order to derive an estimate regarding the effectiveness of treatment for obesity. These and other simplifying assumptions used in the obesity modelling mean that the results need to be treated with a great deal of caution. They are speculative and have weak decision-making value. However, the structure of the model lays the foundation for future research using more reliable estimates. The validity of the model could be improved by the use of data from RCTs with long-term follow-up and UK-specific predictions of excess medical costs and QALY losses due to obesity

If primary prevention strategies are not taken into consideration, the ICER of the monitor, find and treat strategy is approximately £25,000 per QALY (but could be as high as £100,000 per QALY), suggesting that this approach would not be cost-effective according to accepted WTP thresholds. However, there is too much uncertainty around this model to make strong recommendations, and

treatment programmes for the more severely obese should clearly be complementary to other initiatives such as primary prevention.

Even if the incremental cost of monitoring obesity is considered to be close to zero if it is combined with monitoring stature, the result hardly changes. This is because the cost of treating those found to be obese is the major cost in the monitor and treat strategy.

Furthermore, additional screening and treating for obesity are dependent on benefits outweighing harms. The effectiveness of treatment is currently doubtful. The harms of screening are currently poorly researched and, taken in conjunction with a screening measure that is likely to generate false positives, the balance of benefits and harms is unknown. In the light of this, current models of self-referral and attempted treatment for obesity should continue.

Implications for policy and practice

Monitoring for stature-related disorders

This review has indicated the potential utility and cost-effectiveness of growth monitoring in terms of increased detection of stature-related disorders. However, it has also pointed strongly to the need for further research. There is a lack of high-quality evidence on the potential impact of a growth monitoring programme, its acceptability and any potential harms. Gaps and uncertainties in the evidence base mean that growth monitoring does not currently meet all of the NSC criteria. However, it is questionable whether some of the NSC criteria can ever be meaningfully applied to growth monitoring, given that short stature is not a disease in itself, but rather is used as a marker for a range of pathologies and as an indicator of a child's general health status. There is a need to consider the extent to which it is appropriate to evaluate growth monitoring against the NSC criteria. Those considering implementing growth monitoring programmes, whether at a local or national level, need to consider whether the benefits of earlier detection of disorders of stature, as evidenced by the incremental yield observed in the studies included in this review, outweigh the lack of information on other relevant NSC criteria. It may be useful to consider the potential benefits of growth monitoring in the context of overall child health monitoring, (e.g. this review indicates that growth monitoring may detect important, treatable disorders where short stature is a secondary presenting feature in addition to the

primary target conditions). There is considerable potential for the conduct of research, as part of growth monitoring programmes, which would contribute to better understanding of the role of growth monitoring in children's health (see the section below, 'Implications for research'). In the light of the weakness of the current evidence base, those considering implementation of growth monitoring should also consider the consequent opportunities for research.

Monitoring for obesity

Our review found a lack of data on the potential impact of monitoring for obesity and more research is indicated. The cost-effectiveness model incorporated a great deal of uncertainty. Identification of effective interventions for the treatment of obesity is likely to be considered a prerequisite to any monitoring programme that is designed to identify individual overweight and obese children. Similarly, further long-term studies of the predictors of obesity-related co-morbidities in adulthood are warranted. These would clarify the role of screenable parameters such as BMI in determining those children most at risk; at present it is unclear how the target population of any monitoring programme should be defined. There is a need to consider the above points, and also the lack of data on the benefits and harms of the monitoring process, before moving away from the current population-based approach to obesity monitoring, which does not seek to identify and treat individual children.

Implications for research

Monitoring for stature-related disorders

The primary consideration for future research on growth monitoring is the establishment of its clinical and cost-effectiveness as part of the process of detecting and treating stature-related conditions.

The effectiveness of growth monitoring would be most reliably determined by a cluster randomised trial comparing growth monitoring strategies with no growth monitoring in the general population. Such a controlled trial would allow for the evaluation of growth monitoring taking account of confounding factors. The trial would need to be large enough to reflect the prevalence of growth-related conditions such as GHD and TS. It should ideally consider disorders of tall stature in addition to short stature, as there is evidence that referrals for tall stature are becoming increasingly unusual.[23] It might also include multiple arms

comparing different growth monitoring strategies to incorporate single screening and serial measurements at different ages. Different referral strategies such as intermediate referral might usefully be examined. The trial might also examine the effect of different staff training and delivery strategies. Information on resourcing issues and attitudes to the programme and subsequent referral could also be collected. The trial would need to be of a sufficient duration to determine relevant outcomes such as age of referral and treatment outcomes. Overall benefits and harms should be evaluated alongside costs.

Such a trial has large-scale practical implications and costs such that results would not be available for several years. Studies of diagnostic accuracy to determine the diagnostic performance of growth monitoring for the identification of disorders of stature, alongside evidence of effective treatment strategies, could provide an alternative approach to assessing clinical effectiveness. In this context, careful consideration should be given to target conditions and intervention thresholds, e.g. should GH be used in 'short normal' children? Further work is needed to determine the prevalence of psychosocial problems in school-aged children and short stature as a marker of such problems. There is some evidence from outside the review that a considerable proportion of 'short normal' children have significant psychosocial problems.[127] Diagnostic accuracy studies would require long-term follow up of both short and normal children to determine sensitivity and specificity of growth monitoring; again, a variety of ages and thresholds would need to be examined. However, this strategy could be applied either as a pilot study or as a method of post-implementation evaluation.

A third potential approach would be to assess the performance of screening rules and referral criteria using case control studies of existing data sets.[36] These studies have a degree of bias as they only use known cases, but could be used to inform practice whilst awaiting the results of large-scale prospective diagnostic accuracy studies.

Concerning areas of research which would provide additional information pertinent to the NSC criteria, the following issues might best be addressed through qualitative research and surveys:

- What system-related barriers might arise that impact on the implementation of an effective growth monitoring programme? These might include availability of staff and equipment, need for training and need for policies on appropriate referral to specialists.
- What is the most appropriate way to manage and monitor the programme? Any research should build on the lessons learned through existing programmes[128] to ensure best practice.
- What are the most appropriate methods of ensuring the quality of the programme?
- What are the attitudes of healthcare professionals, parents and children to the whole screening programme, including referral and attendance at follow-up appointments?
- Do attitudes vary with the age, gender or ethnicity of children measured? Do attitudes vary with an opt-in/opt-out approach to growth monitoring?

The following issues might best be addressed using interventional studies such as randomised controlled trials:

- How can coverage be optimised?
- Who might best measure children?
- What effective training is needed to ensure the competence of those measuring and to ensure appropriate referrals?
- What is the effect on uptake of providing different levels of information to children and their parents about monitoring for growth disorders?

If evidence from the research described above supported the introduction of growth monitoring, an ongoing programme of evaluation would be required. An ongoing audit of the programme should address the following issues:

- What coverage can be achieved?
- What is the yield of new cases identified?
- What is the workload involved in measuring groups of children?
- What are the costs involved in running the programme?

Concerning more general points, all further research on growth monitoring should ensure relevant reporting of all factors that might impact on detection rates of growth disorders. Studies should clearly report the following details:

- selection criteria for participants
- attempts to contact all those eligible for measurement
- methods to ensure and check the competence of those measuring
- a reproducible protocol to ensure consistency of measurements

- coverage
- measurement error
- all children's measurement results
- follow-up results for all children
- diagnoses of any new or existing growth disorders (organic or non-organic).

Monitoring for obesity

In the absence of evidence of effective interventions, the value of monitoring children in order to identify overweight and obesity will remain questionable. The effectiveness of treatment is currently doubtful and further research in this area, ideally in the form of RCTs, is therefore a priority.

Of equal priority is research on the effectiveness of primary prevention as an alternative or complementary strategy to detection and treatment. Primary prevention for obesity has promising cost-effectiveness but is currently unproven, and larger and longer term trials are urgently needed. The impact of effective prevention strategies on the need to screen and treat should be investigated.

Current knowledge of the long-term medical consequences of childhood obesity is limited. Establishing long-term epidemiological studies is now a high priority. These should help elucidate which children are at most risk of obesity persisting into adulthood and which children are at most risk of adverse outcomes of obesity in adulthood. They are a necessary precursor to research on the effectiveness of monitoring to identify and treat, as they define the target population.

Before implementation of any monitoring programme, with the aim of identifying and treating overweight and obese children, funding for UK research into the benefits and harms of screening and treating for obesity, including long-term outcomes, is a high priority. Should effective treatments for obesity be identified, the effectiveness of monitoring for obesity would be most reliably determined by a cluster RCT comparing monitoring strategies with no monitoring and with alternative preventative strategies. Such a controlled trial would allow for the evaluation of monitoring taking account of confounding factors. As with screening for growth-related conditions, the trial might include multiple arms comparing different monitoring strategies to incorporate single screening and serial measurements at different ages. The impact of different thresholds of obesity and different referral strategies might be examined. The trial might also examine the effect of different staff training and delivery strategies. Information on resourcing issues and attitudes to the programme and subsequent referral could also be collected. Methods of identifying obese children without stigmatisation could be examined. Any trial would need to be of a sufficient duration to determine relevant outcomes such as age of referral and treatment outcomes (sustained weight loss and long-term morbidities and mortality).

Chapter 9
Conclusions

The research objectives of this review were to determine the clinical impact and cost-effectiveness of routinely monitoring growth in children between the ages of 4 and 11 years in order to identify growth-related conditions. This review has indicated the potential utility and cost-effectiveness of growth monitoring in terms of increased detection of stature-related disorders. It has also pointed strongly to the need for further research. There is a lack of high-quality evidence on the potential impact of such a monitoring programme, its acceptability and any potential harms. Plans for managing any proposed growth monitoring programme, making available adequate staffing, facilities and programme information, would need to be developed. Gaps and uncertainties in the evidence base mean that growth monitoring does not currently meet all of the NSC criteria. However, it is questionable whether some of the NSC criteria can ever be meaningfully applied to growth monitoring, given that short stature is not a disease in itself, but rather is used as a marker for a range of pathologies and as an indicator of a child's general health status. There is a need to consider the extent to which it is appropriate to evaluate growth monitoring against the NSC criteria, and also whether the benefits of earlier detection of disorders of stature as evinced by the incremental yield observed in the studies included in this review outweigh the lack of information on other relevant screening programme criteria.

Our review found a lack of data on the potential impact of monitoring for obesity and more research is indicated. The cost-effectiveness model incorporated a great deal of uncertainty. The relative benefits and harms of monitoring have not been determined and the effectiveness of current treatments is doubtful. Given these and other uncertainties, monitoring for obesity does not currently meet the majority of NSC criteria. Further long-term studies of the predictors of obesity-related co-morbidities in adulthood are warranted. This would clarify the role of screenable parameters such as BMI in determining those children most at risk. However, until effective interventions for obesity are identified, there is little to be gained by determining new methods of identifying obese children.

Acknowledgements

References

1. Hall DM, Voss LD. Growth monitoring. *Arch Dis Child* 2000;**82**:10–15.

2. Jotangia D, Moody A, Stamatakis E, Wardle H. *Obesity among children under 11*. London: Joint Health Surveys Unit, National Centre for Social Research; 2005. URL: http://www.dh.gov.uk/ PublicationsAndStatistics/Publications/Publications Statistics/PublicationsStatisticsArticle/fs/en? CONTENT_ID=4109245&chk=WB/AR1

3. House of Commons Health Committee. *Obesity: third report of session 2003–04*. London: Stationery Office; 2004.

4. Department of Health. Children and young people: starting on the right path. In *Choosing health: making healthy choices easier*. London: Stationery Office; 2004. pp. 40–75.

5. Aynsley-Green A, Macfarlane JA. Method for the earlier recognition of abnormal stature. *Arch Dis Child* 1983;**58**:535–7.

6. Job JC. Early diagnosis and early treatment of growth hormone deficiency. *Horm Res* 1989; **31**:149–52.

7. Hyer W, Cotterill AM, Savage MO. Common causes of short stature detectable by a height surveillance programme. *J Med Screen* 1995; **2**:150–3.

8. Sarr M, Job JC, Chaussain JL, Golse B. [Psychogenic growth retardation. Critical study of diagnostic data]. *Arch Fr Pediatr* 1987;**44**:331–8.

9. Lee SG. Does child health surveillance contribute to the detection of growth problems? A study of children referred to the growth clinic in Leeds. Master of Medical Science in Child Health Thesis. Leeds: University of Leeds; 1996.

10. Herber SM, Milner RD. When are we diagnosing growth hormone deficiency? *Arch Dis Child* 1986; **61**:110–12.

11. Massa G, Verlinde F, De Schepper J, Thomas M, Bourguignon JP, Craen M, *et al*. Trends in age at diagnosis of Turner syndrome. *Arch Dis Child* 2005;**90**:267–8.

12. Cotterill AM. An international perspective on height surveillance. *J Med Screen* 1995;**2**:171–3.

13. Thomas M, Massa G, Craen M, de Zegher F, Bourguignon JP, Heinrichs C, *et al*. Prevalence and demographic features of childhood growth hormone deficiency in Belgium during the period 1986–2001. *Eur J Endocrinol* 2004;**151**:67–72.

14. Harris M, Hofman PL, Cutfield WS. Growth hormone treatment in children: review of safety and efficacy. *Paediatr Drugs* 2004;**6**:93–106.

15. Health for all Children. Growth monitoring: the Coventry consensus – website draft 6.1.2000. URL: http://www.healthforallchildren.co.uk/?DO= IMAGE&ID=coventry_concensus

16. Boersma B, Wit JM. Catch-up growth. *Endocr Rev* 1997;**18**:646–61.

17. Savendahl L, Davenport ML. Delayed diagnoses of Turner's syndrome: proposed guidelines for change. *J Pediatr* 2000;**137**:455–9.

18. Hunter I, Greene SA, MacDonald TM. Prevalence and aetiology of hypothyroidism in the young. *Arch Dis Child* 2000;**83**:207–10.

19. Hoey H. Diagnosing psychosocial growth deprivation. *J Pediatr Endocrinol* 1992;**5**:53–8.

20. Gohlke BC, Stanhope R. Final height in psychosocial short stature: is there complete catch-up? *Acta Paediatr* 2002;**91**:961–5.

21. 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 Technol Assess* 2002;**6**(18).

22. Erkula G, Jones KB, Sponseller PD, Dietz HC, Pyeritz RE. Growth and maturation in Marfan syndrome. *Am J Med Genet* 2002;**109**: 100–15.

23. Alexander SD, Hindmarsh P. The evaluation and management of tall stature. *Curr Paediatr* 2004; **14**:414–21.

24. Shankar RR, Pescovitz OH. Precocious puberty. *Adv Endocrinol Metab* 1995;**6**:55–89.

25. Voss LD, Sandberg DE. The psychological burden of short stature: evidence against. *Eur J Endocrinol* 2004;**151**:829–33.

26. Zimet GD, Owens R, Dahms W, Cutler M, Litvene M, Cuttler L. Psychosocial outcome of children evaluated for short stature. *Arch Pediatr Adolesc* Med 1997;**151**:1017–23.

27. British Medical Association Board of Science. *Preventing childhood obesity*. London: BMA; 2005.

28. Rudolf MCJ. The obese child. *Arch Dis Child Ed Pract* 2004;**89**:ep57–62.

83

29. Watson K. A study to assess the feasibility of establishing a robust surveillance system to determine prevalence, trends and variations in childhood overweight and obesity in Wolverhampton. Master in Public Health Thesis. Birmingham: University of Birmingham, 2005.

30. Reilly JJ, Methven E, McDowell ZC, Hacking B, Alexander D, Stewart L, *et al*. Health consequences of obesity. *Arch Dis Child* 2003;**88**:748–52.

31. Summerbell CD, Ashton V, Campbell KJ, Edmunds L, Kelly S, Waters E. Interventions for treating obesity in children. In *The Cochrane Library 2005*, Issue 3. Chichester: Wiley; 2005.

32. Speiser PW, Rudolf MCJ, Anhalt H, Camacho-Hubner C, Chiarelli F, Eliakim A, *et al*. Consensus statement: childhood obesity. *J Clin Endocrinol Metab* 2005;**90**:1871–87.

33. Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ. Interventions for preventing obesity in children. In *The Cochrane Library 2005*, Issue 3. Chichester: Wiley; 2005.

34. National Screening Committee. *Child health sub-group report: growth disorders*. Leeds: National Screening Committee; 2004.

35. Hulse JA, Schilg S. United Kingdom community growth screening 1994: a survey of current practice. *J Med Screen* 1995;**2**:154–6.

36. van Buuren S, van Dommelen P, Zandwijken GR, Grote FK, Wit JM, Verkerk PH. Towards evidence based referral criteria for growth monitoring. *Arch Dis Child* 2004;**89**:336–41.

37. de Onis M, Wijnhoven TMA, Onyango AW. Worldwide practices in child growth monitoring. *J Pediatr* 2004;**144**:461–5.

38. Jellinek D, Hall DM. How are children's growth problems diagnosed? *Child Care Health Dev* 1994; **20**:371–7.

39. Fry T. Monitoring children's growth. Introducing the new child growth standards. *Prof Care Mother Child* 1994;**4**:231–3.

40. Voss LD, Wilkin TJ, Bailey BJR, Betts PR. The reliability of height and height velocity in the assessment of growth (the Wessex Growth Study). *Arch Dis Child* 1991;**66**:833–7.

41. Voss LD. The short normal child: growth and psychosocial functioning in the prepubertal years. The Wessex Growth Study. PhD Thesis. Southampton: University of Southampton, 1997.

42. Cole TJ. A simple chart to identify non-familial short stature. *Arch Dis Child* 2000;**82**:173–6.

43. Wright CM, Cheetham TD. The strengths and limitations of parental heights as a predictor of attained height. *Arch Dis Child* 1999;**81**:257–60.

44. Mulligan J, Voss LD. Non-familial short stature. *Arch Dis Child* 2000;**83**:369–70.

45. Voss LD, Bailey BJ. Equipping the community to measure children's height: the reliability of portable instruments. *Arch Dis Child* 1994;**70**:469–71.

46. Roche AF, Sun SS, editors. Measurement and assessment. In *Human growth: assessment and interpretation*. Cambridge: Cambridge University Press; 2003. pp. 1–74.

47. Scottish Intercollegiate Guidelines Network. *Management of obesity in children and young people: a national clinical guideline*. Edinburgh: SIGN; 2003.

48. Centre for Community Child Health Royal Children's Hospital. *Child health screening and surveillance: a critical review of the evidence*. Melbourne: National Health and Medical Research Council; 2002.

49. Wright CM, Booth IW, Buckler JMH, Cameron N, Cole T, Healey MJH, *et al*. Growth reference charts for use in the United Kingdom. *Arch Dis Child* 2002;**86**:11–14.

50. Karlberg J, Cheung YB, Luo ZC. An update on the update of growth charts. *Acta Paediatr* 1999;**88**:797–9.

51. Freeman JV, Cole TJ, Chinn S, Jones PR, White EM, Preece MA. Cross sectional stature and weight reference curves for the UK, 1990. *Arch Dis Child* 1995;**73**:17–24.

52. Rudolf MC, Cole TJ, Krom AJ, Sahota P, Walker J. Growth of primary school children: a validation of the 1990 references and their use in growth monitoring. *Arch Dis Child* 2000;**83**:298–301.

53. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;**320**:1240–3.

54. Voss LD. Changing practice in growth monitoring. *BMJ* 1999;**318**:344–5.

55. Hall D, Elliman D. *Health for all children*. 4th ed. Oxford: Oxford University Press; 2003.

56. Ahmed ML, Allen AD, Dunger DB, Macfarlane A. The Oxford growth study: a district growth surveillance programme 1988–1994. *J Med Screen* 1995;**2**:160–3.

57. Hulse JA. Referral criteria for growth screening. *J Med Screen* 1995;**2**:168–70.

58. Hall DM. Monitoring children's growth. *BMJ* 1995;**311**:583–4.

59. Banerjee S, Morgan RJ, Rees SA, Latif AH. Height screening at school: ineffective without high standards and adequate resources. *Arch Dis Child* 2003;**88**:477–81.

60. Curran NC, Klimach V, Rankin W, McKinlay I, St Leger AS. Lessons for school health: a survey of facilities for screening tests and medical examinations in schools. *Ambul Child Health* 1997; **2**:211–19.

61. NHS Centre for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews*. 2nd ed. York: NHS Centre for Reviews and Dissemination, 2001.

62. Bavdekar SB, Pawar M. Evaluation of an Internet delivered pediatric diagnosis support system (ISABEL) in a tertiary care center in India. *Indian Pediatr* 2005;**42**:1086–91.

63. Centre for Reviews and Dissemination. *Production of DARE: what are the criteria for the inclusion of reviews on DARE?* URL: http://www.york.ac.uk/inst/crd/faq4.htm. Accessed March 2006.

64. NHS Centre for Reviews and Dissemination. *Improving access to cost-effectiveness information for health care decision making: the NHS Economic Evaluation Database*. 2nd ed. York: NHS Centre for Reviews and Dissemination; 2001.

65. Scottish Intercollegiate Guidelines Network. Scottish Intercollegiate Guidelines Network website. URL: www.sign.ac.uk. Accessed March 2006.

66. National Institute for Clinical Excellence. *Guideline development methods: information for national collaborating centres and guideline developers*. 2004. URL: www.nice.org.uk. Accessed March 2006.

67. 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.

68. Nixon J, Khan KS, Kleijnen J. Summarising economic evaluations in systematic reviews: a new approach. *BMJ* 2001;**322**:1596–8.

69. Hanley JA, Lippman-Hand A. If nothing goes wrong is everything all right? Interpreting zero numerators. *JAMA* 1983;**249**:1743–5.

70. Zhou X, McClish DK, Obuchowski NA. *Statistical methods in diagnostic medicine*. New York: Wiley-Interscience, 2002.

71. Aszkenasy M. *Audit of school entry growth screening*. Middlesbrough: Tees and North East Yorkshire NHS Trust, West Lane Hospital. 2005.

72. Routh K, Denley J, Rao JN. *A low cost method for measuring the prevalence of childhood obesity*. Birmingham: Department of Public Health, North Birmingham Primary Care Trust; 2005.

73. Anthony D, Stevens A. *Growth hormone in children (for growth hormone deficiency, Turner's syndrome, chronic renal failure and idiopathic short stature)*. Southampton: Wessex Institute for Health Research and Development (WIHRD). 1996. URL: http://www.doh.gov.uk/research/swro/rd/publicat/dec/dec57.htm

74. Wang LY, Yang Q, Lowry R, Wechsler H. Economic analysis of a school-based obesity prevention program. *Obes Res* 2003;**11**:1313–24.

75. Goldfield GS, Epstein LH, Kilanowski CK, Paluch RA, Kogut-Bossler B. Cost-effectiveness of group and mixed family-based treatment for childhood obesity. *Int J Obes* 2001;**25**:1843–9.

76. Agwu JC, Leishenring A, Darnley I. Community growth monitoring in practice. *Arch Dis Child* 2004;**89**:394.

77. Ahmed ML, Allen AD, Sharma A, Macfarlane JA, Dunger DB. Evaluation of a district growth screening programme: the Oxford growth study. *Arch Dis Child* 1993;**69**:361–5.

78. Cotterill AM, Majrowski WH, Hearn S, Jenkins S, Preece MA, Savage MO. The potential effect of the UK 1990 height centile charts on community growth surveillance. *Arch Dis Child* 1996;**74**:452–4.

79. Cernerud L, Edding E. The value of measuring height and weight of schoolchildren. *Paediatr Perinat Epidemiol* 1994;**8**:365–72.

80. Cowan FJ, Gregory JW. Assessment of the effects of intensive training of community based staff in the screening of children for disorders of growth and puberty. *J Med Screen* 2001;**8**:15–7.

81. de la Puente ML, Ibanez L, Marcos MV, Peix MA. [Primary care population screening for growth]. *Aten Primaria* 1999;**23**:346–51.

82. Hearn S, Cotterill AM, Majrowski W, Rohan C, Jenkins S, Savage MO. Construction of a community height surveillance programme: the Hackney growth initiative. *J Med Screen* 1995; **2**:31–4.

83. Keller E, Gausche R, Meigen C, Keller A, Burmeister J, Kiess W. Auxological computer based network for early detection of disorders of growth and weight attainment. *J Pediatr Endocrinol* 2002;**15**:149–56.

84. Kiess W, Gausche R, Keller A, Burmeister J, Willgerodt H, Keller E. Computer-guided, population-based screening system for growth disorders (CrescNet) and on-line generation of normative data for growth and development. *Horm Res* 2001;**56** Suppl 1:59–66.

85. Keller E, Burmeister J, Gausche R, Keller A, Hermanussen M, Kiess W. [Model program for the early detection and optimal treatment of disorders of growth and physical development using a medical competence network]. *Z Ärztl Fortbild Qualitätssich* 2000;**94**:695–8.

86. Lacey KA, Parkin JM. Causes of short stature. *Lancet* 1974;**i**:42–5.

87. Lindsay R, Feldkamp M, Harris D, Robertson J, Rallison M. Utah Growth Study: growth standards and the prevalence of growth hormone deficiency. *J Pediatr* 1994;**125**:29–35.

88. Lipman TH, Hench KD, Benyi T, Delaune J, Gilluly KA, Johnson L, *et al*. A multicentre

randomised controlled trial of an intervention to improve the accuracy of linear growth measurement. *Arch Dis Child* 2004;**89**:342–6.

89. Mulligan J, Voss LD, McCaughey ES, Bailey BJ, Betts PR. Growth monitoring: testing the new guidelines. *Arch Dis Child* 1998;**79**:318–22.

90. van Buuren S, Bonnemaijer-Kerckhoffs DJA, Grote FK, Wit JM, Verkerk PH. Many referrals under Dutch short stature guidelines. *Arch Dis Child* 2004;**89**:351–2.

91. Vimpani GV, Vimpani AF, Pocock SJ, Farquhar JW. Differences in physical characteristics, perinatal histories, and social backgrounds between children with growth hormone deficiency and constitutional short stature. *Arch Dis Child* 1981;**56**:922–8.

92. Vimpani GV, Vimpani AF, Lidgard GP, Cameron EH, Farquhar JW. Prevalence of severe growth hormone deficiency. *BMJ* 1977;**ii**:427–30.

93. Voss LD, Mulligan J, Betts PR, Wilkin TJ. Poor growth in school entrants as an index of organic disease: the Wessex Growth Study. *BMJ* 1992;**305**:1400–2.

94. Voss L, Walker J, Lunt H, Wilkin T, Betts P. The Wessex Growth Study: first report. *Acta Paediatr Scand* 1989;**78**:65–72.

95. Welch NM, Saulsbury FT, Kesler RW. The value of the preschool examination in screening for health problems. *J Pediatr* 1982;**100**:232–4.

96. White E, Wilson A, Greene SA, Berry W, McCowan C, Cairns A, *et al*. Growth screening and urban deprivation. *J Med Screen* 1995;**2**:140–4.

97. Bedogni G, Iughetti L, Ferrari M, Malavolti M, Poli M, Bernasconi S, *et al*. Sensitivity and specificity of body mass index and skinfold thicknesses in detecting excess adiposity in children aged 8–12 years. *Ann Hum Biol* 2003; **30**:132–9.

98. Ebbeling CB, Backstrand JR, Rodriguez NR. Screening indices for pediatric obesity. *Nutr Res* 1999;**19**:805–15.

99. Ellis KJ, Abrams SA, Wong WW. Monitoring childhood obesity: assessment of the weight/height^2 index. *Am J Epidemiol* 1999; **150**:939–46.

100. Himes JH, Bouchard C. Validity of anthropometry in classifying youths as obese. *Int J Obesity* 1989; **13**:183–93.

101. Lazarus R, Baur L, Webb K, Blyth F. Body mass index in screening for adiposity in children and adolescents: systematic evaluation using receiver operating characteristic curves. *Am J Clin Nutr* 1996;**63**:500–6.

102. Marshall JD, Hazlett CB, Spady DW, Conger PR, Quinney HA. Validity of convenient indicators of obesity. *Hum Biol* 1991;**63**:137–53.

103. Mast M, Langnase K, Labitzke K, Bruse U, Preuss U, Muller MJ. Use of BMI as a measure of overweight and obesity in a field study on 5–7 year old children. *Eur J Nutr* 2002;**41**:61–7.

104. Reilly JJ, Savage SAH, Ruxton CHS, Kirk TR. Assessment of obesity in a community sample of prepubertal children. *Int J Obes Relat Metab Disord* 1999;**23**:217–19.

105. Reilly JJ, Dorosty AR, Emmett PM. Identification of the obese child: adequacy of the body mass index for clinical practice and epidemiology. *Int J Obes Relat Metab Disord* 2000;**24**:1623–7.

106. Schaefer F, Georgi M, Wuhl E, Scharer K. Body mass index and percentage fat mass in healthy German schoolchildren and adolescents. *Int J Obes Relat Metab Disord* 1998;**22**:461–9.

107. Wickramasinghe VP, Cleghorn GJ, Edmiston KA, Murphy AJ, Abbott RA, Davies PSW. Validity of BMI as a measure of obesity in Australian white Caucasian and Australian Sri Lankan children. *Ann Hum Biol* 2005;**32**:60–71.

108. National Institute for Clinical Excellence. *Guidance on the use of human growth hormone (somatropin) in children with growth failure*. London: National Institute for Clinical Excellence; 2002. Technology Appraisal No. 42. URL: http://www.nice.org.uk/ pdf/HGHinChild-42-ALS.pdf

109. Mulligan J, Voss LD. Identifying very fat and very thin children: test of criterion standards for screening test. *BMJ* 1999;**319**:1103–4.

110. Ahmed ML, Yudkin PL, Macfarlane JA, McPherson K, Dunger DB. Are measurements of height made by health visitors sufficiently accurate for routine screening of growth? *Arch Dis Child* 1990;**65**:1345–8.

111. Cotterill AM, Majrowski WH, Hearn SJ, Jenkins S, Savage MO. Assessment of the reliability of school nurse height measurements in an inner-city population. (The Hackney Growth Initiative). *Child Care Health Dev* 1993;**19**:159–65.

112. National Institute for Clinical Excellence. *Guide to the methods of technology appraisal*. London: National Institute for Clinical Excellence; 2004. URL: http://www.nice.org.uk/pdf/TAP_Methods.pdf

113. Birch S, Gafni A. The 'NICE' approach to technology assessment: an economics perspective. *Health Care Manag Sci* 2004;**7**:35–41.

114. Office for National Statistics. *Birth statistics: review of the Registrar General on births and patterns of family building in England and Wales, 2004*. London: Office for National Statistics; 2005.

115. Briggs A. Handling uncertainty in cost-effectiveness models. *PharmacoEconomics* 2000; **17**:479–500.

116. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves: facts, fallacies and frequently asked questions. *Health Econ* 2004; **13**:405–15.

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

118. Briggs A. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Med Decis Making* 2003;**23**:341–50.

119. 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**(19).

120. Personal Social Services Research Unit. Curtis L, Netten A. *Unit costs of health and social care 2006*. Canterbury: Personal Social Services Research Unit, University of Kent; 2006. http://www.pssru. ac.uk/uc/uc2006contents.htm. Accessed March 2006.

121. Department of Health. *NHS reference costs 2004: NHS trust reference cost index*. 2004. URL: http://www.dh.gov.uk/PublicationsAndStatistics/ Publications/PublicationsPolicyAndGuidance/ PublicationsPolicyAndGuidanceArticle/fs/en? CONTENT_ID=4105545&chk=znAfqu. Accessed March 2006.

122. Tapsfield R, Collier F. *The cost of foster care: investing in our children's future*. London: British Association for Adoption and Fostering and The Fostering Network; 2005. URL: http://www.fostering.net/campaigns/cost_of_foster_ care/CFC_report.pdf

123. Macfarlane A. Should we screen for growth problems in children? The pragmatic answer for 1994–5. *J Med Screen* 1994;**1**:50–9.

124. Joint Formulary Committee. *British National Formulary*. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 2005. URL: http://www.bnf.org/bnf/bnf/current/ openat/index.htm. Accessed March 2006.

125. Organisation for Economic Co-operation and Development. *Purchasing Power Parities (PPP), PPP data*. Paris: OECD; 2005. URL: http://www.oecd.org/linklist/0,2678,en_2649_ 34357_2734617_1_1_1_1,00.html. Accessed March 2006.

126. Voss LD, Bailey BJ, Cumming K, Wilkin TJ, Betts PR. The reliability of height measurement (the Wessex Growth Study). *Arch Dis Child* 1990; **65**:1340–4.

127. Voss LD, Mulligan J, Betts PR. Short stature at school entry – an index of social deprivation? (The Wessex Growth Study). *Child Care Health Dev* 1998;**24**:145–56.

128. Hearn S, Majrowski W, Jenkins S, Savage M, Cotterill AM. Setting up a height surveillance programme: recommendations based on the Hackney growth initiative. *J Med Screen* 1995; **2**:157–9.

Appendix 1

Advisory panel members

The panel members were as follows:
Mr Stephen Palmer, Centre for Health Economics, University of York; Mr Tam Fry, Child Growth Foundation, London; Ms Lynne Morris, Turner Syndrome Support Society; Dr Sarah Lee, Consultant Community

Paediatrician and Director of Leeds Community Growth Monitoring Programme, Leeds; and Sisters Jenny Walker and Amanda Stoner, growth and endocrine specialist nurses directly involved in the delivery of growth training programmes, University of Leeds.

Appendix 2

Protocol changes

A large number of studies examine trends in obesity in order to establish baseline data for a given population. As national statistics exist on prevalence of obesity, we did not consider it to be useful to include these types of study in the review. Therefore, after discussion within the review team, we excluded studies that only provided prevalence data on obesity.

Diagnostic cohort studies are less likely to be conducted in the field of growth monitoring as patients with normal growth will not be subject to full investigation. In the light of this, we decided to relax the study design inclusion criteria to allow diagnostic case control studies in addition to diagnostic cohort studies.

EVPI analysis was not conducted. The quantification of uncertainty was difficult in the stature model given the sources of evidence. For this reason, the probabilities of being offered treatments and dropping out of treatment and the cost estimates were given point estimates and no probability distributions. Furthermore, QALY estimates were given uniform probability distributions. Since EVPI analysis requires the accurate assessment of uncertainty, if applied to the data available it would inevitably be unreliable and was therefore not considered appropriate in this case. The same reasoning applies to the obesity model, but to a greater extent.

Appendix 3
Detailed search strategies

The following searches were run in order to identify published and unpublished literature on screening for child growth disorders.

Databases/resources searched
MEDLINE
MEDLINE In Process
EMBASE
CINAHL
PsycInfo
SIGLE
Sociological Abstracts
LILACS (Latin American and Caribbean Health Sciences Literature)
NHS Economic Evaluation Database
DARE (Database of Abstracts of Reviews of Effects)
HTA database
Pascal
SciSearch
Dissertation Abstracts
Inside Conferences
Science and Technology Proceedings
Cochrane Controlled Trials Register
National Research Register
OHE Health Economic Evaluations Database
metaRegister of Clinical Trials
OMNI
Economic Working Papers Archive
Google
Copernic

Search strategies
MEDLINE (Ovid) 1966–June 2005
Searched: 4 July 2005
Results: 21,994

1. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 measur$).ti,ab.
2. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 monitor$).ti,ab.
3. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 screen$).ti,ab.
4. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 examin$).ti,ab.
5. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 identif$).ti,ab.
6. or/1-5
7. Population Surveillance/
8. Physical Examination/
9. Mass Screening/
10. Early Diagnosis/
11. Child Health Services/
12. School Health Services/
13. surveillance.ti,ab.
14. monitor$ program$.ti,ab.
15. (screen$ or diagnos$).ti,ab.
16. (child$ health service$ or school health service$).ti,ab.
17. (physical$ examin$ or physical check-up or physical checkup).ti,ab.
18. or/7-17
19. ANTHROPOMETRY/
20. Body Weight/
21. Body Height/
22. Body Mass Index/
23. Child Development/
24. Growth/
25. *Growth Disorders/
26. Body Size/
27. (height or weight or stature).ti,ab.
28. (bmi or body mass index or obesity).ti,ab.
29. growth.ti,ab.
30. child$ develop$.ti,ab.
31. anthropometr$.ti,ab.
32. or/19-31
33. 18 and 32
34. 6 or 33
35. exp child/
36. child$.ti,ab.
37. (school-age$ or schoolage$).ti,ab.
38. schoolchild$.ti,ab.
39. (boy or boys or girl or girls).ti,ab.
40. or/35-39
41. 34 and 40
42. Developing Countries/
43. (third world or 3rd world).ti,ab.
44. (developing world or developing countr$ or developing nation$).ti,ab.
45. or/42-44
46. 41 not 45
47. Animals/
48. Humans/
49. 47 not (47 and 48)
50. 46 not 49

MEDLINE In Process (Ovid) 1 July 2005
Searched: 4 July 2005
Results: 462

HMIC (Ovid) May 2005
Searched: 4 July 2005
Results: 133

1. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 measur$).ti,ab.
2. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 monitor$).ti,ab.
3. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 screen$).ti,ab.
4. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 examin$).ti,ab.
5. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 identif$).ti,ab.
6. or/1-5
7. surveillance.ti,ab.
8. monitor$ program$.ti,ab.
9. (screen$ or diagnos$).ti,ab.
10. (child$ health service$ or school health service$).ti,ab.
11. (physical$ examin$ or physical check-up or physical checkup).ti,ab.
12. or/7-11
13. (height or weight or stature).ti,ab.
14. (bmi or body mass index or obesity).ti,ab.
15. growth.ti,ab.
16. child$ develop$.ti,ab.
17. anthropometr$.ti,ab.
18. or/13-17
19. 12 and 18
20. 6 or 19
21. child$.ti,ab.
22. (school-age$ or schoolage$).ti,ab.
23. schoolchild$.ti,ab.
24. (boy or boys or girl or girls).ti,ab.
25. or/21-24
26. 20 and 25
27. (third world or 3rd world).ti,ab.
28. (developing world or developing countr$ or developing nation$).ti,ab.
29. 27 or 28
30. 26 not 29

EMBASE (Ovid) 1980–July 2005
Searched: 4 July 2005
Results: 18,284

1. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 measur$).ti,ab.

2. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 monitor$).ti,ab.
3. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 screen$).ti,ab.
4. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 examin$).ti,ab.
5. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 identif$).ti,ab.
6. or/1-5
7. physical examination/
8. mass screening/
9. early diagnosis/
10. child health care/
11. school health service/
12. surveillance.ti,ab.
13. monitor$ program$.ti,ab.
14. (screen$ or diagnos$).ti,ab.
15. (child$ health service$ or school health service$).ti,ab.
16. (physical$ examin$ or physical check-up or physical checkup).ti,ab.
17. or/7-16
18. anthropometry/
19. Body Weight/
20. Body Height/
21. Body Mass/
22. Child Development/
23. growth/
24. *Growth Disorder/
25. child growth/
26. Body Size/
27. (height or weight or stature).ti,ab.
28. (bmi or body mass index or obesity).ti,ab.
29. growth.ti,ab.
30. child$ develop$.ti,ab.
31. anthropometr$.ti,ab.
32. or/18-30
33. 17 and 32
34. 6 or 33
35. exp child/
36. child$.ti,ab.
37. (school-age$ or schoolage$).ti,ab.
38. schoolchild$.ti,ab.
39. (boy or boys or girl or girls).ti,ab.
40. or/35-39
41. 34 and 40
42. developing country/
43. (third world or 3rd world).ti,ab.
44. (developing world or developing countr$ or developing nation$).ti,ab.
45. or/42-44
46. 41 not 45
47. exp animal/

48. Nonhuman/
49. 47 or 48
50. Human/
51. 49 not (49 and 50)
52. 46 not 51

CINAHL (Ovid) 1982–June 2005
Searched: 4 July 2005
Results: 2175

1. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 measur$).ti,ab.
2. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 monitor$).ti,ab.
3. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 screen$).ti,ab.
4. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 examin$).ti,ab.
5. ((growth or height or weight or bmi or body mass index or obesity or stature) adj3 identif$).ti,ab.
6. or/1-5
7. Physical Examination/
8. Health Screening/
9. Child Health Services/
10. School Health Services/
11. surveillance.ti,ab.
12. monitor$ program$.ti,ab.
13. (screen$ or diagnos$).ti,ab.
14. (child$ health service$ or school health service$).ti,ab.
15. (physical$ examin$ or physical check-up or physical checkup).ti,ab.
16. or/7-15
17. "Body Weights and Measures"/
18. Body Weight/
19. Body Height/
20. Body Mass Index/
21. Child Development/
22. growth/
23. *Growth Disorders/
24. Body Constitution/
25. (height or weight or stature).ti,ab.
26. (bmi or body mass index or obesity).ti,ab.
27. growth.ti,ab.
28. child$ develop$.ti,ab.
29. anthropometr$.ti,ab.
30. or/17-29
31. 16 and 30
32. 6 or 31
33. exp child/
34. child$.ti,ab.
35. (school-age$ or schoolage$).ti,ab.
36. schoolchild$.ti,ab.

37. (boy or boys or girl or girls).ti,ab.
38. or/33-37
39. 32 and 38
40. developing countries/
41. (third world or 3rd world).ti,ab.
42. (developing world or developing countr$ or developing nation$).ti,ab.
43. or/40-42
44. 39 not 43

PsycINFO (WebSPIRS) 1872–June 2005
Searched: 4 July 2005
Results: 1913

SIGLE (WebSPIRS) 1980–December 2004
Searched: 4 July 2005
Results: 8

#1 ((growth or height or weight or bmi or body mass index or obesity or stature) near3 (measur* or monitor* or screen* or examin* or identif*)) in ti,ab
#2 surveillance in ti,ab
#3 monitor* program* in ti,ab
#4 (screen* or diagnos*) in ti,ab
#5 (child* health service* or school health service*) in ti,ab
#6 (physical* examin* or physical check-up or physical checkup) in ti,ab
#7 #2 or #3 or #4 or #5 or #6
#8 (height or weight or stature) in ti,ab
#9 (bmi or body mass index or obesity) in ti,ab
#10 growth in ti,ab
#10 growth in ti,ab
#11 child* develop* in ti,ab
#12 anthropometr* in ti,ab
#13 #8 or #9 or #10 or #11 or #12
#14 #7 and #13
#15 #1 or #14
#16 child* in ti,ab
#17 (school-age* or schoolage*) in ti,ab
#18 schoolchild* in ti,ab
#19 (boy or boys or girl or girls) in ti,ab
#20 #16 or #17 or #18 or #19
#21 #15 and #20
#22 (third world or 3rd world) in ti,ab
#23 (developing world or developing countr* or developing nation*) in ti,ab
#24 #22 or #23
#25 #21 not #24

Sociological Abstracts (CSA Illumina host) 1963–June 2005
Searched: 4 July 2005
Results: 439

#1 (growth or height or weight or bmi or body mass index or obesity or stature) within 3 measur*

95

#2   (growth or height or weight or bmi or body mass index or obesity or stature) within 3 monitor*

#3   (growth or height or weight or bmi or body mass index or obesity or stature) within 3 screen*

#4   (growth or height or weight or bmi or body mass index or obesity or stature) within 3 examin*

#5   (growth or height or weight or bmi or body mass index or obesity or stature) within 3 identif*

#6   #1 or #2 or #3 or #4 or #5

#7   surveillance or monitor* program* or screen* or diagnos*

#8   child* health service* or school health service*

#9   physical* examin* or physical check-up or physical checkup

#10   #7 or #8 or #9

#11   height or weight or stature

#12   bmi or body mass index or obesity

#13   growth or child* develop* or anthropometr*

#14   #11 or #12 or #13

#15   #10 and #14

#16   #6 or #15

#17   child* or school-age* or schoolage* or schoolchild* or boy or boys or girl or girls

#18   #16 and #17

#19   third world or 3rd world

#20   developing world or developing countr* or developing nation*

#21   #19 or #20

#22   #18 not #21

NHS EED – Public access database: All years
Searched: 5 July 2005
Results: 65

NHS EED – CRD Administrative database:
All years
Searched: 5 July 2005
Results: 73

DARE – Public access database: All years
Searched: 8 July 2005
Results: 93

DARE – CRD Administrative database:
All years
Searched: 8 July 2005
Results: 184

HTA database: All years
Searched: 8 July 2005
Results: 11

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)measur$

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)monitor$

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)screen$

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)examin$

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)identif$

s s1 or s2 or s3 or s4 or s5

s surveillance or monitor$(w)program$ or screen$ or diagnos$

s child$(w)health(w)service$ or school(w)health(w)service$

s physical$(w)examin$ or physical(w)check(w)up or physical(w)checkup

s s7 or s8 or s9

s height or weight or stature

s bmi or body(w)mass(w)index or obesity

s growth or child$(w)develop$ or anthropometr$

s s11 or s12 or s13

s s10 and s14

s s6 or s15

s child$ or school(w)age$ or schoolage$ or schoolchild$ or boy or boys or girl or girls

s s16 and s17

s third(w)world or 3rd(w)world

s developing(w)world or developing(w)countr$ or developing(w)nation$

s s19 or s20

s s18 andnot s21

Cochrane CENTRAL database 2005 Issue 2
Searched: 5 July 2005
Records found: 1259

#1   (growth or height or weight or bmi or body next mass next index or obesity or stature) near measur* in All Fields in all products

#2   (growth or height or weight or bmi or body next mass next index or obesity or stature) near monitor* in All Fields in all products

#3   (growth or height or weight or bmi or body next mass next index or obesity or stature) near screen* in All Fields in all products

#4   (growth or height or weight or bmi or body next mass next index or obesity or stature) near examin* in All Fields in all products

#5   (growth or height or weight or bmi or body next mass next index or obesity or stature) near identif* in All Fields in all products

#6 (#1 OR #2 OR #3 OR #4 OR #5)

#7 MeSH descriptor Population Surveillance, this term only in MeSH products

#8 MeSH descriptor Physical Examination, this term only in MeSH products

#9 MeSH descriptor Mass Screening, this term only in MeSH products

#10 MeSH descriptor Early Diagnosis, this term only in MeSH products

#11 MeSH descriptor Child Health Services, this term only in MeSH products

#12 MeSH descriptor School Health Services, this term only in MeSH products

#13 surveillance in All Fields in all products

#14 monitor* next program* in All Fields in all products

#15 screen* or diagnos* in All Fields in all products

#16 (child* next health next service*) or (school next health next service*) in All Fields in all products

#17 (physical* next examin*) or (physical next check next up) or (physical next checkup) in All Fields in all products

#18 (#7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17)

#19 MeSH descriptor Anthropometry, this term only in MeSH products

#20 MeSH descriptor Body Weight, this term only in MeSH products

#21 MeSH descriptor Body Height, this term only in MeSH products

#22 MeSH descriptor Body Mass Index, this term only in MeSH products

#23 MeSH descriptor Child Development, this term only in MeSH products

#24 MeSH descriptor Growth, this term only in MeSH products

#25 MeSH descriptor Body Size, this term only in MeSH products

#26 height or weight or stature in All Fields in all products

#27 bmi or (body next mass next index) or obesity in All Fields in all products

#28 growth in All Fields in all products

#29 child* next develop* in All Fields in all products

#30 anthropometr* in All Fields in all products

#31 (#19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30)

#32 (#18 AND #31)

#33 (#6 OR #32)

#34 MeSH descriptor Child explode all trees in MeSH products

#35 child* in All Fields in all products 45040 edit delete

#36 school next age* or schoolage* in All Fields in all products

#37 schoolchild* in All Fields in all products

#38 boy or boys or girl or girls in All Fields in all products

#39 (#34 OR #35 OR #36 OR #37 OR #38)

#40 MeSH descriptor Developing Countries, this term only in MeSH products

#41 MeSH descriptor Developing Countries, this term only in MeSH products

#42 third next world in All Fields in all products

#43 developing next world or developing next countr* or developing next nation* in All Fields in all products

#44 (#40 OR #41 OR #42)

#45 (#33 AND #39)

#46 (#45 AND NOT #44)

National Research Register 2005 Issue 2
Searched: 5 July 2005
Records found:

#1. ((growth near measure*) or (height near measure*) or (weight near measure*) or (bmi near measure*) or ((body next mass next index) near measure*) or (obesity near measure*) or (stature near measure*))

#2. ((growth near monitor*) or (height near monitor*) or (weight near monitor*) or (bmi near monitor*) or ((body next mass next index) near monitor*) or (obesity near monitor*) or (stature near monitor*))

#3. ((growth near screen*) or (height near screen*) or (weight near screen*) or (bmi near screen*) or ((body next mass next index) near screen*) or (obesity near screen*) or (stature near screen*))

#4. ((growth near examin*) or (height near examin*) or (weight near examin*) or (bmi near examin*) or ((body next mass next index) near examin*) or (obesity near examin*) or (stature near examin*))

#5. ((growth near identif*) or (height near identif*) or (weight near identif*) or (bmi near identif*) or ((body next mass next index) near identif*) or (obesity near identif*) or (stature near identif*))

#6. (#1 or #2 or #3 or #4 or #5)

#7. POPULATION SURVEILLANCE single term (MeSH)

#8. PHYSICAL EXAMINATION single term (MeSH)

#9. MASS SCREENING single term (MeSH)

#10. EARLY DIAGNOSIS single term (MeSH)

#11. CHILD HEALTH SERVICES single term (MeSH)

#12. SCHOOL HEALTH SERVICES single term (MeSH)
#13. surveillance
#14. (monitor* next program*)
#15. (screen* or diagnos*)
#16. ((child* next health next service*) or (school next health next service*))
#17. ((physical* next examin*) or (physical next check next up) or (physical next checkup))
#18. (#7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17)
#19. ANTHROPOMETRY single term (MeSH)
#20. BODY WEIGHT single term (MeSH)
#21. BODY HEIGHT single term (MeSH)
#22. BODY MASS INDEX single term (MeSH)
#23. CHILD DEVELOPMENT single term (MeSH)
#24. GROWTH single term (MeSH)
#25. BODY CONSTITUTION single term (MeSH)
#26. (height or weight or stature)
#27. (bmi or (body next mass next index) or obesity)
#28. growth
#29. (child* next develop*)
#30. anthropometr*
#31. (#19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30)
#32. (#18 and #31)
#33. (#6 or #32)
#34. CHILD explode tree 1 (MeSH)
#35. child*
#36. ((school next age*) or schoolage*)
#37. schoolchild*
#38. (boy or boys or girl or girls)
#39. (#34 or #35 or #36 or #37 or #38)
#40. (#33 and #39)
#41. DEVELOPING COUNTRIES single term (MeSH)
#42. (third next world)
#43. ((developing next world) or (developing next countr*) or (developing next nation*))
#44. (#41 or #42 or #43)
#45. (#40 and (not #44))

Pascal (Dialog) 1973–June 2005
Searched: 5 July 2005
Results: 3312

SciSearch (Dialog) 1974–June 2005
Searched: 5 July 2005
Results: 5648

Dissertation Abstracts (Dialog) 1861–June 2005
Searched: 5 July 2005
Results: 618

s ((growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(measure or measures or measured or measurement? or measuring))/ti,ab
s ((growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(monitor or monitors or monitored or monitoring))/ti,ab
s ((growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(screen or screens or screened or screening))/ti,ab
s ((growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(examine or examines or examined or examination or examining))/ti,ab
s ((growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(identify or identifies or identified or identification or identifying))/ti,ab
s s1 or s2 or s3 or s4 or s5
s (surveillance or ((monitor or monitors or monitored or monitoring)(w)(program or programs or programme or programmes)) or screen or screens or screened or screening or diagnose or diagnosis or diagnosing)/ti,ab
s (((child or children?)(w)health(w)(service or services)) or (school(w)health(w)(service or services)))/ti,ab
s ((physical?(w)(examine or examines or examined or examination or examining)) or physical(w)check(w)up or physical(w)checkup)/ti,ab
s s7 or s8 or s9
s (height or weight or stature)/ti,ab
s (bmi or body(w)mass(w)index or obesity)/ti,ab
s (growth or ((child or children?)(w)(develop or develops or developed or development or developing)) or anthropometr?)/ti,ab
s s11 or s12 or s13
s s10 and s14
s s6 or s15
s (child or children? or school(w)age or school(w)aged or schoolage? or schoolchild? or boy or boys or girl or girls)/ti,ab
s s16 and s17
s (third(w)world or 3rd(w)world)/ti,ab
s (developing(w)world or developing(w)(country or countries) or developing(w)(nation or nations))/ti,ab
s s19 or s20
s s18 not s21

Inside Conferences (Dialog) 1993–June 2005
Searched: 5 July 2005
Results: 22

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or

stature)(3w)(measure or measures or measured or measurement? or measuring)

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(monitor or monitors or monitored or monitoring)

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(screen or screens or screened or screening)

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(examine or examines or examined or examination or examining)

s (growth or height or weight or bmi or body(w)mass(w)index or obesity or stature)(3w)(identify or identifies or identified or identification or identifying)

s s1 or s2 or s3 or s4 or s5

s surveillance or ((monitor or monitors or monitored or monitoring)(w)(program or programs or programme or programmes)) or screen or screens or screened or screening or diagnose or diagnosis or diagnosing

s ((child or children?)(w)health(w)(service or services)) or (school(w)health(w)(service or services))

s (physical?(w)(examine or examines or examined or examination or examining)) or physical(w)check(w)up or physical(w)checkup

s s7 or s8 or s9

s height or weight or stature

s bmi or body(w)mass(w)index or obesity

s growth or ((child or children?)(w)(develop or develops or developed or development or developing)) or anthropometr?

s s11 or s12 or s13

s s10 and s14

s s6 or s15

s child or children? or school(w)age or school(w)aged or schoolage? or schoolchild? or boy or boys or girl or girls

s s16 and s17

s third(w)world or 3rd(w)world

s developing(w)world or developing(w)(country or countries) or developing(w)(nation or nations)

s s19 or s20

s s18 not s21

Science and Technology Proceedings (ISI Web of Knowledge) 1990–2005
Searched: 5 July 2005
Results: 1424

((growth or height or weight or bmi or body mass index or obesity or stature) same (measur* or monitor* or screen* or examin* or identify*)) and

(child* or school-age* or schoolage* or schoolchild* or boy or boys or girl or girls)

or

(surveillance or monitor* program* or screen* or diagnos* or child* health service* or school health service* or physical* examin* or physical check-up or physical checkup) and (height or weight or stature or bmi or body mass index or obesity or growth or child* develop* or anthropometr*) and (child* or school-age* or schoolage* or schoolchild* or boy or boys or girl or girls)

not

(third world or 3rd world or developing world or developing countr* or developing nation*)

LILACS All years
Date searched: 16 May 2005
Results: 309

growth or height or weight or bmi or body mass index or obesity or stature

and

screening or screen or diagnose or diagnosis or surveillance or examination or examine or measure or measurement or identify or identification

and

child or children or school or boys or boy or girls or girl

OHE Health Economic Evaluation Database (HEED) All years
Date searched: 20 May 2005
Results: 82

growth or height or weight or bmi or body mass index or obesity or stature

and

screening or screen or diagnose or diagnosis or surveillance or examination or examine or measure or measurement or identify or identification

and

child or children or school or boys or boy or girls or girl

**meta Register of Clinical Trials (mRCT)
(http://controlled-trials.com/mrct/) All years
Searched: 20 May 2005
Results: 0**

(growth or height or weight or obes%) and
(screen% or diagnos% or surveillance or examin%
or measure% or identify%) and (child%)

**OMNI (http://omni.ac.uk/) All years
Date searched: 20 May 2005
Results: 0**

(growth or height or weight or obese or obesity)
and (screen or screening or diagnose or diagnosis
or surveillance or examine or examination or
measure or measurement or identify or
identification) and (child or children)

**Economic Working Papers Archive
(http://econwpa.wustl.edu/) All years
Searched: 23 May 2005
Results: 0**

growth or height or weight or obesity or bmi or
body mass index or stature or child development

**Copernic (http://www.copernic.com/) All years
Date searched: 23 May 2005
Results: 3**

growth or height or weight or bmi or body mass
index or obesity or stature

search within results:
(screening or screen or diagnose or diagnosis or
surveillance or examination or examine or
measure or measurement or identify or
identification) and (child or children or school or
boys or boy or girls or girl)

Quality of life searches

The search strategy was designed to retrieve
records on QoL in children with the target
conditions assessed in this study. The strategy uses
a sensitive QoL search filter, combined with
specific search terms for each of the target
conditions to be assessed. The results were limited
to those studies referring to children only.

Databases searched
MEDLINE
MEDLINE In Process
EMBASE
CINAHL

CENTRAL
NHS Economic Evaluation Database
OHE Health Economic Evaluations Database

Limits
No date limits.
Studies referring to children only.
No study design limits.

Search strategies
**MEDLINE (Ovid) 1966–October 2005
Searched: 21 October 2005
Records found: 478**

1. (sf36 or sf 36).ti,ab.
2. (eq5d or eq 5d or euroqol).ti,ab.
3. (short form 36 or shortform 36 or sf thirtysix
   or sf thirty six or shortform thirtysix or
   shortform thirty six or short form thirtysix or
   short form thirty six).ti,ab.
4. (hrql or hrqol or qol or hql or hqol).ti,ab.
5. (hye or hyes or health$ year$ equivalent$ or
   health utilit$).ti,ab.
6. rosser.ti,ab.
7. (person trade off$ or person tradeoff$ or
   standard gamble$ or time trade off or time
   tradeoff or tto).ti,ab.
8. (disutilities or disutility or daly or disability
   adjusted life).ti,ab.
9. (qaly$ or qualy$ or quality adjusted life or
   quality of life or life quality).ti,ab.
10. qwb.ti,ab.
11. (quality of wellbeing or quality of well being
    or index of well being or index of
    wellbeing).ti,ab.
12. factor analysis.ti,ab.
13. preference based.ti,ab.
14. (health status or health state$).ti,ab.
15. (state adj2 (value or values or valuing or
    valued)).ti,ab.
16. hspv.ti,ab.
17. quality adjusted life year/
18. "Quality of Life"/
19. Health Status/
20. Health Status Indicators/
21. Sickness Impact Profile/
22. (utilit$ approach$ or health gain or hui or
    hui2 or hui 2 or hui3 or hui 3).ti,ab.
23. (categor$ scal$ or linear scal$ or linear
    analog$ scal$ or visual scal$ or magnitude
    estimat$).ti,ab.
24. (multiattribute$ health or multi attribute$
    health).ti,ab.
25. health measurement$.ti,ab.
26. health survey questionnaire$.ti,ab.
27. (general health questionnaire$ or
    ghq).ti,ab.

28. (multiattribute$ theor$ or multi attribute$ theor$ or multiattribute$ analys$ or multi attribute$ analys$).ti,ab.
29. classification illness state$.ti,ab.
30. (health adj2 utilit$).ti,ab.
31. (multiattribute$ utilit$ or multi attribute$ utilit$).ti,ab.
32. willingness pay.ti,ab.
33. theory utilit$.ti,ab.
34. or/1-33
35. Growth Disorders/
36. (growth hormone disease$ or growth hormone deficienc$).ti,ab.
37. growth hormone treatment$.ti,ab.
38. idiopathic short stature.ti,ab.
39. Prader-Willi Syndrome/
40. prader-willi syndrome.ti,ab.
41. Kidney Failure, Chronic/
42. chronic renal failure.ti,ab.
43. Turner-Syndrome/
44. turner$ syndrome.ti,ab.
45. juvenile hypothyroidism.ti,ab.
46. Celiac Disease/
47. (celiac disease or caeliac disease).ti,ab.
48. inflammatory bowel diseases/ or crohn disease/
49. (inflammatory bowel disease$ or crohn$ disease).ti,ab.
50. Craniopharyngioma/
51. (cranial tumo?r$ or craniopharyngioma).ti,ab.
52. Puberty, Precocious/
53. precocious sexual maturation.ti,ab.
54. Marfan Syndrome/
55. marfan$ syndrome.ti,ab.
56. Klinefelter Syndrome/
57. klinefelter$ syndrome.ti,ab.
58. psychosocial growth failure.ti,ab.
59. Noonan Syndrome/
60. noonan$ syndrome.ti,ab.
61. (psychosocial$ depriv$ or emotional$ depriv$).ti,ab.
62. russell-silver$ syndrome.ti,ab.
63. or/35-62
64. 34 and 63
65. exp child/
66. child$.ti,ab.
67. (school-age$ or schoolage$).ti,ab.
68. schoolchild$.ti,ab.
69. (boy or boys or girl or girls).ti,ab.
70. or/65-69
71. 64 and 70

MEDLINE In Process October 2005
Searched: 21 October 2005
Records found: 11

1. (sf36 or sf 36).ti,ab.

2. (eq5d or eq 5d or euroqol).ti,ab.
3. (short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab.
4. (hrql or hrqol or qol or hql or hqol).ti,ab.
5. (hye or hyes or health$ year$ equivalent$ or health utilit$).ti,ab.
6. rosser.ti,ab.
7. (person trade off$ or person tradeoff$ or standard gamble$ or time trade off or time tradeoff or tto or willingness to pay).ti,ab.
8. (disutilities or disutility or daly or disability adjusted life).ti,ab.
9. (qaly$ or qualy$ or quality adjusted life or quality of life or life quality).ti,ab.
10. qwb.ti,ab.
11. (quality of wellbeing or quality of well being or index of well being or index of wellbeing).ti,ab.
12. factor analysis.ti,ab.
13. preference based.ti,ab.
14. (health status or health state$).ti,ab.
15. (state adj2 (value or values or valuing or valued)).ti,ab.
16. hspv.ti,ab.
17. (utilit$ approach$ or health gain or hui or hui2 or hui 2 or hui3 or hui 3).ti,ab.
18. (categor$ scal$ or linear scal$ or linear analog$ scale$ or visual scal$ or magnitude estimat$).ti,ab.
19. (multiattribute$ health or multi attribute$ health).ti,ab.
20. health measurement$.ti,ab.
21. health survey questionnaire$.ti,ab.
22. (general health questionnaire$ or ghq).ti,ab.
23. (multiattribute$ theor$ or multi attribute$ theor$ or multiattribute$ analys$ or multi attribute$ analys$).ti,ab.
24. classification illness state$.ti,ab.
25. (health adj2 utilit$).ti,ab.
26. (multiattribute$ utilit$ or multi attribute$ utilit$).ti,ab.
27. willingness pay.ti,ab.
28. theory utilit$.ti,ab.
29. or/1-28
30. (growth hormone disease$ or growth hormone deficienc$).ti,ab.
31. growth hormone treatment$.ti,ab.
32. idiopathic short stature.ti,ab.
33. prader-willi syndrome.ti,ab.
34. chronic renal failure.ti,ab.
35. turner$ syndrome.ti,ab.
36. juvenile hypothyroidism.ti,ab.
37. (celiac disease or caeliac disease).ti,ab.
38. (inflammatory bowel disease$ or crohn$ disease).ti,ab.

39. (cranial tumo?r$ or craniopharyngioma).ti,ab.
40. precocious sexual maturation.ti,ab.
41. marfan$ syndrome.ti,ab.
42. klinefelter$ syndrome.ti,ab.
43. psychosocial growth failure.ti,ab.
44. noonan$ syndrome.ti,ab.
45. (psychosocial$ depriv$ or emotional$ depriv$).ti,ab.
46. russell-silver$ syndrome.ti,ab.
47. or/30-46
48. child$.ti,ab.
49. (school-age$ or schoolage$).ti,ab.
50. schoolchild$.ti,ab.
51. (boy or boys or girl or girls).ti,ab.
52. or/48-51
53. 29 and 47 and 52

EMBASE (Ovid) 1980–October 2005
Searched: 21 October 2005
Records found: 444

1. (sf36 or sf 36).ti,ab.
2. (eq5d or eq 5d or euroqol).ti,ab.
3. (short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab.
4. (hrql or hrqol or qol or hql or hqol).ti,ab.
5. (hye or hyes or health$ year$ equivalent$ or health utilit$).ti,ab.
6. rosser.ti,ab.
7. (person trade off$ or person tradeoff$ or standard gamble$ or time trade off or time tradeoff or tto).ti,ab.
8. (disutilities or disutility or daly or disability adjusted life).ti,ab.
9. (qaly$ or qualy$ or quality adjusted life or quality of life or life quality).ti,ab.
10. qwb.ti,ab.
11. (quality of wellbeing or quality of well being or index of well being or index of wellbeing).ti,ab.
12. factor analysis.ti,ab.
13. preference based.ti,ab.
14. (health status or health state$).ti,ab.
15. (state adj2 (value or values or valuing or valued)).ti,ab.
16. hspv.ti,ab.
17. quality adjusted life year/
18. "Quality of Life"/
19. Health Status/
20. health survey/
21. (hui or hui2 or hui 2 or hui3 or hui 3).ti,ab.
22. (utilit$ approach$ or health gain).ti,ab.
23. (categor$ scal$ or linear scal$ or linear analog$ scal$ or visual scal$ or magnitude estimat$).ti,ab.

24. (multiattribute$ health or multi attribute$ health).ti,ab.
25. health measurement$.ti,ab.
26. health survey questionnaire$.ti,ab.
27. (general health questionnaire$ or ghq).ti,ab.
28. (multiattribute$ theor$ or multi attribute$ theor$ or multiattribute$ analys$ or multi attribute$ analys$).ti,ab.
29. classification illness state$.ti,ab.
30. (health adj2 utilit$).ti,ab.
31. (multiattribute$ utilit$ or multi attribute$ utilit$).ti,ab.
32. willingness pay.ti,ab.
33. theory utilit$.ti,ab.
34. or/1-33
35. Growth Disorder/
36. (growth hormone disease$ or growth hormone deficienc$).ti,ab.
37. growth hormone treatment$.ti,ab.
38. idiopathic short stature.ti,ab.
39. Prader Willi Syndrome/
40. prader-willi syndrome.ti,ab.
41. Chronic Kidney Failure/
42. chronic renal failure.ti,ab.
43. Turner Syndrome/
44. turner$ syndrome.ti,ab.
45. juvenile hypothyroidism.ti,ab.
46. Celiac Disease/
47. (celiac disease or caeliac disease).ti,ab.
48. Crohn Disease/
49. (inflammatory bowel disease$ or crohn$ disease).ti,ab.
50. CRANIOPHARYNGIOMA/
51. (cranial tumo?r$ or craniopharyngioma).ti,ab.
52. Precocious Puberty/
53. precocious sexual maturation.ti,ab.
54. Marfan Syndrome/
55. marfan$ syndrome.ti,ab.
56. Klinefelter Syndrome/
57. klinefelter$ syndrome.ti,ab.
58. psychosocial growth failure.ti,ab.
59. Noonan Syndrome/
60. noonan$ syndrome.ti,ab.
61. (psychosocial$ depriv$ or emotional$ depriv$).ti,ab.
62. Silver Russell Syndrome/
63. russell-silver$ syndrome.ti,ab.
64. or/35-63
65. exp child/
66. child$.ti,ab.
67. (school-age$ or schoolage$).ti,ab.
68. (boy or boys or girl or girls).ti,ab.
69. schoolchild$.ti,ab.
70. or/65-69
71. 34 and 64 and 70

CINAHL (Ovid) 1982–October 2005
Searched: 21 October 2005
Records found: 64

1.   (sf36 or sf 36).ti,ab.
2.   (eq5d or eq 5d or euroqol).ti,ab.
3.   (short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform or thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab.
4.   (hrql or hrqol or qol or hql or hqol).ti,ab.
5.   (hye or hyes or health$ year$ equivalent$ or health utilit$).ti,ab.
6.   rosser.ti,ab.
7.   (person trade off$ or person tradeoff$ or standard gamble$ or time trade off or time tradeoff or tto).ti,ab.
8.   (disutilities or disutility or daly or disability adjusted life).ti,ab.
9.   (qaly$ or qualy$ or quality adjusted life or quality of life or life quality).ti,ab.
10. qwb.ti,ab.
11. (quality of wellbeing or quality of well being or index of well being or index of wellbeing).ti,ab.
12. factor analysis.ti,ab.
13. preference based.ti,ab.
14. (health status or health state$).ti,ab.
15. (state adj2 (value or values or valuing or valued)).ti,ab.
16. hspv.ti,ab.
17. exp "Quality of Life"/
18. exp Health Status/
19. Health Status Indicators/
20. Sickness Impact Profile/
21. (utilit$ approach$ or health gain or hui or hui2 or hui 2 or hui3 or hui 3).ti,ab.
22. (categor$ scal$ or linear scal$ or linear analog$ scal$ or visual scal$ or magnitude estimat$).ti,ab.
23. (multiattribute$ health or multi attribute$ health).ti,ab.
24. health measurement$.ti,ab.
25. health survey questionnaire$.ti,ab.
26. (general health questionnaire$ or ghq).ti,ab.
27. (multiattribute$ theor$ or multi attribute$ theor$ or multiattribute$ analys$ or multi attribute$ analys$).ti,ab.
28. classification illness state$.ti,ab.
29. (health adj2 utilit$).ti,ab.
30. (multiattribute$ utilit$ or multi attribute$ utilit$).ti,ab.
31. willingness pay.ti,ab.
32. theory utilit$.ti,ab.
33. or/1-32
34. Growth Disorders/
35. (growth hormone disease$ or growth hormone deficienc$).ti,ab.
36. growth hormone treatment$.ti,ab.
37. idiopathic short stature.ti,ab.
38. Prader-Willi Syndrome/
39. prader-willi syndrome.ti,ab.
40. Kidney Failure, Chronic/
41. chronic renal failure.ti,ab.
42. Turner-Syndrome/
43. turner$ syndrome.ti,ab.
44. juvenile hypothyroidism.ti,ab.
45. Celiac Disease/
46. (celiac disease or caeliac disease).ti,ab.
47. inflammatory bowel diseases/ or crohn disease/
48. (inflammatory bowel disease$ or crohn$ disease).ti,ab.
49. (cranial tumo?r$ or craniopharyngioma).ti,ab.
50. Puberty, Precocious/
51. precocious sexual maturation.ti,ab.
52. Marfan Syndrome/
53. marfan$ syndrome.ti,ab.
54. klinefelter$ syndrome.ti,ab.
55. psychosocial growth failure.ti,ab.
56. Noonan Syndrome/
57. noonan$ syndrome.ti,ab.
58. (psychosocial$ depriv$ or emotional$ depriv$).ti,ab.
59. russell-silver$ syndrome.ti,ab.
60. KLINEFELTER'S SYNDROME/
61. or/34-60
62. 33 and 61
63. exp child/
64. child$.ti,ab.
65. (school-age$ or schoolage$).ti,ab.
66. schoolchild$.ti,ab.
67. (boy or boys or girl or girls).ti,ab.
68. or/63-67
69. 62 and 68

Cochrane Library 2005 Issue 4
Searched: 21 October 2005
Records found: 6 reviews on CDSR; 42 RCTs CENTRAL

#1   sf36 or "sf 36" in All Fields in all products
#2   eq5d or "eq 5d" or euroqol in All Fields in all products
#3   "short form 36" or "shortform 36" or "sf thirtysix" or "sf thirty six" or "shortform thirtysix" or "shortform thirty six" or "short form thirtysix" or "short form thirty six" in All Fields in all products
#4   hrql or hrqol or qol or hql or hqol in All Fields in all products
#5   hye or hyes or "health* year* equivalent*" or "health utilit*" in All Fields in all products

#6 rosser in All Fields in all products

#7 "person trade off*" or "person tradeoff*" or "standard gamble*" or "time trade off" or "time tradeoff" or tto in All Fields in all products

#8 disutilities or disutility or daly or "disability adjusted life" in All Fields in all products

#9 qaly* or qualy* or "quality adjusted life" or "quality of life" or "life quality" in All Fields in all products

#10 qwb in All Fields in all products

#11 "quality of wellbeing" or "quality of well being" or "index of well being" or "index of wellbeing" in All Fields in all products

#12 "factor analysis" in All Fields in all products

#13 "preference based" in All Fields in all products

#14 "health status" or "health state*" in All Fields in all products

#15 state near (value or values or valuing or valued) in All Fields in all products

#16 hspv in All Fields in all products

#17 MeSH descriptor Quality-Adjusted Life Years explode all trees in MeSH products

#18 MeSH descriptor Quality of Life explode all trees in MeSH products

#19 MeSH descriptor Health Status explode all trees in MeSH products

#20 MeSH descriptor Health Status Indicators explode all trees in MeSH products

#21 MeSH descriptor Sickness Impact Profile explode all trees in MeSH products

#22 "utilit$ approach*" or "health gain" or hui or hui2 or "hui 2" or hui3 or "hui 3" in All Fields in all products

#23 "categor* scal*" or "linear scal*" or "linear analog* scal*" or "visual scal*" or "magnitude estimat*" in All Fields in all products

#24 "multiattribute* health" or "multi attribute* health" in All Fields in all products

#25 "health measurement*" in All Fields in all products

#26 "health survey questionnaire*" in All Fields in all products

#27 "general health questionnaire*" or ghq in All Fields in all products

#28 "multiattribute* theor*" or "multi attribute* theor*" or "multiattribute* analys*" or "multi attribute* analys*" in All Fields in all products

#29 "classification illness state*" in All Fields in all products

#30 health near utilit* in All Fields in all products

#31 "multiattribute* utilit*" or "multi attribute* utilit*" in All Fields in all products

#32 "willingness pay" in All Fields in all products

#33 "theory utilit*" in All Fields in all products

#34 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33)

#35 MeSH descriptor Growth Disorders explode all trees in MeSH products

#36 "growth hormone disease*" or "growth hormone deficienc*" in All Fields in all products

#37 "growth hormone treatment*" in All Fields in all products

#38 "idiopathic short stature" in All Fields in all products

#39 MeSH descriptor Prader-Willi Syndrome explode all trees in MeSH products

#40 "prader willi syndrome" in All Fields in all products

#41 MeSH descriptor Kidney Failure, Chronic explode all trees in MeSH products

#42 "chronic renal failure" in All Fields in all products

#43 MeSH descriptor Turner Syndrome explode all trees in MeSH products

#44 "turner* syndrome" in All Fields in all products

#45 "juvenile hypothyroidism" in All Fields in all products

#46 MeSH descriptor Celiac Disease explode all trees in MeSH products

#47 "celiac disease" or "caeliac disease" in All Fields in all products

#48 MeSH descriptor Inflammatory Bowel Diseases explode all trees in MeSH products

#49 MeSH descriptor Crohn Disease explode all trees in MeSH products

#50 "inflammatory bowel disease" or "crohn* disease" in All Fields in all products

#51 MeSH descriptor Craniopharyngioma explode all trees in MeSH products

#52 "cranial tumor*" or "cranial tumour*" or craniopharyngioma in All Fields in all products

#53 MeSH descriptor Puberty, Precocious explode all trees in MeSH products

#54 "precocious sexual maturation" in All Fields in all products

#55 MeSH descriptor Marfan Syndrome explode all trees in MeSH products

#56 "marfan* syndrome" in All Fields in all products

#57 MeSH descriptor Klinefelter Syndrome explode all trees in MeSH products

#58 "klinefelter* syndrome" in All Fields in all products
#59 "psychosocial growth failure" in All Fields in all products
#60 MeSH descriptor Noonan Syndrome explode all trees in MeSH products
#61 "noonan* syndrome" in All Fields in all products
#62 "psychosocial* depriv*" or "emotional* depriv*" in All Fields in all products
#63 "russell silver* syndrome" in All Fields in all products
#64 (#35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53 OR #54 OR #55 OR #56 OR #57 OR #58 OR #59 OR #60 OR #61 OR #62 OR #63) 4718
#65 (#34 AND #64)
#66 MeSH descriptor Child explode all trees in MeSH products
#67 child* in All Fields in all products
#68 "school age*" or schoolage* in All Fields in all products
#69 schoolchild* in All Fields in all products
#70 boy or boys or girl or girls in All Fields in all products
#71 (#66 OR #67 OR #68 OR #69 OR #70)
#72 (#65 AND #71)

NHS Economic Evaluation Database (public database) All years
Searched: 21 October 2005
Records found: 10

Health Technology Assessment (HTA) Database (public database) All years
Searched: 21 October 2005
Records found: 11

Database of Abstracts of Reviews of Effects (DARE) (public database) All years
Searched: 21 October 2005
Records found: 11

s growth(w)hormone(w)disease$ or growth(w)hormone(w)deficienc$
s growth(w)hormone(w)treatment$
s idiopathic(w)short(w)stature
s prader(w)willi(w)syndrome
s chronic(w)renal(w)failure
s turner$(w)syndrome
s juvenile(w)hypothyroidism
s celiac(w)disease or caeliac(w)disease
s inflammatory(w)bowel(w)disease$ or crohn$(w)disease

s cranial(w)tumor$ or cranial(w)tumour$ or craniopharyngioma
s precocious(w)sexual(w)maturation
s marfan$(w)syndrome
s klinefelter$(w)syndrome
s psychosocial(w)growth(w)failure
s noonan$(w)syndrome
s psychosocial$(w)depriv$ or emotional$(w)depriv$
s russell(w)silver$(w)syndrome
s s1 or s2 or s3 or s4 or s5 or s6 or s7 or s8 or s9 or s10 or s11 or s12 or s13 or s14 or s15 or s16 or s17
s child$
s school(w)age$ or schoolage$
s schoolchild$
s boy or boys or girl or girls
s s19 or s20 or s21 or s22
s s18 and s23

OHE Health Economic Evaluations Database September 2005
Searched: 21 October 2005
Records found: 4

'growth hormone disease' or 'growth hormone deficiency' or 'growth hormone treatment' or 'idiopathic short stature' or 'prader willi syndrome' or 'chronic renal failure' or 'turner syndrome' or 'juvenile hypothyroidism' or 'celiac disease' or 'caeliac disease' or 'inflammatory bowel disease' or 'crohns disease' or 'cranial tumors' or 'cranial tumours' or craniopharyngioma or 'precocious sexual maturation' or 'marfan syndrome' or 'klinefelter syndrome' or 'psychosocial growth failure' or 'noonan syndrome' or 'psychosocial deprivation' or 'emotional deprivation' or 'russell silver syndrome'

and

child or 'school age' or schoolage* or schoolchild* or boy or boys or girl or girls

Child growth – oxandrolone/ Turner's syndrome searches

The search strategy was designed to retrieve records on the use of oxandrolone in the treatment of TS.

Databases searched
MEDLINE
MEDLINE In Process
EMBASE
CINAHL

CENTRAL
Web of Knowledge Science Citation Index

Limits
1990 – 2005.
No study design limits.

Search strategies
MEDLINE (Ovid) 1990–October 2005
Searched: 24 October 2005
Records found: 66

1. OXANDROLONE/
2. (oxandrolone or oxandrin).ti,ab.
3. 1 or 2
4. Turner Syndrome/
5. turner$ syndrome.ti,ab.
6. 4 or 5
7. 3 and 6
8. limit 7 to yr="1990 - 2005"

MEDLINE In Process October 2005
Searched: 24 October 2005
Records found: 0

1. (oxandrolone or oxandrin).ti,ab.
2. turner$ syndrome.ti,ab.
3. 1 and 3
4. limit 3 to yr="1990 - 2005"

EMBASE (Ovid) 1990–October 2005
Searched: 24 October 2005
Records found: 136

1. OXANDROLONE/
2. (oxandrolone or oxandrin).ti,ab.
3. 1 or 2
4. Turner Syndrome/
5. turner$ syndrome.ti,ab.
6. 4 or 5

7. 3 and 6
8. limit 7 to yr="1990 - 2005"

CINAHL (Ovid)
1990–October 2005
Searched: 24 October 2005
Records found: 1

1. OXANDROLONE/
2. (oxandrolone or oxandrin).ti,ab.
3. 1 or 2
4. Turner Syndrome/
5. turner$ syndrome.ti,ab.
6. 4 or 5
7. 3 and 6
8. limit 7 to yr="1990 - 2005"

Cochrane Library – CENTRAL 2005 Issue 4
Searched: 24 October 2005
Records found: 18

#1 MeSH descriptor Oxandrolone explode all trees in MeSH products
#2 oxandrolone or oxandrin in All Fields in all products
#3 (#1 OR #2)
#4 MeSH descriptor Turner Syndrome explode all trees in MeSH products
#5 "turner* syndrome" in All Fields in all products
#6 (#4 OR #5)
#7 (#3 AND #6), from 1990 to 2005

Science Citation Index (Web of Knowledge)
1990–October 2005
Searched: 24 October 2005
Records found: 68

TS=(oxandrolone or oxandrin) and (turner* syndrome)

Appendix 4

Quality assessment tool for diagnostic yield studies

1. **Were the selection criteria clearly described?**
Yes Enough details are provided of how children were selected so that the selection process could be replicated. As a minimum, details of the school or community from which the children were selected should be given
No Insufficient details are presented

2. **Was an attempt described to contact all children who were eligible for measurement?**
Yes An explicit statement that all those eligible were invited to attend growth screening. Follow-up of those not attending is not necessary to answer *Yes*
No No details of participant contact are presented

3a **Were methods described to ensure the competence of those measuring?**
Yes Details of the competence of those measuring (e.g. trained auxologist) or details of training provided to ensure competence
No Insufficient details are presented

3b **Were methods described to check the competence of those measuring?**
Yes Details of methods to check the competence including, for example, comparisons with experts
No Insufficient details are presented

4 **Was there a defined, reproducible protocol to ensure accuracy and consistency of measurements?**
Yes Enough details are provided so that the measurement process could be replicated. Details are given of attempts to standardise good practice in measuring equipment, charts used and cut-offs
No Insufficient details are presented

5 **Was >80% of the sample actually measured?**
Yes Statement of the percentage measured or data available to calculate this
No No percentage stated and no means of calculation possible

6 **Were details of measurement error provided?**
Yes Details of the number incorrectly measured (either unexpected values or when re-measured by expert)
No Insufficient details provided

7 **Were all children measured accounted for in measurement results?**
Yes Details of all measurements and missing data are provided and numbers tally
No Insufficient details provided

8 **Were all children identified as needing follow-up accounted for in terms of diagnosis/false positive/lost to follow-up?**
Yes All children accounted for and numbers tally
No Some discrepancies in numbers followed up or insufficient details provided

9 **Were sufficient details provided concerning those diagnosed with a growth disorder?**
Yes Named condition and details of whether new or existing case or not known for all those diagnosed
No Insufficient details provided

Appendix 5

Modified QUADAS tool to assess the quality of diagnostic accuracy studies

1	**Were the selection criteria clearly described?**
Yes	Enough details are provided of how children were selected so that the selection process could be replicated. As a minimum, details of the school or community from which the children were selected should be given
No	Insufficient details are presented
2	**Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?**
Yes	If the whole sample or a random selection of the sample received the same reference standard
No	If only a selected sample received the reference standard
Unclear	If it is not clear whether all the children received the reference standard
3	**Did patients receive the same reference standard regardless of the index test result?**
Yes	If all patients received the same reference standard
No	If some patients received a different reference standard
Unclear	If it is not clear whether all patients received the same reference standard
4	**Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?**
Yes	If the index test and reference standard were independent
No	If the index test formed part of the reference standard
Unclear	If it is not clear if the index test and reference standard were independent
5a	**Was the execution of the index test described in sufficient detail to permit replication of the test?**
5b	**Was the execution of the reference standard described in sufficient detail to permit its replication?**
Yes	If sufficient details of test/reference standard execution are reported so that the test/reference standard could reasonably be replicated. This should include a description of the method used and by whom
No	If sufficient details are not reported
6a	**Were the index test results interpreted without knowledge of the results of the reference standard?**
6b	**Were the reference standard results interpreted without knowledge of the results of the index test?**
Yes	If the index test was interpreted without knowledge of the results of the reference standard and vice versa. If one test was clearly interpreted before the results of the other test were available then this should be scored as *Yes*
No	If the person interpreting the index test was aware of the results of the reference standard or vice versa
Unclear	If no information is provided regarding whether tests were interpreted independently
7	**Were uninterpretable/intermediate test results reported?**
Yes	If details are provided on uninterpretable/intermediate test results
No	If there appear to be some uninterpretable/intermediate but the results of these are not reported
Unclear	If it is not clear whether there were any uninterpretable/intermediate test results
8	**Were withdrawals from the study explained?**
Yes	If all children recruited into the study were accounted for
No	If there appear to be children who were recruited into the study who are not accounted for
Unclear	If it is not clear whether any withdrawals occurred
9a	**Was the cut-off threshold predefined for the index test?**
Yes	If the threshold was derived from the study population
No	If the threshold was derived from a separate population to that included in the study
Unclear	If the method of defining the threshold is not clearly defined
9b	**Was the cut-off threshold predefined for the reference standard?**
Yes	If the threshold was based on a value derived from the study population
No	If the threshold was derived from a separate population to that included in the study
Unclear	If the method of defining the threshold is not clearly defined

Appendix 6

Quality assessment tool used for RCTs

Study identification

Guideline topic:	**Key question no:**
Checklist completed by:	

SECTION 1: INTERNAL VALIDITY

In a well-conducted RCT study:	**In this study this criterion is:** *(Circle one option for each question)*	
1.1 The study addresses an appropriate and clearly focused question	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.2 The assignment of subjects to treatment groups is randomised	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.3 An adequate concealment method is used	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.4 Subjects and investigators are kept 'blind' about treatment allocation	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.5 The treatment and control groups are similar at the start of the trial	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.6 The only difference between groups is the treatment under investigation	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.7 All relevant outcomes are measured in a standard, valid and reliable way	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.8 What percentage of the individuals or clusters recruited into each treatment arm of the study dropped out before the study was completed?		
1.9 All the subjects are analysed in the groups to which they were randomly allocated (often referred to as intention-to-treat analysis)	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.10 Where the study is carried out at more than one site, results are comparable for all sites	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable

---

SECTION 2: OVERALL ASSESSMENT OF THE STUDY

2.1 How well was the study done to minimise bias? *Code* ++, + or −	
2.2 If coded as + or −, what is the likely direction in which bias might affect the study results?	
2.3 Taking into account clinical considerations, your evaluation of the methodology used and the statistical power of the study, are you certain that the overall effect is due to the study intervention?	
2.4 Are the results of this study directly applicable to the patient group targeted by this guideline?	

111

Appendix 7

Detailed description of identified growth monitoring programmes

TABLE 22 Detailed description of growth monitoring programmes

Study ID and programme delivery	Cost and referrals
Study ID: Agwu (2004)[76] **Method:** Each child was measured by a school nurse or school health nurse support worker using the Leicester height measure as part of a routine growth monitoring programme. Height data were entered on to a screening form. Those with heights below the 0.4th centile on the UK 1990 charts were referred to a growth clinic and were re-measured and underwent further investigation (MPH, bone age assessment, thyroid function tests, basic biochemistry, IGF-1, coeliac screen, chromosome screen, provocative tests for GH assessment) Those involved in measurement were issued with appropriate measuring equipment and were made aware of the screening guidelines **Training:** School health nurses and school health nurse support workers attended an update course on practical and theoretical aspects of growth measuring and monitoring	No relevant data reported
Study ID: Ahmed (1995)[56] **Method:** Height was measured by health visitor (HV) or GP using a Microtoise, Minimetre or Oxford screening wall charts. Height data were entered on to a database. Based on the Tanner and Whitehouse charts, any child with a height –2SDs or below the mean was invited to attend one of three community growth clinics (CGCs) staffed by an experienced auxologist, and any child with a height –3SDs of the mean was referred directly to the paediatric endocrinologist. At the CGC height was measured using a Magnimetre and monitored at 6-monthly intervals for 1–2 years and growth rate was determined. Bone assessment was undertaken at annual visits and a karyotype assessment was undertaken in girls. If growth rate and parental height suggested familial short stature or delay, children were discharged. Those with an annual height velocity less than the 25th centile were referred to the endocrinologist **Staff:** The CGCs were staffed by an auxologist and a project coordinator. Supervision was provided by a consultant paediatrician and a paediatric endocrinologist. The project coordinator input data into a computer to generate an SD for each child measured. This process took approximately 5 hours per week. The GP had to consent to growth referral and was asked to provide any relevant information **Training:** Letters were sent to GPs and HVs to advise on the importance of accurate height measurements and early identification of growth disorders. Initial meetings with HVs were held to discuss the measuring technique and the return of recording forms. All HVs were provided with Microtoises, height recording forms, measuring instructions and letter informing importance of growth assessment.	**Referrals:** Screening of children in the community clinics minimised the number of false positives and over-investigation and reduced travel and anxiety for families. 41/80 (51%) children referred to the auxologist were found to be growing normally **Cost:** The first 4 years of the study were funded to include the cost of the co-ordinator's salary (18.5 h/week), mileage, on-going computer costs, stationery and postage, telephone usage, Minimetres, data entry, patient travel claims, printouts and photocopying. The total cost was approx. £10,000/year. The use of a triage system was found to be cost-effective. The revised cost of running the surveillance to include an estimate of the auxologist's time. They found the costs for 180 children who were found to be growing slowly were £9630, including an initial consultant appointment. The cost of direct referral of 180 children to the paediatric endocrinologist at £90/visit with a minimum of three visits to determine growth rate gives a total of £48,600

TABLE 22 *Detailed description of growth monitoring programmes (cont'd)*

Study ID and programme delivery	Cost and referrals
Accuracy of measurement: A prior publication relating to the Oxford study[110] found that HVs were sufficiently accurate at measuring growth and that the Microtoise (or Minimetre) gave the most reproducible results. However, despite receiving training in using the Microtoises, some of the HVs used either the Oxford wall chart (used in 9.6% of the 3-year-olds and 4.3% of the 4.5-year-olds) or a tape measure (used to measure 23% of the 3-year-olds and 5% of the 4.5-year-olds), which may have introduced greater inaccuracy. On average, 93% of HVs returned forms, of which 84% contained usable measurements	
Study ID: Aszkenasy (2005)[71] **Method:** Each child was measured at school entry by a school nurse using a Leicester height measure in line with regional policy. Data were converted to a centile score for each child. Those with heights on or below the 0.4th centile based on the UK 1990 charts and who were not already known to the system were invited to attend a clinic for further investigation. Three audit cycles were performed in 1999–2000, 2000–1 and 2001–2. After each audit cycle the results were presented to school nurses **Staff:** A new form for school nurses was developed and introduced after the first audit cycle (1999–2000). The new form does not require a graphical plot to determine whether the height of a child was below the threshold for referral	**Referrals:** The author stated that two children in the cohorts were subsequently diagnosed with GHD, one did not attend for screening and the other child was on the 2nd centile at the time of screening and was therefore not referred for further investigation
Study ID: Banerjee (2003)[59] **Method:** An audit of height and weight data of all children born in 1992–3 and measured in 1998–9 from the relevant National Child Health System (NCHS) was undertaken. The NCHS was set up to identify children with heights or weights >98th centile or <2nd centile. Case notes of these children were reviewed and a manual system was used to identify each child with a height <0.4th centile based on the UK 1990 charts. Case notes of children of those initially referred were reviewed. Those who were not referred initially were recalled and invited to attend a clinic (either school medical clinic with a community paediatrician or an outpatient clinic with the consultant paediatrician in the local hospital/health centre) where they were re-measured and MPH and height velocity were assessed. Laboratory investigations were performed if needed **Staff:** The authors note that there was a shortage of school nurses in the region so other healthcare personnel were recruited to deliver routine school health surveillance. It was estimated that 130 sessions were needed to carry out measurements in 103 primary schools **Training:** No data on training were given in the report although the authors state that training and awareness of primary healthcare professionals were inadequate. Following the audit, training sessions on accurate measurement and interpretation of growth data were organised **Accuracy of measurement:** Height measurements were not transcribed on to centile charts in 75% of the case notes reviewed	**Referrals:** Of the 15 children identified with a height below the 0.4th centile, 10 were not previously referred **Costs:** Total cost of the programme was £14,550. The breakdown of these costs was: • personnel cost per session: £50 for health visitor and £25 for nursery nurse based on possible 10 sessions per week • training costs £1200 per academic session • administration costs £250 and miscellaneous costs £2000 Costs of further investigation were not included

TABLE 22 *Detailed description of growth monitoring programmes (cont'd)*

Study ID and programme delivery	Cost and referrals
Study ID: Cernerud (1994)[79] **Method:** Height and weight measurements were taken by school nurses in accordance with existing instructions. Those children with measurements outside predefined criteria for normal growth were followed up by the school physician, who then referred certain children with obvious deviations for specialist investigation. Attitudes of health professionals to growth screening were assessed by open interviews using an expert panel consisting of experienced school nurses and senior paediatricians **Staff:** The mean time required for measurements and associated procedures (preparation, plotting growth records, later examination of absent children and follow-up consultations by school doctor) was 5 minutes per child. The time used for measurements was less than 2% of the total time available for school health work by school nurses. The time used by the school doctor for follow-up examinations was negligible **Accuracy of measurement** The equipment used in the schools was not subject to systematic quality control, although the members of the panel recommended a more systematic quality control and feedback of the measurement technique, the choice of instruments and an annual calibration of the instruments	**Costs:** The costs of equipment were very low (less than US$15 per instrument per year). The personnel costs for the measurements represented a very small part of the salary for a nurse
Study ID: de la Puente (1999)[81] **Method:** Primary care teams linked to three hospitals in the province of Barcelona were invited to participate in the study. The eight health centres which elected to participate had to screen all children born between 1986 and 1987 under their jurisdiction. All children had previously attended the health centres. Most children were aged 6 or 7 years, although a small number were aged 5 or 8 years at the time of measurement. Both height and weight measurements were taken and a demographic questionnaire was given to the person accompanying the child. Any child with a height less than or equal to the 3rd centile based on the growth charts for Catalonia was identified and invited to attend an appointment with one of three paediatric endocrinologists. The protocol included assessment of parental height, growth velocity and bone age. If appropriate, GH tests, skeletal assessments and karyotyping were also conducted. The maximum follow-up was 1 year **Accuracy of measurement:** One person, who had previously trained, acted as a reference for the rest of the team. Team members assessed the height of small samples of children and were only permitted to participate in the growth monitoring programme when their results did not differ significantly from the reference	**Referrals:** The authors suggested that to improve the efficiency of growth monitoring in primary care, only those under the 0.4 percentile should be referred to the specialist. The paediatrics team should deal with non-organic growth disorders. There was a suggestion that school-based screening might only be necessary for those who do not regularly attend health centres
Study ID: Hearn (1995)[82] **Method:** Height and weight measurements were taken by school nurses at school entry (ages 5 and 11 years) using a Minimeter as part of a school medical programme. Data were inputted into a computer database and any child with a height less than the 3rd centile on the Tanner and Whitehouse charts was identified and invited to attend one of three new community growth clinics (CGCs). The clinics were held weekly in existing local health centres and were staffed by a community paediatrician or paediatric endocrinologist, a research sister and an auxologist. Children were measured and a detailed medical history and examination and pubertal assessment was undertaken. Blood, creatinine and liver function tests were also performed. Those with a confirmed height below the	**Referrals:** New procedures for referral of short children and audit were introduced The authors state that the upward secular trend in the socio-economic deprived population included in the study and the use of the Tanner and Whitehouse centile as a threshold for referral may have contributed to a low referral rate and the subsequent low yield of new cases of children with GHD

continued

115

TABLE 22 *Detailed description of growth monitoring programmes (cont'd)*

Study ID and programme delivery	Cost and referrals
3rd centile were offered repeat appointments at 4, 8 and 12 months to determine height velocity. Those with initial HSDS less than –3SDs were immediately referred to a paediatric endocrinologist at the hospital growth clinic **Staff:** Monthly support visits were given to each nurse to set agreed targets and evaluate performance. Referral appointments and clinic organisation were coordinated by the research sister **Training:** School nurses were given training by a clinical auxologist. Teaching seminars were organised covering growth, aims and rationale and key role of school nurses in height surveillance. Training sessions were given on plotting height data and use of growth standards. Reference manuals were provided giving guidelines on measuring, recording of data and referral procedure. Participating schools and health centres were visited to discuss any practical difficulties. The authors state that prior to the study school nurses were not equipped with appropriate equipment or training in growth assessment **Accuracy of measurement:** In a related publication concerning the Hackney study, the accuracy and reproducibility of the nurses' measurements were found to be satisfactory.[111] Minimeters were checked by the auxologist, and those found to be inaccurate by more than 0.3 mm were repositioned and recalibrated or replaced	The authors stated that the use of community-based growth clinics was useful as it allowed for initial assessment of short children prior to referral, leading to a higher identification of cases of abnormal growth in those attending referral clinics
Study ID: Keller (2002)[83] **Method:** This study presented an auxological computer-based network (CrescNet) developed through the collaboration of practising paediatricians and paediatric endocrinologists to detect growth disorders. Children and adolescents who were included in the network had attended children's health clinics for routine check-ups. Each child was measured on at least one occasion. Data were entered into a computerised database and those above the 97th or below the 3rd percentile were identified **Staff:** Height and weight data obtained from each child were entered on to a tear-off bar code ticket and sent to a central database centre at the University of Leipzig. An experienced auxologist or endocrinologist would review the data and forward an individual case recommendation to the paediatrician Annual meetings involving all participants of the network were held. Research meetings were held every 4 weeks and quality assessment was routinely performed by email or telephone interview. **Training:** Staff who were responsible for measuring height and weight in the practices were trained at least once by experienced staff members from the growth clinic **Accuracy of measurement:** No specific details were reported, although regular meetings (detailed above) were used for the purposes of quality assessment. Plausibility of the data was checked using computerised thresholds and comparison with reference data	**Referrals:** Data present in an earlier publication by Keller and colleagues[85] containing 1999 data found that 132/30,182 entries were mistakes, 916/30,182 were classified as above the 97th centile and 1140/30,182 were below the 3rd centile

TABLE 22 *Detailed description of growth monitoring programmes (cont'd)*

Study ID and programme delivery	Cost and referrals
Study ID: Lacey (1974)[86] **Method:** Children who were measured as part of the Newcastle Survey of Child Development (NSCD) were included. Children born during 1960–2 to mothers living in Newcastle who attended local schools were measured, usually by a school nurse, at school entry and at the age of 10 years Children were selected for inclusion based on recorded heights at the age of 10 years. Two groups were selected: children born in 1960 with a recorded height below the 3rd centile based on the Tanner and Whitehouse charts at age 10 years, and children born in 1961 or 1962 with a recorded height more than 3SDs below the mean (complete data were only available for 1960 cohort; some data on the number of cases found in the 1961, 1962 cohorts are presented, although it is not clear how many children were included in the study) Each child was visited and consent was sought to undergo further clinical investigation, including additional height measurements **Accuracy of measurement:** In the 1960 cohort, six children whose reported heights placed them below the 3rd centile at 10 years were found to be above the 3rd centile at the clinical assessment and at school entry and were excluded based on the assumption that their heights were erroneous	
Study ID: Lindsay (1994)[87] **Method:** Measurements of elementary school children were taken by trained volunteers using an Accustat stadiometer over two consecutive years to determine height and height velocity. In the first year of the study, children with a height <-2SDs (in the first year of the study) or those with a height $<$3rd centile and growth rate <5 cm/year (in the second year of the study) were examined by the study physicians or family physicians to determine the cause of short stature. Follow-up investigations included review of earlier growth, physical examination, bone age assessment and X-ray. Additional assessments were recommended to ascertain cause of short stature, where necessary **Staff:** Each child could be measured in less than 30 seconds when the process was done properly, therefore a full classroom took less than 20 minutes. The project coordinator supervised 27% of schools to ensure accuracy and reproducibility. Family physicians obtained advice from study physicians as needed **Training:** The project coordinator met with each parent–teacher group to provide training workshops. These involved practice measurements and presentation of information describing abnormal and normal growth and the importance of detecting abnormal growth early **Accuracy of measurement:** Pilot surveys were performed in a sample of schools to assess the efficiency of the procedure. This involved two measurements one month apart to assess reproducibility and accuracy. Unacceptable measurements resulted in extensive training by demonstration and use of stadiometer. This was found to improve accuracy of measurement	**Referrals:** Referral to an endocrinologist had previously been suggested for 50% of children with GHD and 17% of children with TS, and only 25% of the first and none of the latter had actually been seen by an endocrinologist prior to the study

117

TABLE 22 *Detailed description of growth monitoring programmes (cont'd)*

Study ID and programme delivery	Cost and referrals
Study ID: Vimpani (1981)[91] **Method:** Each child was measured in the classroom using a portable anthropometer and the smallest children were re-measured using a wall-mounted stadiometer in a purpose-built mobile clinic by an investigator. Any child who was absent on the day of measurement and who was considered short by the class teachers was measured on a later visit. Children with a height <2.5SDs based on the Tanner and Whitehouse charts who had not undergone any previous investigations or did not have any apparent organic basis for short stature underwent a series of diagnostic tests to ascertain the cause of short stature **Accuracy of measurement:** The accuracy of the stadiometer in the mobile clinic was assessed at each school. Measurements were considered reliable when 28 children attending a single school were measured on successive days with 95% of repeat heights being within 4.5 mm	
Study ID: Voss (1992)[93] **Method:** Each child was measured by a school nurse at school entry medical using a portable wall-mounted Microtoise. The heights of those near or below the 3rd centile based on the Tanner and Whitehouse charts were verified by a trained auxologist using a stadiometer. Parents of short children identified were requested to take their child to the nearest clinic for blood tests, bone age assessment and measurement of urinary GH. If the test results were abnormal, children were examined by an paediatric specialist to ascertain the cause of short stature **Training:** No details were given on the training of school nurses involved in the initial measurement of child height. The auxologist involved in the verification of heights and subsequent measurements had received training in anthropometric techniques **Accuracy of measurement:** The Microtoise used by school nurses was checked for accuracy using a meter rod, and the stadiometer used by the trained auxologist was regularly checked for accuracy. Several studies were conducted prior to the monitoring programme to determine the reliability of height measurement	**Referrals:** The authors state that the school nurses were encouraged to err on the side of caution and refer all children for re-measurement that were near the 3rd centile The authors suggest that the unexpectedly low percentage of children with heights below the 3rd centile (1.3% of school entry population compared with an expected 3%) may be due to the Tanner and Whitehouse standards being out of date given the upward secular trend
Study ID: White (1995)[96] **Method:** Children aged 3, 5 and 14 years were measured as part of a statutory medical screening by their GP, clinical medical officer, health visitor or school nurse. Children aged 7, 9 and 11 years were measured separately for other health surveillance reasons by the school nurse or clinical medical officer. Measurements were made using a portable stadiometer (Microtoise, Raven Minimeter). Height and weight data were entered into a computer package that identified children outwith the 3rd and 97th centiles according to the Tanner and Whitehouse charts. No assessment was undertaken on the causes of short stature **Staff:** Dovetailing new measurements with existing health services minimised the workload of primary healthcare staff. More than 100 community staff took part in the study, but approximately 75% of children were measured by 26 school nurses	

TABLE 22 *Detailed description of growth monitoring programmes (cont'd)*

Study ID and programme delivery	Cost and referrals
Training: In-service training on an ongoing basis was provided for the community staff by instruction in a standard method. This was provided for the benefit of new staff and to maintain appropriate standards in existing staff **Accuracy of measurement:** The accuracy of measurement was validated by comparing the community measurements with a single auxologist's. No significant bias was found between measurements taken by primary health workers and auxologists. The SD of the difference was 0.5 cm, meaning that 95% of measurements made by the observers in the community fell within 1 cm of the auxologist's measurement. Measuring equipment and techniques were standardised for the region. Schools and clinics installed portable stadiometers	

Appendix 8

Detection of growth conditions

TABLE 23 *Conditions identified by the screening programmes*

Study	Conditions identified by the screening programme (new cases, existing cases or unclear)
Agwu (2004)[76]	Autoimmune hypothyroidism (1 new), idiopathic GHD (2 new), intrauterine growth retardation with major psychosocial problems (1 new), intrauterine growth retardation (1 existing), Batter's syndrome (1 existing), cystic fibrosis (1 existing), familial short stature (5 new, 2 existing)
Ahmed (1995)[56]	Short normals (69 new, 14 existing), congenital abnormalities/syndromes (4 new, 44 existing), intrauterine growth retardation (3 new, 8 existing), bony dysplasia (1 new, 2 existing), GHD (1 new, 1 existing), emotional deprivation (1 new), coeliac disease (1 new)
Aszkenasy (2005)[71]	William's syndrome (1 existing), myoclonic epilepsy and immune deficiency (1 existing), GHD (1 existing), chronic lung disease (1 existing), severe eczema (1 existing), non-organic failure to thrive (3 existing), osteogenesis imperfecta (1 existing), muscular dystrophy (1 existing), Down's syndrome (1 existing), familial short stature (9 new cases, 1 existing)
Banerjee (2003)[59]	Neurofibromatosis (1 existing), arthrogryphosis (1 existing)
Cernerud (1994)[79]	None
de la Puente (1999)[81]	GHD (1 existing), intrauterine growth retardation (2 new, 4 existing), TS (1 existing), neonatal anoxia (2 existing), achondroplasia (1 existing), familial short stature (19 new, six existing, two unclear), constitutional growth retardation (three new, four existing), combined growth retardation (2 new, 1 existing, 1 unclear)
Hearn (1995)[82]	Genetic short stature (30 new), ISS (12 new), intrauterine growth retardation (5 new), constitutional growth delay (6 new), GHD (2 new), congenital hypothyroidism (1 new), psychosocial deprivation (2 new), chronic liver disease (1 new)
Keller (2002)[83]	*Tall stature:* disorders of puberty (10 new), adiposogigantism (3 new), Marfan's syndrome (1 new), suspected acromegaly (1 new), polycystic ovary syndrome (1 new), Klinefelter's syndrome (1 new) *Short stature:* achondroplasia (3 new), hypochondroplasia (2 new), skeletal dysplasia (1 new), small for gestational age (14 new), disorders of puberty (8 new), gastroenterological disorders (5 new), bronchial asthma (4 new), psychosocial (3 new), hypothyroidism (2 new), neurodermatitis (2 new), myelomeningocele (1 new), Duchenne muscular dystrophy (1 new), arthritic (1 new), anaemia (1 new), nephrogenous (1 new), Ullrich–Turner syndrome (4 new), PWS (1 new), Silver–Russell syndrome (1 new), trichorhinopharyngeal syndrome (1 new), William–Beuren syndrome (1 new), total GHD (3 new), partial GHD (23 new), neurosecretory dysfunction (7 new), suspected GHD (5 new)
Lacey (1974)[86]	Down's syndrome (5 existing), cystic fibrosis (1 existing), chromosome abnormality (1 new), chronic renal disease (1 new), GHD (1 new), mental subnormality (4 existing), Fallot's tetralogy (1 existing), Still's disease (1 existing), Hurler's syndrome (1 existing), reversibly low GHD due to emotional deprivation (1 new, described in text)
Lindsay (1994)[87]	Familial short stature (207 unclear), constitutional growth delay (149 unclear), familial short stature/constitutional growth delay (94 unclear), other medical causes (53 unclear), ISS (27 unclear, three of which may be new cases of neurosecretory GHD), GHD (16 new), TS (6 new), hypothyroidism (3 new)

TABLE 23 *Conditions identified by the screening programmes (cont'd)*

Study	Conditions identified by the screening programme (new cases, existing cases or unclear)
Vimpani (1981)[91]	Spina bifida (14 unclear), cerebral palsy (13 unclear), microcephaly (3 unclear), arrested hydrocephalus (2 existing), epilepsy and severe mental retardation (2 unclear), hemiplegia (1 existing), neurofibromatosis (1 existing), meningitis and nerve deafness (1 existing), non-specific mental retardation (2 unclear), scoliosis (1 existing, 4 unclear), achondroplasia (3 unclear), achondroplasis and spina bifida occulta (1 unclear), hypochondroplasia (1 unclear), sacral tumour (2 unclear), Perthes' disease (2 existing), Klippel–Feil abnormality (1 unclear), metaphyseal dystosis (1 existing), rickets (1 existing), Down's syndrome (15 unclear), trisomy 18 (1 unclear), Turner's mosaic (1 new), asthma (2 existing, 3 unclear), chronic bronchitis (2 existing), cystic fibrosis (1 unclear), pulmonary stenosis (2 existing, 1 unclear), ventricular septal defect (1 existing), patent ductus arteriosus (1 unclear), intersex and chronic heart disease (1 unclear), coeliac disease (1 existing, 2 unclear), Hirchsprung's disease (1 unclear), imperforate anus and mental retardation (1 unclear), tracheo-oesophageal fistula (1 existing), glycogen storage disease (1 existing), Hurler's syndrome (1 unclear), Morquio's syndrome (1 unclear), non-specific storage disorder (1 unclear), acute lymphoblastic leukaemia (1 existing), chronic idiopathic thrombocytopenic purpura (1 unclear), renal hypoplasia (1 existing), severe recurrent urinary tract infections (1 existing), hypothyroidism (1 new), gonadal hypoplasia (1 unclear), Russell–Silver syndrome (1 existing, 1 unclear), ataxia telangiectasia (1 unclear), progeria (1 unclear), smaller of discordant twins (1 existing), severe GHD (9 new, 4 existing), partial GHD (25 unclear), low birthweight short stature (34 unclear), constitutional short stature (178 unclear)
Voss (1992)[93]	Down's syndrome (4 existing), mental retardation (4 existing), gross physical deformities (3 existing), phocomelia (2 existing), cerebral palsy (2 existing), coeliac disease (1 new, 1 existing), cystic fibrosis (1 existing), hypochondroplasia (1 existing), phenytoin toxicity syndrome (1 existing), Rubinstein–Taybi syndrome (1 existing), Russell–Silver syndrome (1 existing), spina bifida (1 existing), TS (1 existing), Vater association (1 existing), Down's syndrome with GHD (1 existing), Noonan's syndrome (1 new), hypothyroidism (1 new), lead poisoning (1 new), neurofibromatosis (1 new), neurofibromatosis with brachydactyly (1 new), trichorhinophalangeal syndrome (1 new), GHD (with known mental retardation) (1 new), described in text as asthmatic (23 unclear), described in text as having psychosocial deprivation (32 unclear)

TABLE 24 *Detection rate of growth-related conditions*

	Agwu (2004)[76]	Ahmed (1995)[56]	Aszkenasy (2005)[71]	Banerjee (2003)[59]	Cernerud (1994)[79]	de la Puente (1999)[81]	Hearn (1995)[82]	Keller (2002)[83]	Lacey (1974)[86]	Lindsay (1994)[87]	Vimpani (1981)[91]	Voss (1992)[93]
Number of participants	3,465	20,338	9,338	1,592	7,129	2,084	9,549	6,0984	2,256[a]	114,881	48,221	14,346
Number meeting referral criteria and were referred[b]	18	260	24[c]	15	20	73	146[d]	SS = 346 TS = 130	111	630	449	180[e]
Number who were followed up/data available	14	149	9	9	20	49	59	SS = 346 TS = 130	98	555	358	Unclear
Total number diagnosed prior to programme/already seeing a paediatrician	5	69	12	2	None	20[f]	17 out of first 100 referred	N/A[g]	13	Unclear	Unclear	25
Number meeting referral criteria who were lost to follow-up or no data were available	4	76	15	6	None	20	At least 2 out of first 100 referred	N/A[g]	13	75	91	Unclear
Error in initial measurement	Unclear	35	10[h]	2[i]	Unclear	4	25 out of first 100 referred	Not reported	Not reported	Not reported	Unclear	51
Specified conditions												
GHD	2N 1:1,733	IN (IE) 1:20,338	IE			IE	2N 1:4,775	38N 1:1,605	IN 1:2,256	16N 1:7,180	9N (4E 25U) 1:5,358	IN (IE) 1:14,346
TS						IE		4N 1:15,246		6N 1:19,147	IN 1:48,221	IE
JH	IN 1:3,465		3E					2N 1:30,492		3N 1:38,293	IN 1:48,221	IN 1:14,346
Psychosocial growth failure	IN 1:3,465	IN 1:20,338					2N 1:4,775	3N 1:20,328	IN 1:2,256			32U[j]
Conditions of tall stature	N/A	N/A	N/A	N/A		N/A	N/A	6N 1:10,164	N/A	N/A	N/A	N/A

continued

TABLE 24 *Detection rate of growth-related conditions (cont'd)*

	Agwu (2004)[76]	Ahmed (1995)[56]	Aszkenasy (2005)[71]	Banerjee (2003)[59]	Cernerud (1994)[79]	de la Puente (1999)[81]	Hearn (1995)[82]	Keller (2002)[83]	Lacey (1974)[86]	Lindsay (1994)[87]	Vimpani (1981)[91]	Voss (1992)[93]
Other conditions	3E	9N 54E	7E	2E	0N	2N 7E	7N	59N[k]	2N 13E	53U	25E 115U	6N 23E 23U[i]
Total number of cases of all conditions	4N (3E) *1:866*	1N (54E) *1:1,849*	0N (11E)	0N (2E)	0N	2N (9E) *1:1,042*	11N *1:868*	112N *1:545*	4N (13E) *1:564*	25N (53U) *1:4595*	1N (29E 140U) *1:4,384*	8N (25E, 23U) *1:1,793*
Short normal (ISS/familial short stature/constitutional delay)	2E 5N	14E 69N	1E 9N			11E 24N 3U	48N	251N	81U	477U	178U	140U[j]

E, existing case; N, new case; U, unclear if new or existing; diagnostic yield (in italics) was calculated for total number of new cases.
a Based on data from the 1960 cohort only as insufficient details on the 1961 and 1962 cohorts were given in the report.
b Based on the assumption that all children meeting the referral criteria were referred, unless stated otherwise.
c 36 children were found to be below the 0.4th centile but 12 children who were already known to paediatricians were not referred.
d The authors state that 59 children have been fully assessed at the growth clinic, but it is not clear if this is based on all 146 children referred for further assessment.
e The authors state that 239/14,346 were found to be on or below the 3rd centile. Of these, eight were excluded as they were not in the age range of interest and 51 were excluded after re-measurement found them to be above the 3rd centile.[41]
f A further three cases were unclear.
g Study was based on the development of a computer-based network for the detection of growth disorders.
h The computer analysis identified 36 children with heights less than the 0.4th centile of which the school nurse identified 23.
i Two of the children who were not initially referred were found to have a height greater than the 2nd centile when measured by a paediatrician.
j Among the short normal children, 32 were judged to have psychosocial deprivation and 23 were asthmatic.
k Figure includes 11 cases associated with tall stature.

Appendix 9

Studies of diagnostic accuracy of obesity

TABLE 25 Studies of diagnostic accuracy of obesity

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Bedogni (2003)[97] **Index tests:** 1. BMI. Centiles were internally derived 2. 4SF = sum of four skinfold measurements (triceps, biceps, subscapular and suprailic). Centiles were internally derived **Reference standard:** 1. Fat mass (FM): weight (wt) DEXA was used to develop a population-specific algorithm to obtain fat-free mass (FFM) from BIA. FM was obtained by subtracting FFM from wt. Centiles were internally derived	BMI >85th centile	FM: Wt >85th centile	Whole sample	91 49	40 754	0.65	0.57 to 0.72	0.95	0.93 to 0.96
	BMI >95th centile	FM: Wt >85th centile	Whole sample	55 85	8 786	0.39	0.32 to 0.48	0.99	0.98 to 0.99
	4SF >85th centile	FM: Wt >85th centile	Whole sample	105 35	48 746	0.75	0.67 to 0.81	0.94	0.92 to 0.95
Ebbeling (1999)[98] **Index tests:** 1. BMI. 85th centile from NHANES I 2. TSF. 85th centile from NHANES I 3. Weight for height (WH). 120% of median from NCHS growth curves 4. WH and TSF. Thresholds as above **Reference standard:** 1. % BF Estimated using Williams equations incorporating age, TSF and subscapular skinfold. The obesity cut off point used for boys was 25% BF and that for girls was 30% BF	BMI >85th centile	%BF>25%	Boys	43 8	64 470	0.84	0.72 to 0.92	0.88	0.85 to 0.91
		%BF>30%	Girls	43 0	76 467	1.00	0.92 to 1.00	0.86	0.83 to 0.89
	TSF >85th centile	%BF>25%	Boys	48 3	53 481	0.94	0.84 to 0.98	0.90	0.87 to 0.92
		%BF>30%	Girls	42 1	65 478	0.98	0.88 to 1.00	0.88	0.85 to 0.90
	WH >120% median	%BF>25%	Boys	38 10	16 512	0.79	0.66 to 0.88	0.97	0.95 to 0.98
		%BF>30%	Girls	26 2	24 465	0.93	0.77 to 0.98	0.95	0.93 to 0.97
	WH >120% median and TSF >85th centile	%BF>25%	Boys	35 13	11 517	0.73	0.59 to 0.83	0.98	0.96 to 0.99
		%BF>30%	Girls	26 2	10 479	0.93	0.77 to 0.98	0.98	0.96 to 0.99

TABLE 25 *Studies of diagnostic accuracy of obesity (cont'd)*

Study details	Index test	Reference standard	Sample	TP / FN	FP / TN	Sensitivity	95% CI	Specificity	95% CI
Ellis (1999)[99] **Index test:** 1. BMI. Centiles were internally derived **Reference standard:** 1. %BF DEXA was used to estimate %BF. Thresholds for 'excess BF' (≥85th centile) and 'obese' (≥95th centile) were used. Centiles were internally derived	BMI ≥85th centile	%BF ≥85th centile	Boys	55 / 6	57 / 288	0.90	0.80 to 0.95	0.83	0.79 to 0.87
			Girls	81 / 5	85 / 402	0.94	0.87 to 0.97	0.83	0.79 to 0.86
	BMI ≥95th centile	%BF ≥95th centile	Boys	15 / 6	35 / 350	0.71	0.50 to 0.86	0.91	0.88 to 0.93
			Girls	26 / 5	42 /	0.90	0.74 to 0.96	0.92	0.90 to 0.94
Himes (1989)[100] **Index tests:** 1. BMI 2. %BF estimated from 4 skinfold measurements 3. Subscapular skinfold (SSF) 4. TSF 5. Weight To define the obesity cut-off for each measure, 85th centiles for age and gender were derived from reference data	BMI ≥85th centile	%BF ≥90th centile	Boys	6 / 15	1 / 137	0.29	0.14 to 0.50	0.99	0.96 to 1.00
			Girls	7 / 23	0 / 127	0.23	0.12 to 0.41	1.00	0.97 to 1.00
	%BF from 4SF ≥85th centile	%BF ≥90th centile	Boys	12 / 9	21 / 117	0.57	0.37 to 0.76	0.85	0.78 to 0.90
			Girls	24 / 6	23 / 104	0.80	0.63 to 0.90	0.82	0.74 to 0.88
	SSF ≥85th centile	%BF ≥90th centile	Boys	8 / 13	1 / 137	0.38	0.21 to 0.59	0.99	0.96 to 1.00
			Girls	9 / 21	1 / 126	0.30	0.17 to 0.48	0.99	0.96 to 1.00
	TSF ≥85th centile	%BF ≥90th centile	Boys	5 / 16	0 / 138	0.24	0.11 to 0.45	1.00	0.97 to 1.00
			Girls	7 / 23	4 / 123	0.23	0.12 to 0.41	0.97	0.92 to 0.99
	Weight ≥85th centile	%BF ≥90th centile	Boys	9 / 12	7 / 131	0.43	0.24 to 0.63	0.95	0.90 to 0.98
			Girls	5 / 25	3 / 124	0.17	0.07 to 0.34	0.98	0.93 to 0.99

Reference standard:
1. %BF
%BF was measured using densitometry 90th centiles for age and gender, defining the obesity cut-off, were derived from reference data for Cincinnati youths

127

128

TABLE 25 Studies of diagnostic accuracy of obesity (cont'd)

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Lazarus (1996)[101] **Index test:** 1. BMI. Centiles were internally derived **Reference standard:** 1. %BF %BF was estimated using DEXA and adjusted for age and gender. 85th centiles, representing excess adiposity, were internally derived	BMI ≥85th centile	%BF ≥85th centile	All	24 10	10 186	0.71	0.54 to 0.83	0.95	0.91 to 0.97
			Boys	12 6	6 95	0.67	0.44 to 0.84	0.94	0.88 to 0.97
			Girls	12 4	4 91	0.75	0.51 to 0.90	0.96	0.90 to 0.98
	BMI ≥90th centile	%BF ≥85th centile	All	21 13	2 194	0.62	0.45 to 0.76	0.99	0.96 to 1.00
			Boys	11 7	1 100	0.61	0.39 to 0.80	0.99	0.95 to 1.00
			Girls	10 6	1 94	0.63	0.39 to 0.82	0.99	0.94 to 1.00
	BMI ≥95th centile	%BF ≥85th centile	All	10 24	2 194	0.29	0.17 to 0.46	0.99	0.96 to 1.00
			Boys	5 13	1 100	0.28	0.12 to 0.51	0.99	0.95 to 1.00
			Girls	5 11	0 95	0.31	0.14 to 0.56	1.00	0.96 to 1.00

TABLE 25 *Studies of diagnostic accuracy of obesity (cont'd)*

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Marshall (1991)[102]									
Index tests:									
1. CSTF = sum of five skinfolds according to the Canadian Standardized Test of Fitness (1986). 85th centiles were from reference data	CSTF ≥85th centile	%BF ≥20%	Boys	36 9	23 198	0.80	0.66 to 0.89	0.90	0.85 to 0.93
		%BF ≥25%	Girls	30 1	23 220	0.97	0.84 to 0.99	0.91	0.86 to 0.94
2. Relative BMI. Calculated using reference data	Relative BMI ≥120%	%BF ≥20%	Boys	31 14	16 205	0.69	0.54 to 0.80	0.93	0.89 to 0.95
3. Relative weight. Calculated using reference data		%BF ≥25%	Girls	23 8	21 222	0.74	0.57 to 0.86	0.91	0.87 to 0.94
4. TSF. 85th centiles were internally derived	Relative weight ≥120%	%BF ≥20%	Boys	22 23	11 210	0.49	0.35 to 0.63	0.95	0.91 to 0.97
Reference standard:		%BF ≥25%	Girls	18 13	12 231	0.58	0.41 to 0.74	0.95	0.92 to 0.97
1. % BF Body density was measured through hydrostatic weighing. %BF was calculated from this using age- and gender-specific equations. Obesity cut-offs were 20% for boys and 25% for girls	TSF ≥85th centile	%BF ≥20%	Boys	29 16	12 209	0.64	0.50 to 0.77	0.95	0.91 to 0.97
		%BF ≥25%	Girls	21 10	16 227	0.68	0.50 to 0.81	0.93	0.90 to 0.96

129

TABLE 25 *Studies of diagnostic accuracy of obesity (cont'd)*

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Mast (2002)[103] **Index test:** 1. BMI. Centiles are from German reference data	BMI ≥90th centile	%BF(BIA) ≥90th centile	Boys	86 19	62 979	0.82	0.73 to 0.88	0.94	0.92 to 0.95
			Girls	84ᵃ 24	52 980	0.78	0.69 to 0.85	0.95	0.94 to 0.96
Reference standards: 1. %BF as estimated using BIA 2. %BF as estimated using skinfold	BMI ≥90th centile	%BF (skinfolds) ≥90th centile	Boys	86 21	62 977	0.80	0.72 to 0.87	0.94	0.92 to 0.95
measurements 3. TSF			Girls	79 40	61 960	0.66	0.58 to 0.74	0.94	0.92 to 0.95
For all three measures 90th centiles (overweight cut-off) and 97th centiles (obese cut-off) were derived internally	BMI ≥90th centile	TSF ≥90th centile	Boys	80 32	72 962	0.71	0.62 to 0.79	0.93	0.91 to 0.94
			Girls	67 45	72 956	0.60	0.51 to 0.68	0.93	0.91 to 0.94
	BMI ≥97th centile	%BF (BIA) ≥97th centile	Boys	29 5	22 1090	0.85	0.70 to 0.94	0.98	0.97 to 0.99
			Girls	27 7	33 1073	0.79	0.63 to 0.90	0.97	0.96 to 0.98
	BMI ≥97th centile	%BF (skinfolds) ≥97th centile	Boys	27 7	56 1056	0.79	0.63 to 0.90	0.95	0.94 to 0.96
			Girls	25 9	44 1062	0.74	0.57 to 0.85	0.96	0.95 to 0.97
	BMI ≥97th centile	TSF ≥97th centile	Boys	28 6	56 1056	0.82	0.66 to 0.92	0.95	0.94 to 0.96
			Girls	21 13	44 1062	0.62	0.45 to 0.76	0.96	0.95 to 0.97

TABLE 25 Studies of diagnostic accuracy of obesity (cont'd)

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Reilly (1999)[104]									
Index tests:	BMI SDS >85th centile	%BF >25%	Boys	9 2	13 99	0.82	0.52 to 0.95	0.88[a]	0.81 to 0.93
1. BMI. 85th centile (SD = 1.04) taken from UK reference data. SD >2.00 internally derived		%BF >32%	Girls	5 0	12 98	1.00	0.57 to 1.00	0.89[a]	0.82 to 0.94
2. Ideal body weight (IBW). Using WHO reference values	IBW >120%	%BF >25%	Boys	7 4	18 95	0.64	0.35 to 0.85	0.84	0.76 to 0.90
Reference standard:		%BF >32%	Girls	5 0	12 99	1.00	0.57 to 1.00	0.89	0.82 to 0.94
1. %BF	BMI SD >2.00	%BF >25%	Boys	4 7	2 110	0.36	0.15 to 0.65	0.98	0.94 to 1.00
Derived from triceps and subscapular skinfold measurements. Obesity cut-offs were 25% for boys and 32% for girls		%BF >32%	Girls	3 2	1 109	0.60	0.23 to 0.88	0.99	0.95 to 1.00

131

TABLE 25 *Studies of diagnostic accuracy of obesity (cont'd)*

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Reilly (2000)[105] **Index test:** I. BMI. Centiles were taken from UK 1990 reference data. 'Obesity' cut-offs from the IOTF were 19.3 and 19.2 for 7-year-old boys and girls, respectively; 'overweight' cut-offs were also used	BMI >86th centile	%BF >95th centile	All	187 10	541 3210	0.95	0.91 to 0.97	0.86	0.84 to 0.87
	BMI >88th centile		All	187 10	461 3290	0.95	0.91 to 0.97	0.88	0.87 to 0.89
	BMI >90th centile		All	186 11	392 3359	0.94	0.90 to 0.97	0.90	0.89 to 0.90
Reference standard: I. %BF Estimated using BIA methods; 95th centile, the obesity cut-off, was internally derived	BMI >92nd centile		All	181 16	318 3433	0.92	0.87 to 0.95	0.92	0.91 to 0.92
	BMI >94th centile		All	176 21	244 3507	0.89	0.84 to 0.93	0.93	0.93 to 0.94
	BMI >95th centile		All	173 24	225 3526	0.88	0.83 to 0.92	0.94	0.93 to 0.95
	BMI >96th centile		All	163 34	157 3594	0.83	0.77 to 0.87	0.96	0.95 to 0.96
	BMI >98th centile		All	140 57	65 3686	0.71	0.64 to 0.77	0.98	0.98 to 0.99
	BMI IOTF 'obese'	%BF >95th centile	Boys	46 54	9 1901	0.46	0.37 to 0.56	1.00	0.99 to 1.00
			Girls	70 27	28 1813	0.72	0.63 to 0.80	0.98	0.98 to 0.99
	BMI IOTF 'overweight'	%BF >95th centile	Boys	90 10	154 1756	0.90	0.83 to 0.94	0.92	0.91 to 0.93
			Girls	94 3	298 1543	0.97	0.91 to 0.99	0.84	0.82 to 0.85

continued

TABLE 25 Studies of diagnostic accuracy of obesity (cont'd)

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Schaefer (1998)[106] **Index test:** I. BMI. Centiles were derived internally	BMI >85th centile	%BF>75th centile	All	313 326	77 1838	0.49	0.45 to 0.53	0.96	0.95 to 0.97
Reference standard: I. % BF Calculated from skinfolds measurements. Centiles were derived internally		%BF>85th centile	All	249 134	130 2041	0.65	0.60 to 0.70	0.94	0.93 to 0.95
		%BF>90th centile	All	194 61	207 2092	0.76	0.70 to 0.81	0.91	0.90 to 0.92
		%BF>95th centile	All	104 24	291 2135	0.81	0.74 to 0.87	0.88	0.87 to 0.89
	BMI >90th centile	%BF>75th centile	All	230 409	38 1877	0.36	0.32 to 0.40	0.98	0.97 to 0.99
		%BF>85th centile	All	195 188	65 2106	0.51	0.46 to 0.56	0.97	0.96 to 0.98
		%BF>90th centile	All	163 92	92 2207	0.64	0.58 to 0.70	0.96	0.95 to 0.97
		%BF>95th centile	All	93 35	170 2256	0.73	0.64 to 0.80	0.93	0.92 to 0.94
	BMI >95th centile	%BF>75th centile	All	128 511	19 1896	0.20	0.17 to 0.23	0.99	0.98 to 0.99
		%BF>85th centile	All	119 264	22 2149	0.31	0.27 to 0.36	0.99	0.98 to 0.99
		%BF>90th centile	All	107 148	46 2253	0.42	0.36 to 0.48	0.98	0.97 to 0.98
		%BF>95th centile	All	70 58	73 2353	0.55	0.46 to 0.63	0.97	0.96 to 0.98

continued

133

TABLE 25 Studies of diagnostic accuracy of obesity (cont'd)

Study details	Index test	Reference standard	Sample	TP FN	FP TN	Sensitivity	95% CI	Specificity	95% CI
Wickramasinghe (2005)[107]									
Index test:	BMI SD >2	%BF >20%	Caucasian boys	2 27	1 14	0.07	0.02 to 0.22	0.93	0.70 to 0.99
1. BMI. Three different age- and gender-specific obesity cut-offs were used: SD >2 from British Growth standards; BMI >95th centile from CDC/NCHS; IOTF BMI values for obesity		%BF >30%	Caucasian girls	1 18	0 33	0.05	0.01 to 0.25	1.00	0.90 to 1.00
	BMI >95th centile	%BF >20%	Caucasian boys	1 28	0 15	0.03	0.01 to 0.17	1.00	0.80 to 1.00
		%BF >30%	Caucasian girls	4 15	0 33	0.21	0.09 to 0.43	1.00	0.90 to 1.00
Reference standard:	BMI IOTF 'obese'	%BF >20%	Caucasian boys	0 29	0 15	0.00	0.00 to 0.12	1.00	0.80 to 1.00
1. %BF		%BF>30%	Caucasian girls	0 19	0 33	0.00	0.00 to 0.17	1.00	0.90 to 1.00
Total body water was measured using an isotope dilution method. From this, fat-free mass and fat mass (and hence %BF) were derived using age- and gender-specific constants. Obesity thresholds used were 20% for boys and 30% for girls									

[a] Data are inconsistently reported in the original study.
FN, false negatives; FP, false positives; TN, true negatives; TP, true positives.

Appendix 10

Details of studies reporting human resources data

TABLE 26 *Studies evaluating human resource issues of growth monitoring*

Study details	Intervention	Population	Results
Aszkenasy (2005)[71] **Design:** Audit	**Aim:** To ascertain and improve the uptake of screening, the proportion of children with heights below the 0.4th centile detected, and the proportion of those below this threshold with a treatable growth disorder, and the proportion of children who have a treatable problem who are detected **Intervention:** Data were converted to a centile score for each child and school notes were retrieved for each child on or below the 0.4th centile. Three audit cycles were performed in 1999–2000, 2000–1 and 2001–2. After each audit cycle the results were presented to school nurses. A new form was developed and introduced after the first audit cycle (1999–2000). The new form does not require a graphical plot to determine whether the height of a child was below the threshold for referral **Outcome:** The number of children detected by the computer analysis and the number who were detected by the school nurse were obtained	**Location:** Middlesbrough **Selection procedure:** School entry cohorts in 1999–2000, 2000–1 and 2001–2. Height measurements were obtained in 83, 83 and 82% of eligible children in the respective cohorts	The number of children correctly identified by the school nurse improved over the duration of the study from 43% (6/14) in 1999–2000, 70% (7/10) in 2000–1 to 83% (10/12) in 2001–2 school entry cohorts
Cowan (2001)[80] **Design:** Before-and-after study	**Aim:** To assess the effects of an intensive training programme in screening for growth disorders for community based staff **Intervention:** Community nurses (health visitors and school nurses) and community paediatricians attended a 1-day training course in techniques and importance of measuring and recording growth. Lectures were also given on the medical conditions associated with growth abnormalities **Outcome:** An audit of the clinic database was undertaken to identify referrals to the growth clinic in the 18 months before and after the intervention. Referral rates of community based medical and nursing staff were compared with those from GPs (who had not received the training intervention and acted as control)	**Location:** Cardiff, Wales **Selection procedure:** 254 referrals were potentially eligible. 45 case notes could not be identified and 112 referrals were excluded. The remaining 97 referrals were included in the final audit **Included children:** Age of referred children with a diagnosis ranged from 0.4 to 17.2 years. Gender is not given	In the 18 months prior to the training there were 15 referrals to the growth clinic from community staff (2 with significant pathology) and 25 referrals from the GP (1 with significant pathology) In the 18 months after training there were 30 referrals from the community staff (6 with significant pathology) and 27 from the GP (2 with significant pathology) Post-training, 20 (4) of the referrals from community staff were made in the first 6 months; in the following 6-month periods there were only 4(0) and 6 (3) referrals, respectively, suggesting that the increase in referral rate post-training was not sustained. The number of referrals from GPs in each period was 10 (1), 8 (0) and 9 (1) No difference was found in the age or height of children referred before or after the intervention

TABLE 26 Studies evaluating human resource issues of growth monitoring (cont'd)

Study details	Intervention	Population	Results
Lipman (2004)[88] **Design:** Cluster RCT **Quality assessment results:** (see Appendix 6 for tool)	**Aim:** To determine whether instructing healthcare providers on the correct measurement and the use of accurate equipment results in more accurate linear measurement	**Location:** USA: Philadelphia, PA; Galveston, TX; St Louis, MO; New Orleans, LA; Providence, RI; Broward County, FL; Albany, NY	**Measurement technique:** At baseline, 29% of children in the control group and 32% of those in the intervention group were measured using an appropriate technique
1.1. Research question — Adequately addressed	**Intervention:** All PCPs were visited at baseline and the technique and accuracy of staff measurements of children's height/length (using their own technique and equipment) were compared with those of study coordinators (using gold standard equipment/technique)	**Selection procedure:** 55 PCPs (44 paediatric and 11 family practice providers) from each of the locations who responded to a mailed survey were randomly selected to participate in the study. After obtaining consent the PCPs were assigned to intervention or control	Significantly more children were measured using the correct technique in the intervention group compared with the control group at 3 months (54 versus 23%; $p < 0.0005$) and at 6 months (74 versus 26%; $p < 0.0005$).
1.2. Randomisation — Adequately addressed			**Accuracy:** At baseline the mean difference between PCP staff and study coordinators was 1.3 cm in the control group and 1.2 cm in the intervention group
1.3. Concealment — Not reported	PCPs assigned to the intervention were given a permanently affixed wall-mounted plastic Accustat with a level rolling head plate to measure height and a paediatric length board to measure length. A training session was delivered by site coordinators, who were paediatric endocrine nurses with an average of 8.2 years of experience of measuring children. The session included a written pre-test of knowledge of the appropriate linear growth assessment, a presentation and handouts on growth disorders and linear growth techniques. The session also included information on the accurate installation of equipment, a demonstration of correct measurement and a written assessment		Significantly more children were measured accurately in the intervention group compared with the control group at 3 months (55 versus 37%; $p = 0.003$) and at 6 months (70 versus 34%; $p < 0.005$)
1.4. Blinding — Poorly addressed			The accuracy of measurement in the intervention group improved from a mean difference (SD) between PCP staff and study coordinators of 1.2 (1.2) cm at baseline to 0.5 (0.8) cm at 6 months
1.5. Groups at baseline — Adequately addressed			
1.6. Differences in treatment — Not reported		**Included children:** Children measured were aged 0–18 years (mean 5.0 years). 46% of children were male	
1.7. Outcome measurement — Adequately addressed			The accuracy of measurement was significantly greater in the intervention group compared with the control group at 3 months (0.6 versus 1.0 cm; $p = 0.0001$) and at 6 months (0.5 versus 1.1 cm; $p < 0.0005$)
1.8. Percentage drop-out — No cluster drop-outs reported; 15/127 (12%) measurers dropped out	PCPs assigned to the control were not provided with the training session, the equipment or information on linear growth measurement		
1.9. Intention-to-treat analysis — Not applicable			Multivariate analyses found that accuracy of measurement was 1.3 times more likely with a 5-year increase in child's age and that registered nurses were twice as accurate as measurers with other educational backgrounds
1.10. Site comparability — Not addressed	**Outcomes:** Accuracy of PCP staff measurement and use of correct measurement technique was assessed at 3 and 6 months. Equipment was considered accurate based on recommended criteria, measurement technique was considered correct by observation of the measurement and measurements were considered accurate if the measurement made by the PCP staff and study coordinator did not differ by more than 0.5 cm		
2.1. Minimisation of bias — +			

PCP, Primary care practice.

138

TABLE 26 *Studies evaluating human resource issues of growth monitoring (cont'd)*

Study details	Intervention	Population	Results
Welch (1982)[95] **Study design:** Comparative study	**Aim:** To compare the effectiveness of physicians with trained volunteers and public health nurses in the detection of abnormalities in height, weight, vision, hearing, blood pressure and dental caries **Intervention:** Physician (pre-school) programme: the pre-school examination included measurement categories described above and was a prerequisite for enrolment School screening programme: physical education teachers measured height and weight and public health nurses plotted values on growth charts. Professionally trained volunteers performed vision and hearing testing and abnormalities were confirmed by public health nurses. Dental caries and blood pressure were determined by public health nurses. Those involved in the school screening programme were unaware of the results of the pre-school examination **Outcomes:** The number of abnormalities detected by the pre-school examination was compared with those detected by the school screening programme. Comparisons were also made by physician group Only results for height and weight are presented herein. Height or weight values greater than the 95th centile or less than the 5th centile were considered abnormal	**Location:** Virginia, USA **Selection procedure:** Children registered in a school district for the 1977 to 1978 school year were included in the study **Included children:** 1158 kindergarten children	There were 19 children with height-related abnormalities. The pre-school examination identified 13 and the school screening programme identified 16. The difference was not statistically significant There were 64 children with weight-related abnormalities. The pre-school examination identified 45 and the school screening programme identified 52. The difference was not statistically significant There was no significant difference between types of physician in abnormality detection rate

TABLE 26 Studies evaluating human resource issues of growth monitoring (cont'd)

	Methodology	Results
Cotterill (1996)[78]	**Aim:** The study was based on the data obtained in the Hackney growth study.[82] The aims were (1) to determine whether the UK 1990 charts more accurately represent the current child population of Hackney than the Tanner and Whitehouse charts, (2) if the UK 1990 charts are considered appropriate then what would be the effect of the change from the Tanner and Whitehouse charts to the UK 1990 charts on the workload of community health workers and (3) to determine the effect on the referral of children with short stature if the referral were changed from the 3rd to the 0.4th centile **Method:** Height data obtained from the Hackney growth study[82] were examined using the Tanner and Whitehouse and UK 1990 charts Children with a diagnosis of GHD attending a growth clinic for treatment who had a height below the 2nd centile were identified. Height at diagnosis was obtained and compared with the UK height standards. Children were divided according to the 0.4th centile and the characteristics of those below the 0.4th centile and above the 0.4th centile were compared **Outcomes:** The number of children with height below the Tanner and Whitehouse 3rd centile was compared with the number below the UK 1990 3rd and 0.4th centiles. The height at diagnosis of children with GHD was also determined to establish the consequence of using the UK 1990 charts	At age 5 years, using the Tanner and Whitehouse charts, 1% of children had a height below the 3rd centile whereas using the UK 1990 charts 3.3% of children had a height below the 3rd centile and 0.47% below the 0.4th centile At age 11 years, the Tanner and Whitehouse charts identified 2.1% and the the UK 1990 charts 2.93% of children with a height below the 3rd centile and 0.34% below the 0.4th centile This suggests that the use of the UK 1990 charts with a cut-off for referral for short stature below the 3rd centile would lead to an increase in workload of 2–3-fold, and the use of 0.4th centile would reduce workload by 50%, which may exclude a number of children with abnormalities from further assessment Of the 68 children with a diagnosis of GHD receiving treatment and height below the 2nd centile, 28 had a height at diagnosis between the 0.4th and 2nd centile (mean height SD score −2.32). The authors state that this suggests a significant proportion of children with pathology would not be referred for further investigation if the 0.4th centile were used as a threshold for referral The authors suggest that the UK 1990 2nd centile should replace the Tanner and Whitehouse charts. Children below the 0.4th centile should be directly referred and children between the 2nd and 0.4th centiles should undergo an intermediary assessment to confirm height to avoid referral of children with mild familial short stature

continued

TABLE 26 *Studies evaluating human resource issues of growth monitoring (cont'd)*

	Methodology	Results
Mulligan (1998)[89]	**Aim:** At the time of this study, the UK joint working party on child health surveillance recommended that all children should be measured at the age of 5 and again between 7 and 9 years. The aims of the study were: 1. to use longitudinal data of children measured in the community to determine the number of school-age children in the normal population who are likely to be identified for referral according to the above guidelines using the UK 1990 charts 2. to compare the results of children measured in the community with those from the Wessex growth study who were measured at 6-monthly intervals in a research setting by a single trained observer **Method:** *Community setting:* 11 schools across the Southampton health district were included in the study. Each school had a different nurse, and some changed during the period of the study. School nurses had received routine training in the technique of measuring height and relevance of height monitoring Height data were obtained on 486 children (247 boys and 239 girls) at age 5 and again at age 8 years. Height measurements were converted to HSDSs using the 1990 UK reference data. The pre-pubertal change in height standard deviation score (ΔHSDS) was derived by subtracting the measurement at age 8 years from the HSDS of the measurement at age 5 years *Research setting:* 140 children identified at school entry as being <3rd centile according to the Tanner and Whitehouse charts (short normal) were matched with children whose heights were within 10th to 90th centiles (controls). These children were measured every 6 months by a single observer in the research setting. The height data of 212 children (109 short normal and 103 controls) with measurement in the first year of school (mean age 6.05 years) and 3 years later (mean age 9 years) were converted to HSDS using the UK 1990 charts **Outcomes:** The number of children identified as having abnormal stature (defined as <0.4th centile or >99.6th centile) were obtained, along with the number who had a abnormal growth rate (defined as an HSDS change >0.67, the equivalent of one centile on the UK 1990 charts)	At school entry seven children met the referral criteria for abnormal stature (four were <0.4th centile and three were >99.6th centile). At age 8 years there were no new referrals The HSDS of 466 children (96%) did not change by more than 0.67. 11 (2.3%) children had a decrease by more than 0.67 and were considered to be 'slow growing', and nine (1.9%) had an increase by more than 0.67 293 (60%) of children were measured by a different nurse and the ΔHSDS was higher for two observers compared with one, although this was not statistically significant (0.36 versus 0.32; *p* = 0.16). Nine of the 11 children who were considered slow growing were measured by different nurses. Correlation between the two measurements was 0.948 for a single observer and 0.933 for two observers The variance of the change in HSDS was less in ideal research conditions (the Wessex study) than for these community measurements. However, the number of children whose HSDS changed by more than 0.67 was not significantly different and the data obtained in the community setting were comparable to those collected in the research setting

TABLE 26 Studies evaluating human resource issues of growth monitoring (cont'd)

	Methodology	Results
van Buuren (2004)[90]	**Aim:** At the time of the study, the Dutch Institute for Health Care Improvement had published consensus guidelines for the early diagnosis and treatment of short stature. The aim of this study was to estimate the number of referrals following strict adherence to the proposed guidelines **Method:** The guidelines proposed six screening rules based on HSDSs. Referral was recommended for boys younger than 10 years and girls younger than 9 years if one or more of the following criteria were met: 1. HSDS is lower than −2.5 (absolute height SD) 2. HSDS is lower than −1.3 and HSDS is 1.3 lower than the target height SDS (parental height corrected) 3. The growth curve deflects by more than 0.25 SD per year (deflection: the guidelines recommend that at least three measurements should be taken at least 6 months apart) 4. HSDS decreases by more than 1 SD over several years (slow SDS loss) 5. In children born small for dates (birth length SD lower than −1.88) HSDS is lower than −1.88 after the age of two (No catch up) 6. HSDS is lower than −1.3 and the child has disproportion or dysmorphic features (clinical symptoms) Longitudinal height data were obtained for all children (481 boys and 489 girls) born in 1989 and 1990 in Landgraaf and Kerkrade, The Netherlands. The number of recorded heights was 14,310 **Outcomes:** The percentage of referrals according to each of the referral criteria was calculated	The percentage of children meeting the referral criteria were as follows: 1. Absolute height SD = 6.2% 2. Parental height corrected = 5.9% 3. Deflection = 31.5% 4. Slow SD loss = 5.5% 5. No catch up = 1.4% 6. Clinical symptoms = no data available Screening rules 1–4 combined = 38.2% This corresponds to 77,000 children in The Netherlands each year, which is 30 times higher than what was expected by the consensus committee. The authors state that strict adherence to the proposed guidelines would result in a substantial number of false positives and would impair regular practice and lead to avoidable anxiety

Appendix 11

Details of studies reporting attitudinal data

144

TABLE 27 *Attitudes to growth monitoring*

Study	Attendance at initial measurement	Attendance at referral	Other attitude related data
Agwu (2004)[76]	3474/3864 (90%) of eligible children were measured and 3465/3864 (89.6%) were judged to have been appropriately screened (measured and screening forms were completed correctly)	Data were only available on 14/18 of the children meeting the criteria for referral. No reasons were given for the failure to follow up the remaining four	
Ahmed (1995)[56]	20,338 children were measured. Coverage for 3-year-olds was approximately 61% in the 1st year and 71% in the 2nd year, and for 4.5 years was 61 and 65% in the respective years. Reasons why measurements were not provided include relocation of child, child had transferred to different practice, child did not attend child health clinic, child was too uncooperative to measure, the parent refused or the measurement was not recorded	76/260 (29%) children referred were lost to follow-up. Reasons stated were 33 refused referral (13%), 11 moved out of the area, nine were followed up by health visitors only, 10 did not attend (4%) and 13 attended clinic only once	The authors state that ascertainment could be improved and there were continual problems with encouraging and reminding some of the health visitors to measure children or including the measurement on the appropriate forms. This suggests a need for regular training
Aszkenasy (2005)[71]	9338/11316 (83%) of eligible children were measured. Reasons for non-measurement were not provided	9/24 (38%) children who were invited to attend assessment attended the clinic. Of those who did not attend one child moved out of the area (4%), parents of two children had no concerns and did not wish their children to be seen by a paediatrician (8%) and 12 did not attend (50%)	The authors state that there is a large proportion of children who do not attend the clinic and reasons for this need to be explored
Banerjee (2003)[59]	1862/2354 (79.1%) of eligible children had at least one measurement of height or weight, 2.8% refused and 18.1% had no data for unexplained reasons. 1592 (67.6%) had a height measurement and 1862 (79.1%) had a weight measurement	6 (40%) failed to attend the initial referral despite being offered two appointments. Input from health visitors (home visit/telephone call) led to four of these non-attendees being assessed	The authors proposed that the reason for only 80% of eligible children being measured may be due to problems with school attendance on any given day (average 92% in the LEA), the opt-in consent policy of the NHS trust and the shortage of school nurses, meaning opportunistic measurement of defaulters was impractical

TABLE 27 *Attitudes to growth monitoring (cont'd)*

Study	Attendance at initial measurement	Attendance at referral	Other attitude related data
Cernerud (1994)[79]	7129 children were measured. These represent random samples of the school classes	20/7129 of children were referred for investigation of stature. It is not clear if any children were lost to follow-up, although the authors state that no new diseases were identified in the children referred	The benefits of the surveillance were assessed by an expert panel. No member of the panel stated that there was any single procedure they wanted to promote at the cost of the school health anthropometry or which they had to abandon due to growth surveillance The benefits of growth surveillance were assessed by an expert panel of experienced school nurses and senior paediatricians who also work as school doctors. The most important benefits were: ● the opportunity to meet every child under structured and uncontroversial conditions ● conditions such as malnutrition and drug abuse may be revealed during measurement ● growth data can be used in discussions with children, young persons and their parents when they worry about growth and what is normal ● provides data for future medical consultations in school health service, in outpatient clinics and in hospitals ● data can be used in the field of public health research
de la Puente (1999)[81]	2084/4612 (45%) were measured. Reasons for low coverage of measurements were not given	6/73 (8.2%) did not attend the appointment with the endocrinologist, four (5.5%) had been wrongly measured on assessment by the endocrinologist and for 14 (19.2%) no definite diagnosis could be reached	
Hearn (1995)[82]	9549/12051 (79%) of eligible children were measured. This corresponds to a coverage in the first year of primary and secondary school entrants of 73 and 35%, respectively, in the second year was 91 and 86%, respectively and in the third year was 89 and 87%, respectively. The authors attribute the increase to improvements in school nurse working practice	Data were given only on the first 100 children referred for further investigation. 26 out of the first 100 (26%) referred did not attend the community growth clinic and were offered an appointment at the hospital clinic. A further two children had moved out of the area prior to attending the growth clinic	

continued

145

146

TABLE 27 *Attitudes to growth monitoring (cont'd)*

Study	Attendance at initial measurement	Attendance at referral	Other attitude related data
Keller (2002)[83]	Outcome data are presented on 60,984 children. There are no data on coverage. An additional related study provides data on 83,721 children and adolescents but diagnosis of growth disorders was not given[84]	476 cases were referred and examined for further investigation	
Lacey (1974)[86]	2256/5000 (45%) children born in 1960 had a measurement at the age of 10 years and were eligible for inclusion in the study. The authors state that the remaining children had moved out of the area	13/111 (12%) did not undergo further investigation as parents did not give consent	
Lindsay (1994)[87]	114,881 measurements were available for evaluation. 4.9% of the eligible population were absent due to illness at the time of measurement, and 7.3% of measurements at tabulation were considered incomplete. Only 69% of children had a second height measurement to detect growth rate. 14% of children were unavailable for remeasurement due to interstate or intrastate migration	630 children were identified as having a possible growth problem. 75 were lost to follow-up: 49 moved, 26 (4%) refused evaluation. 555 children attended for further investigation	The authors state that some children at risk of having a growth problem were not examined by the study physicians due to lack of parental concern or refusal to allow follow-up. 88% of children with GHD and 60% of those with TS had seen a physician on at least one occasion. Physicians had raised concerns regarding the heights of 60% of GHD and 67% TS, but only 25% of GHD and no children with TS had seen an endocrinologist prior to the study. The authors suggest the need for further education for physicians on growth disorders
Vimpani (1981)[91]	48,221 child height measurements were included in the study. The number of potentially eligible children is not reported 554/630 (88%) of the absentees thought to be small according to class teachers were measured on a later date. Up to nine short children may have been among those who were not measured	The parents of 77/449 (17%) children with short stature refused to consent to further investigation. Ten children had emigrated before being able to participate in further investigation. Of the 264 children included in the further investigation, 14 did not undergo an insulin hypoglycaemia test and could not be classified due to emigration (1 child) or parents' refusal	Parents of five of the six children with a new diagnosis of GHD and in whom genetic cause of disease could be excluded had previously consulted their GP but none had been referred for further investigation. The parents of one child had not raised concerns regarding their child's height

TABLE 27 Attitudes to growth monitoring (cont'd)

Study	Attendance at initial measurement	Attendance at referral	Other attitude related data
Voss (1992)[93]	14,346 children were screened. It is not clear if these represent all eligible children	Of 180 referred, 19 declined further investigation (11%)	128/152 underwent further investigation – 24 did not (i.e. five more than reported as refusing referral) Some independent schools were not willing to participate as height measurements were not routinely made or it was perceived that height 'problems' were not their concern and should be dealt with by the school's GP. The authors state that it was decided not to include children from these schools in the study
White (1995)[96]	23,046/26,700 eligible children were measured and included in the study. This represents 91% of school-age children and 63% of 3-year-olds. Data on 52 children were incomplete and not included in the analysis. Parents of 36 children refused to participate in the screening programme. The most common reason stated by the parent was that the child was overweight and they did not wish attention drawn to this condition	Not applicable as no further data were given on those who were outwith the 3rd or 97th centile	

LEA, local education authority.

TABLE 28 Studies of attitudes to monitoring for obesity

Study	Intervention	Population	Results
Routh (2005)[72]	**Aim:** To test a simple, low-cost method of obtaining data on the prevalence of childhood obesity. The method involved taking height and weight measurement within a numeracy and data handling lesson for 9–11-year-old children **Intervention:** The lesson was delivered to year 5 classes (age 9–10 years) in seven schools. School nurses were involved in the lesson along with the investigator, and their queries and concerns were addressed in a pre-intervention meeting. School health link workers attended some lessons to help with measurements. Parental consent was obtained on an opt-out basis. To avoid an emphasis on weight, weighing was just one of a range of activities, each child was weighed separately and their weight kept private from the rest of the class. Weight and height data were anonymised, and substituted with 'dummy' results for the classes' use **Outcome:** BMI was calculated for each child and compared with UK BMI reference charts and the number of boys and girls who were normal weight, overweight or obese were determined A questionnaire was sent to participating schools to evaluate the acceptability of the methods used	**Location:** Birmingham **Selection procedure:** All primary schools in north Birmingham were approached. Seven agreed to take part	Parents of 3/255 (1.2%) children withdrew children from the study and children were not referred after screening Data were retrieved on 252 children. 20% were overweight and 7% obese. 71% were normal and 2% underweight. There was a marked gender difference; among boys, 15% were overweight and 5% obese; among girls 24% were overweight and 9% obese. There was some evidence of higher levels of overweight and obesity in areas with higher levels of deprivation Results of the questionnaire were based on 5/8 questionnaires returned by teachers and 4/7 questionnaires returned by nurses. The results suggested that both teachers and nurses were happy with the method, the weighing was done in a sensitive manner and there were no subsequent instances of stigmatism of overweight

Appendix 12

Full structured abstracts of included economic evaluations

1. Cost-effectiveness of group and mixed family-based treatment for childhood obesity

Goldfield GS, Epstein LH, Kilanowski CK, Paluch RA, Kogut-Bossler B. *Int J Obesity* 2001; **25**:1843–9.

Health technology
The use of mixed treatment for childhood obesity, incorporating both group and individual approaches to treatment. The comparator was group treatment only.

Disease
Nutritional and metabolic diseases.

Type of intervention
Treatment.

Hypothesis/study question
The aim of the study was to determine the cost-effectiveness of two protocols for the delivery of family-based behavioural treatment for childhood obesity. The health technology studied was mixed treatment, which incorporated both group and individual treatment approaches. The comparator was group treatment only. The common components of the treatments were: a 13-session programme on diet, activity, behavioural change techniques, parenting and coping with psychosocial problems; the Traffic Light Diet; reinforcement for physical activity; self-monitoring; and stimulus control.

The patients in the mixed treatment group had individual sessions of 15–20 minutes' duration with a therapist and 40 minutes of group therapy. The patients in the group treatment had sessions of 35–40 minutes' duration. The group treatment was justified because it involved fewer staff. The study was conducted from the perspective of the health service.

Economic study type
CEA.

Study population
The study population comprised families with obese children aged from 8–12 years. The families had to meet the following inclusion criteria: the child was between 20 and 100% overweight; neither parent was greater than 100% overweight; one parent was willing to attend treatment meetings; no family member was participating in an alternative weight control programme; no child or parent had current psychiatric problems; and there were no dietary or exercise restrictions on the child or parent.

Setting
The setting was community. The economic analysis was carried out in the USA.

Dates to which data relate
The dates during which the effectiveness, resource use and cost data were obtained were not reported. The price year was not reported.

Source of effectiveness data
The effectiveness data were derived from a single study.

Links between effectiveness and cost data
The costing was carried out prospectively on the same patient sample as that used in the effectiveness analysis.

Study sample
Thirty-one families were enrolled in the study. The families were recruited through newspaper advertisements and by physician referrals. No power calculations were performed to determine the sample size.

Study design
This was a randomised controlled study carried out in the community. The method of randomisation was not stated. The families were seen at 6 and 12 months after the treatment was started. Two families dropped out before treatment began and five families refused to participate in the follow-up

assessments. There were 24 families for which complete data were available.

Analysis of effectiveness

The clinical study was analysed on an intention-to-treat basis. The primary health outcomes used were the reduction in standardised BMI (Z-BMI) and the percentage overweight. The groups were comparable in terms of the demographic and anthropometric characteristics of the children and parents, with the exception of the parents' height.

Effectiveness results

There was a significant change over time in terms of how overweight the participant was (percentage overweight), ($p < 0.001$).

For children ($n = 24$), the percentage overweight changed by –9.97 from baseline to 6 months, and by –8.04 from baseline to 12 months. The Z-BMI changed by –0.59 from baseline to 6 months and by –0.64 from baseline to 12 months.

For parents ($n = 24$), the percentage overweight changed by –6.67 from baseline to 6 months and by –5.31 from baseline to 12 months. The Z-BMI changed by –0.31 from baseline to 6 months and by –0.29 from baseline to 12 months.

For obese parents ($n = 18$), the percentage overweight changed by –7.03 from baseline to 6 months and by –5.70 from baseline to 12 months. The Z-BMI changed by –0.39 from baseline to 6 months and by –0.31 from baseline to 12 months.

There were no main effects or interactions due to the type of group or generation.

Clinical conclusions

The two groups were similar in terms of the Z-BMI and the percentage overweight for the children and their parents.

Measure of benefits used in the economic analysis

The benefit measures used in the economic analysis were the reductions in Z-BMI and percentage overweight.

Direct costs

The direct costs were not discounted due to the short time frame of the study (less than 1 year). The quantities and costs were reported separately. The direct costs related to the orientation costs and the treatment costs. The orientation costs included advertising, materials and salary. The treatment costs included materials, travel and salary. The quantity/cost boundary adopted was that of the health service. The source of the cost data was not reported. The price year was not reported.

Indirect costs

The indirect costs were not included.

Currency

US dollars ($).

Statistical analysis of costs

The costs were analysed using a one-way analysis of variance.

Sensitivity analysis

No sensitivity analyses were reported.

Estimated benefits used in the economic analysis

See 'Effectiveness results' above.

Cost results

The total costs amounted to $491.48 per family for the group only treatment and $1390.72 per family for the mixed treatment. The difference was statistically significant.

Synthesis of costs and benefits

A cost-effectiveness ratio was calculated for each benefit measure by dividing the charges over 12 months by the total cost. At 12 months, a decrease of 0.005 percentage overweight units per dollar was observed for the mixed group, compared with a decrease of 0.014 percentage overweight units per dollar with the group treatment.

Authors' conclusions

The authors argued that family-based behavioural treatment for childhood obesity was more cost-effective when provided in a group format than when provided in a combined group and individual approach. The cost-effectiveness of the treatment extended to parents.

CRD commentary

Selection of comparators

The choice of comparator was justified on the grounds that it involved fewer staff. You should decide if these health technologies are relevant to your setting.

Validity of estimate of measure of effectiveness

The analysis was based on a randomised controlled study, which was appropriate for the study question, and should have good validity

although the sample size was small partly due to drop-outs and refusals (to participate). The method of randomisation was not stated. The inclusion criteria were reported, as were the demographic characteristics of the children and parents. The treatment groups were comparable in terms of the demographic and anthropometric characteristics of the children and parents, with the exception of the parents' height. The study sample was representative of the study population. Appropriate statistical analyses were undertaken.

Validity of estimate of measure of benefit
The benefits were estimated directly from the effectiveness analysis. Two measures of health benefit were therefore used in the economic analysis.

Validity of estimate of costs
The positive features of the cost analysis were that all relevant direct cost categories were included and that the quantities and costs were reported separately. This made it possible to replicate the cost results in other settings. Further, statistical analyses were performed on the cost estimates. However, the price year was not reported. The authors did not conduct sensitivity analyses on the quantities or costs, which may have limited the generalisability of the results. The authors included the costs of recruiting patients, but did not quantify the costs to the families who participated in the study.

Other issues
The authors made appropriate comparisons of their findings with those from other studies, but did not address the issue of generalisability to other settings. The authors did not seem to present their results selectively. The study considered families with obese children aged from 8 to 12 years, and this was reflected in the authors' conclusions. Given that the effectiveness results were shown to be similar between groups, the authors could have carried out a cost-minimisation analysis. It should be noted that the study had a small sample size and no power calculations were reported.

The dates during which the effectiveness, resource use and cost data were collected were not reported, nor was the price year. Further research is needed to determine if the current results are generalisable to more obese children, since the population studied was mild to moderately obese. The authors stated that there may be some self-selection bias inherent in recruiting families for a study that provides family-based treatment. They also stated that the effects of the interventions may change over time, implying that analyses over a long time frame would be useful.

Implications of the study
The authors argued that family-based behavioural treatment for childhood obesity was more cost-effective when provided in a group format than when provided in a combined group plus individual approach. In addition, the cost-effectiveness of the treatment extended to parents. Further research is needed to determine if the current results generalise to more obese children, and also to determine the cost-effectiveness of treating only one member of the family versus concurrent treatment of both parent and child.

2. Economic analysis of a school-based obesity prevention program

Wang LY, Yang Q, Lowry R, Wechsler H. *Obesity Res* 2003;**11**:1313–24.

Health technology
The use of a school-based obesity prevention programme, referred to as Planet Health. Planet Health was designed to reduce obesity in youth of middle-school age. The programme was an interdisciplinary curriculum, whereby intervention material was infused into four major subject areas (language arts, mathematics, science and social studies) and into physical education. Sessions focused on decreasing television viewing, decreasing consumption of high-fat foods, increasing fruit and vegetable intake and increasing moderate and vigorous physical activity.

Disease
Nutritional and metabolic diseases; healthcare: health promotion.

Type of intervention
Primary prevention.

Hypothesis/study question
The objective of the study was to assess the cost-effectiveness and cost–benefit of Planet Health, a school-based intervention designed to reduce obesity in youth of middle-school age. The school-based intervention was compared with a no-intervention alternative, whereby students received the usual curricula and physical education classes. A societal perspective was adopted in the economic analysis.

151

Economic study type
CUA.

Study population
The study population comprised male and female students of middle-school age.

Setting
The setting was the community. The economic study was conducted in the USA.

Dates to which data relate
The effectiveness data were derived from a study reporting efficacy data on the Planet Health intervention that was conducted in 1995 and published in 1999 [Gortmaker and colleagues; see the section 'Other publications of related interest' (p. 155)], a study published in 1997 [Whitaker and colleagues; see the section 'Other publications of related interest' (p. 155)], and from the National Health and Nutrition Examination Study Epidemiological Follow-up Study (NHANES I EFS) conducted from 1971 to 1992. The resource use data were derived from studies and sources published between 1997 and 1999. The price year was 1996.

Source of effectiveness data
The effectiveness data were derived from the study reporting efficacy data on the Planet Health intervention, a study predicting obesity in young adulthood from childhood and parental obesity [Whitaker and colleagues; see the section 'Other publications of related interest' (p. 155)] and from the NHANES I EFS.

Modelling
A decision model was created to calculate the cost-effectiveness of the health intervention over 25 years. A two-stage overweight progression model was used to determine the expected number of adulthood overweight cases by age 40 years among the 310 female students in the intervention, compared with the same 310 students in a hypothetical no-intervention condition. Overweight was defined as a BMI of at least 25 kg/m^2.

Study sample
Efficacy data on the Planet Health intervention was derived from the study by Gortmaker and colleagues. The authors provided brief details of this study. In 1995, 10 middle schools in four communities in the Boston metropolitan area were randomly assigned to either the intervention (five schools) or control condition (five schools). A total of 1203 students in the schools were randomised to receive the intervention.

Study design
The study was based on an RCT that was undertaken in 10 schools. Although 1203 students in schools were randomised to receive the intervention, only 310 girls and 331 boys completed the follow-up for two school years.

Analysis of effectiveness
The analysis of the clinical study was conducted on the basis of treatment completers only. The measure of outcomes used was the change in obesity from baseline (autumn 1995) to follow-up (spring 1997) among students in the intervention and control schools. Obesity was defined as a composite indicator based on having both BMI and a TSF value greater than or equal to age- and gender-specific 85th percentiles.

Effectiveness results
The trial found that, during the 2-year intervention, the prevalence of obesity among girls declined from 23.6 to 20.4% in the intervention schools ($n = 310$), but increased from 21.5 to 23.7% in the control schools.

After controlling for baseline covariates, the prevalence of obesity among girls in the intervention schools was reduced significantly compared with girls in the control schools (odds ratio 0.47, 95% CI: 0.24 to 0.93; $p = 0.03$).

No significant differences were found among boys.

Clinical conclusions
The study concluded that the Planet Health intervention was efficacious in reducing the prevalence of obesity in female students of middle-school age.

Outcomes assessed in the review
The outcomes assessed were: the probability of a 14-year-old overweight female student becoming an overweight young woman by 21–29 years of age; the probability of a 14-year-old non-overweight female student becoming an overweight young woman by 21–29 years of age; the probability of an overweight young woman aged 21–29 years becoming an overweight woman by age 40 years; the probability of a non-overweight young woman aged 21–29 years becoming an overweight woman by age 40 years; the years of healthy life scores by BMI for women aged 40–64 years; the probability of dying during the 25-year period by BMI; and the expected number of years of life after age 40 years by BMI for women who died during the 25-year period.

Study designs and other criteria for inclusion in the review

Not reported.

Sources searched to identify primary studies

Not reported.

Criteria used to ensure the validity of primary studies

Not reported.

Methods used to judge relevance and validity and for extracting data

Not reported.

Number of primary studies included

The effectiveness data were derived from a study predicting obesity in young adulthood from childhood and parental obesity (Whitaker and colleagues) and from the NHANES I EFS. Three further published studies were used to derive utility values and life expectancy.

Method of combination of primary studies

Not relevant.

Investigation of differences between studies

Not relevant.

Results of the review

The proportion overweight at ages 21–29 years was 75.4% (95% CI: 64.4 to 86.2) among those who were overweight at age 14 years, and 9.8% (95% CI: 7.5 to 12.1) among those who were not overweight at age 14 years.

The estimated proportion overweight at 40 years was 91.2% (95% CI: 85.0 to 97.5) among those who were overweight at age 21–29 years and 39.3% (95% CI: 33.5 to 45.1) among those who were not overweight at age 21–29 years.

The years of healthy life scores per woman aged 40–65 years were 0.835 (95% CI: 0.827 to 0.842) for non-overweight females and 0.753 (95% CI: 0.743 to 0.764) for overweight females.

The probability of dying during the 25-year period was 0.117 for non-overweight females and 0.152 for overweight females.

The expected number of years of life after age 40 years among women who died during the 25-year period was 16.50 for overweight females and 16.05 for non-overweight females.

Measure of benefits used in the economic analysis

The measure of benefits used was the QALYs. In the present study, QALYs were calculated using the Healthy People 2000 years of healthy life measure, in conjunction with the 1990 National Health Interview Survey (NHIS) for women aged 40–64 years. These estimates of years of healthy life were then combined with the life expectancy estimates from a published study (Gorsky and colleagues), to calculate QALYs for overweight and non-overweight women. The benefits were discounted at an annual rate of 3%.

Direct costs

The resource use and quantities were reported separately for some resource categories only. The direct costs to the third-party payer were included in the analysis. These costs included the intervention costs of Planet Health, such as teacher training workshops, wellness sessions and fitness funds and the medical costs of being overweight. The costs of teacher training included salaries for a trainer and assistant trainer for delivering the training, teachers' stipends for attending the training and cost of food provided during the training. The medical costs due to being overweight included the direct healthcare and medication costs associated with women who were currently 40 years of age and who maintained an overweight status through age 65 years. The medical costs estimated were those associated with events of fatal and non-fatal coronary heart disease, hypertension, diabetes, symptomatic gallstones and osteoarthritis. These costs were based on a published study (Gorsky and colleagues). Since the costs were incurred until the female students were aged 65 years, discounting was relevant and was appropriately performed using a rate of 3% per annum. The study reported the incremental costs. The price year was 1996.

Indirect costs

The indirect costs due to lost productivity consisted of the costs associated with lost or impaired ability to work or to engage in leisure activities because of morbidity and lost economic productivity because of death. In this study, the authors estimated the excess costs associated with excess work-days lost and excess life-years lost per overweight woman, compared with a non-overweight woman, for a period of 25 years from 40–65 years of age. The authors used the 1990 NHIS of the Health Promotion and Disease

prevention sample person file to estimate the mean work-loss days among employed women aged 40–64 years by their BMI status. The median weekly earnings per age group were derived from the Bureau of Labour Statistics of the US Department of Labour. Since the costs were incurred until the female students were 65 years old, discounting was relevant and was appropriately performed at a rate of 3% per annum. The study reported the incremental costs. The price year was 1996. However, for the CEA, the authors did not include the costs of lost productivity averted, and only included the intervention costs and the medical care costs averted.

Currency
US dollars ($).

Statistical analysis of costs
The costs were treated as point estimates (i.e. the data were deterministic).

Sensitivity analysis
To test whether the results of the base-case analysis were dependent on the accuracy of the parameter estimates derived from either the efficacy study or published studies, the authors conducted sensitivity analyses by varying 10 parameters (e.g. the conditional probabilities of being overweight, the years of healthy life scores, the expected number of years of life and the annual workdays lost). Both one- and multi-way sensitivity analyses were used. For medical cost per care prevented the authors used Gorsky and colleagues' estimates as a plausible range. For discount rate, the range was 0–5%. For other parameters, the authors used 95% CIs for each parameter and a Monte Carlo simulation (using 10,000 iterations) was performed. Further, to test whether the Planet Health programme would be cost-effective in other locations, the authors performed separate univariate analyses to examine the sensitivity of the results to the variation of intervention costs.

Estimated benefits used in the economic analysis
The number of QALYs saved due to the Planet Health Intervention was 4.13.

Cost results
The intervention costs of Planet Health were $33,677, the medical care costs averted because of Planet Health were $15,887 and the costs of lost productivity averted because of Planet Health were $25,104. Hence the authors estimated that Planet Health was associated with savings of $7313.

Synthesis of costs and benefits
The costs and benefits were combined using an incremental cost-utility ratio (i.e. the cost per extra QALY gained). The costs of lost productivity averted because of Planet Health were not included in this analysis. The authors found that the incremental cost per QALY gained was $4305 when Planet Health was compared with no intervention.

The univariate results showed that the cost-effectiveness of the programme remained relatively unaffected by changes in most of the parameter variations, but was relatively more sensitive to the annual discount rate. The authors also found that the results remained cost saving to society under most scenarios. The results of the Monte Carlo simulation resulted in 95% CIs between $1,612 and $9,010 per QALY saved.

The results of the univariate sensitivity analysis on intervention costs showed that, while teachers' stipends varied from $15 to $29, the cost-effectiveness of the intervention fell in a range of $2666–4964 per QALY saved, and the costs to society remained a net saving of $4602–14,094.

Authors' conclusions
The Planet Health programme was cost-effective and cost saving. The authors also concluded that school-based prevention programmes of this type were likely to be cost-effective uses of public funds.

CRD commentary
Selection of comparators
A justification was given for using a no-intervention programme as the comparator. It represented current practice in the authors' settings. You should decide if this is a widely used health technology in your own setting.

Validity of estimate of measure of effectiveness
The analysis was based on a trial that investigated the efficacy of a school-based obesity prevention programme on the prevalence of obesity over 2 years. The authors then constructed a model to extrapolate these results up to the age of 65 years. Two-year efficacy data on Planet Health were derived from an RCT. This was appropriate for the study question as well-conducted RCTs are the 'gold standard' study design when comparing different health interventions. Even though the authors provided only some details of this study, it would appear that the study was well conducted, with the study sample being representative of the study population and the analysis of efficacy being handled credibly.

Other data to supplement the model were based on a synthesis of published studies. The authors did not report that a systematic review of the literature was conducted, or the methodology of the review. However, the studies used appear to be very relevant and credible, as they included results of US-wide surveys on nutrition. The methods of combining the efficacy data from Planet Health, the conditional probabilities of becoming overweight, life expectancy and QoL were clearly described, and further details were reported in appendices.

Validity of estimate of measure of benefit
The estimation of benefits was modelled using a decision tree analytic model, which was appropriate for the study question. As the benefits were incurred over a long period, future QALYs were discounted at a rate of 3% per annum.

Validity of estimate of costs
All the categories of cost relevant to the societal perspective adopted were included in the analysis. Further, all relevant costs for these categories appear to have been included in the analysis. Importantly, though, the authors did not take all the direct and indirect costs associated with obesity during adolescence and young adulthood into consideration. However, the authors reported that the inclusion of these costs would have made the cost-effectiveness and savings to society due to Planet Health even larger. The authors did not include the indirect costs in the incremental cost analysis, but clearly showed that when they were included the Planet Health intervention became cost saving. The costs and the quantities were only reported separately for the intervention costs, which will enhance the generalisability of the authors' results. The intervention costs were derived from the actual costs of the Planet Health intervention. For these costs, the authors performed separate univariate analyses to test whether the intervention programme would be cost-effective in other settings. Other costs were derived from the literature, with appropriate sensitivity analyses being undertaken using appropriate ranges. Discounting was necessary, as the costs were incurred during a long period, and was appropriately undertaken. The price year was reported, which will aid any possible inflation exercises.

Other issues
The authors did not compare the results of their study with those from other studies, as no cost-effectiveness study had been published in the field of obesity prevention. The issue of generalisability to other settings was addressed in the sensitivity analyses. The authors do not appear to have presented their results selectively. In their conclusions, the authors reported that Planet Health was cost-effective, implying it was cost-effective generally. However, the authors did not mention that the study investigating the efficacy of the Planet Health intervention found it to be effective only on girls, and the subsequent analyses in this study were based on females only.

The authors reported a number of further limitations to their study. First, the study was retrospective, with intervention costs being modelled rather than measured. Second, only single data sources were available for most of the model parameters, therefore 95% CI estimates had to be used for sensitivity analyses. Third, the authors did not consider overweight relapse among students who lost weight during the 2-year study period. Fourth, although the definition of childhood obesity in the study of Planet Health was based on both BMI and body fat measures, the only progression probability estimates from the literature used only BMI measures. Fifth, intervention effectiveness was estimated from 310 female students, but the intervention costs were estimated to include all participants in the study. Sixth, the authors did not take all the direct and indirect costs associated with obesity during adolescence and young adulthood into consideration. However, the authors reported that the inclusion of these would have made the cost-effectiveness and savings to society due to Planet Health even larger.

Implications of the study
The authors reported that more research is needed on the relationship between overweight status in children and obesity in adults, and the QALYs and costs due to lost productivity of overweight and non-overweight adults. In addition, the authors recommended that future school-based obesity programmes should routinely collect programme cost information so that more cost-effectiveness calculations can be conducted. The authors also suggested that school-based obesity prevention programmes should be included in portfolios of obesity prevention programmes to reduce efficiently the burden of obesity to society.

Other publications of related interest
Gorsky RD, Pamuk E, Williamson DF, *et al*. The 25-year health care costs of women who remain overweight after 40 years of age. *Am J Prev Med* 1996;**12**:388–94.

Gortmaker SL, Peterson K, Wiecha J, *et al*. Reducing obesity via a school-based interdisciplinary intervention among youth. Planet Health. *Arch Pediatr Adoles Med* 1999; **153**:409–18.

Whitaker RC, Wright JA, Pepe MS, *et al*. Predicting obesity in young adulthood from childhood and parental obesity. *N Engl J Med* 1997;**337**:869–73.

3. Clinical effectiveness and cost-effectiveness of growth hormone in children: a systematic review and economic evaluation

Bryant J, Cave C, Mihaylova B, Chase D, McIntyre L, Gerard K, *et al. Health Technology Assess* 2002;**6**(18).

Health technology
The use of GH was evaluated in children suffering from GHD, TS, CRF, PWS or ISS. At the time of writing, the authors stated that GH was not currently licensed (in the UK) for use in ISS.

Disease
Neonatal diseases and abnormalities; endocrine diseases; urological and male genital diseases; female genital disease and pregnancy complications; nervous system diseases; nutritional and metabolic diseases.

Type of intervention
Treatment.

Hypothesis/study question
The objective of the review was to carry out an incremental CEA comparing the use of GH with placebo or no intervention during the assessment of growth in five conditions in which individuals were characterised by short stature. GH represented the standard active treatment and 'no intervention' was specified as growth monitoring. It was considered that the choice of the comparators in this review would enable the best evidence to be provided on the clinical effectiveness of GH. The economic analysis was conducted from the perspectives of the NHS and Personal Social Services (England and Wales).

Economic study type
CEA.

Study population
The study population comprised children (younger than 17 years) who were suffering from one of five conditions (GHD, TS, CRF, PWS or ISS).

Setting
The setting was secondary care. The economic study was carried out in Southampton, UK.

Dates to which data relate
The effectiveness data were derived from published studies dating from 1989 to 2000. The resource use data were derived from published and unpublished data in 1999 and 2001. The costs were presented at year 2000 prices.

Source of effectiveness data
The effectiveness data were derived from a systematic review of the literature.

Modelling
A separate model was developed for the CEA of the five conditions in the UK setting. A similar, deterministic decision tree approach was used. The same period of childhood growth was assessed, although the period varied under different scenarios.

Outcomes assessed in the review
The primary effectiveness outcomes assessed in the systematic review were: the height (cm) at a given point in time, or at completion of growth (cm, standard deviation, or relative to adult norms); the height standard deviation score (HtSDS); the growth velocity (GV); the GV relative to norms for same age children (GVSDS); bone age; body composition; and psychological outcomes.

QoL measures were also eligible for inclusion, although none were found. Compliance rates and adverse events were not included in the analysis.

The primary epidemiological outcomes assessed by an *ad hoc* review of the literature were population data (i.e. weight, age, gender distribution), incidence, prevalence and current treatment patterns associated with each of the five conditions in children under the age of 17 years.

The parameters used in the model for the economic analysis included population data, the outcome measure, drop-out rate, average age at the start of treatment, average length of treatment, costs of treatment, drug doses and discount rate.

Study designs and other criteria for inclusion in the review

RCTs, or systematic reviews of RCTs, assessing the effects of GH in comparison with placebo or no intervention were eligible for inclusion. Where final height did not feature as an outcome in at least one of the trials for a particular condition, the authors considered other study designs (controlled studies, case-controlled studies and case series) to assess this measure. Economic evaluations also formed part of the inclusion criteria, although none were found. The final cohort of included studies comprised a mixture of RCTs and non-RCTs.

Sources searched to identify primary studies

Published studies and statistics were consulted for epidemiological data and current treatment patterns. The Cochrane Database of Systematic Reviews, the Cochrane Controlled Trials Register, CRD databases (DARE, NHS EED and HTA), MEDLINE, PubMed, EMBASE, the National Research Register, the Science Citation Index, BIOSIS Previews, Econlit, the MRC Trials database, Early Warning System and Current Controlled Trials were searched for effectiveness data. All databases were searched from inception to April 2001 and were limited to articles reported in English. Further studies were identified by consultation with experts, and through bibliographies and industry submissions or trials (via NICE).

Criteria used to ensure the validity of primary studies

The validity of RCTs was assessed using the Jadad checklist. The validity of non-RCTs was judged using a modified version of the Spitzer criteria.

Methods used to judge relevance and validity and for extracting data

One reviewer undertook the data extraction and validity assessment, with a second reviewer checking them. Any disagreements were resolved by discussion.

Number of primary studies included

Thirty-two studies were included in the review. There were 21 RCTs and 11 non-RCTs.

Method of combination of primary studies

The primary studies were combined in a narrative format and the results were structured according to the five health conditions. The authors stated that, owing to the heterogeneity amongst the studies, a meta-analysis was not possible. The range of results among the included studies was used for a sensitivity analysis of some of the parameters.

Investigation of differences between studies

Differences between the studies were discussed in a narrative.

Results of the review

Higher quality evidence existed for studies measuring short-term height outcomes (full details in report). However, model parameters were based on final height gains. Base-cases were used in this analysis to reflect variations in measures of clinical effectiveness. Base-case 1 represented a larger effect size, whereas base-case 2 represented a more cautious estimate of effectiveness.

Effectiveness and epidemiological values used in the economic model are reported below (values for costs and discount rates are reported later). The data for final height assumed that the benefit was evenly spread over the treatment period.

GH in GHD

In base-case 1, the length of treatment (assuming the child was aged 9 years) was 8 years and the final height gain was 10.28 cm. In base-case 2, the length of treatment was 5 years (assuming the child was aged 12 years) and the final height gain was 8.58 cm. The drug dose (based on average age- and sex-related weight at 50th percentile and not adjusted during puberty) was 0.175 mg/kg per week (range: 0.175–0.35). The assumed drop-out rate was 9.3% after the first year of treatment. Population data suggested that 63% were boys.

GH in TS

In base-cases 1 and 2, the length of treatment (assuming the child was aged 11 years) was 5 years. The final height gain was 4.8 cm in base-case 1 and 4.4 cm in base-case 2. The drug dose (based on average age- and sex-related weight at 50th percentile and not adjusted during puberty) was 0.30 mg/kg per week (range: 0.17–0.70). The assumed drop-out rate was 17% after the first year of treatment and 41% from monitoring after the first year of monitoring. Population data suggested that all were girls.

GH in CRF

In base-case 1, the length of treatment (assuming the child was aged 14 years) was 3 years and the final height gain was 8.82 cm. In base-case 2, the length of treatment (assuming the child was aged

11 years) was 5 years and the final height gain was 3.48 cm. The drug dose (based on age- and sex-related weight at 50th percentile and not adjusted during puberty) was 0.33 mg/kg per week. The drop-out rate was 16% after the first year of treatment and 28% from monitoring after the first year of monitoring. Population data suggested that 68% were boys.

GH in PWS
In base-case 1, the length of treatment (assuming the child was aged 11 years) was 5 years (modellers' assumption). The height outcome was 1.4 HtSDS at 1 year and the drug dose was 0.233 mg/kg per week. In base-case 2, the length of treatment (assuming the child was aged 8 years) was 5 years (modellers' assumption). The height outcome was 1.0 HtSDS at 1 year and the drug dose was 0.35 mg/kg per week. A third base-case considered one study reporting on final height gained. In base-case 3, the length of treatment (assuming the child was aged 8 years) was 8 years. The final height gain was 10.38 cm (based on the distribution of final height in the general population) and the drug dose was 0.23 mg/kg per week. The drop-out rate was nil. The modellers assumed that 50% were boys.

GH in ISS
In base-case 1, the length of treatment (assuming the child was aged 10 years) was 6 years. The final height gain was 7.5 cm and the drug dose was 0.35 mg/kg per week (30 IU/m^2 per week; range: 0.35–0.70 mg/kg per week). In base-case 2, the length of treatment (assuming the child was aged 9 years) was 7 years. The final height gain was 2.68 cm and the drug dose was 0.233 mg/kg per week (20 IU/m^2 per week; range: 0.35–0.70 mg/kg per week). The drop-out rate was 29% after the first year of treatment and 30% from monitoring after the first year of monitoring. The modellers assumed that 60% were boys.

Measure of benefits used in the economic analysis
The measure of benefits used for four of the five conditions (GHD, TS, CRF, ISS) was centimetres gained. In the analysis of PWS, the measure of benefit was centimetres or HtSDS at 1 year.

Direct costs
All direct costs relating to the health service perspective were included in the analysis. These were for drugs, outpatient and day admissions, district nurse, X-ray, magnetic resonance imaging and laboratory tests. The prices were taken from the BNF (2001), the Personal Social Services Unit (University of Kent, 1999) and the Contracting Unit of Southampton University Hospitals Trust (2001). The resource quantities and the costs were reported separately. Discounting was conducted at 6.0% for costs and 1.5% for benefits (according to NICE guidelines).

Indirect costs
In line with the chosen perspective, the indirect costs were not reported.

Currency
UK pounds sterling (£).

Statistical analysis of costs
The data were deterministic.

Sensitivity analysis
A range of parameter values were tested using one- and two-way sensitivity analyses. Such parameters included the length of treatment (range: 1–13 years), final height effect (range: 10–300% of the effect from the base-case from trials), GH dose (by indication), GH cost (range: £15–25/mg) and the annual range of discounting costs (range: 0–12%). The analysis was conducted according to the two chosen base-cases and then subdivided into any of four scenarios that reflected important cost and practical factors that could influence successful treatment (full details were provided).

Estimated benefits used in the economic analysis
The incremental benefits were not reported separately [see the section 'Synthesis of costs and benefits' (p. 159)].

Cost results
The authors provided a breakdown of costs (model inputs) reflecting UK practice conditions that were common to all five conditions. Full details were provided in the report. The discount rate for the costs was 6.0%. The results were reported for each of the five conditions according to specific event pathways derived from expert consensus. Additional model parameters were applied separately according to resource use for diagnosis and treatment (full details in report) as follows.

GH and GHD
For base-case 1, the mean total cost of GH treatment was £55,712 and the mean incremental total cost per patient was £53,373. For base-case 2, these were £44,990 and £43,086, respectively. The mean costs of growth monitoring were £2339 for base-case 1 and £1904 for base-case 2.

GH and TS

For base-cases 1 and 2, the mean total cost of GH treatment was £62,621 and the mean incremental total cost per patient was £61,770. The mean cost of growth monitoring was £852.

GH and CRF

For base-case 1, the mean total cost of GH treatment was £54,617 and the mean incremental total cost per patient was £58,006. For base-case 2, these were £69,390 and £68,425, respectively. The mean cost of growth monitoring was £611 in base-case 1 and £965 in base-case 2.

GH and PWS

For base-case 1, the mean total cost of GH treatment was £56,663 and the mean incremental total cost per patient was £55,453. For base-case 2, these were £84,055 and £82,845, respectively, and for base-case 3, £70,882 and £69,263. The mean cost of growth monitoring was £1210 for base-cases 1 and 2 and £1620 for base-case 3.

GH and ISS

For base-case 1, the mean total cost of GH treatment was £70,674 and the mean incremental total cost per patient was £69,234. For base-case 2, these were £51,023 and £49,488, respectively. The mean cost of growth monitoring was £1440 in base-case 1 and £1535 in base-case 2.

Synthesis of costs and benefits

The results were reported as ICERs according to the five conditions. All units were reported as the cost per centimetre gained, except for PWS, which was the cost per HtSDS at 1 year.

GH and GHD

The incremental cost per unit gained was £6029 (range: 1385–11,853) for base-case 1 and £5708 (range: 1660–11,209) for base-case 2.

GH and TS

The incremental cost per unit gained was £15,997 (range: 4690–36,855) for base-case 1 and £17,429 (range: 5116–40,205) for base-case 2.

GH and CRF

The incremental cost per unit gained was £7403 (range: 2468–15,530) for base-case 1 and £24,093 (range: 7455–50,538) for base-case 2.

GH and PWS

The incremental cost per unit gained was £40,815 (range: 10,873–121,341) for base-case 1, £85,368 (range: 17,760–169,877) for base-case 2 and £7030 (range: 1466–20,897) for base-case 3.

GH and ISS

The incremental cost per unit gained was £13,498 (range: 4295–134,978) for base-case 1 and £27,202 (range: 8096–272,019) for base-case 2.

The sensitivity analysis results confirmed the sensitivity of the cost-effectiveness estimates. The most important factors were the measure of effectiveness, GH dose, and costs associated with length of treatment.

Authors' conclusions

GH treatment can potentially increase short-term growth and improve final height, but it is an expensive alternative to growth monitoring. The utility of small gains in these areas will be dependent on other factors, such as height in relation to peers and any psychological and health outcomes arising. More reliable evidence exists for short-term outcomes. Caution is required for final height results, owing to the limited number of very small, poorer quality studies measuring this outcome.

CRD commentary

Selection of comparators

Although no explicit justification was provided for the study of the GH drug, it would appear to represent current practice for the treatment of the five conditions in the UK setting. You should decide if this represents current practice in your own setting. The authors chose placebo (or no intervention, defined as growth monitoring) as the comparator for the intervention drug. This allowed the active value of the treatment to be evaluated.

Validity of estimate of measure of effectiveness

A systematic review was undertaken to derive the clinical effectiveness parameters. The epidemiological parameters were taken selectively from the literature. The review was supported by an extensive database search, and some appropriate steps to minimise bias were employed. However, the lack of independent data extraction and validity assessment potentially limits the reliability of the findings. A narrative synthesis was adopted to derive estimates of effectiveness and, although this was supported by a discussion of study weighting according to methodological rigour, the quality of the studies was generally poor. There was a great deal of heterogeneity amongst the included studies and this was appropriately identified in the analysis of the results. To address these weaknesses in the effectiveness data,

the authors undertook appropriate sensitivity analyses.

Validity of estimate of measure of benefit
The measures of benefit were centimetres and HtSDS gained. Final height was considered to be the more valid measure of effectiveness for the modelling, but very few good-quality studies were found to address this outcome.

Validity of estimate of costs
It appears that categories of costs relevant to the NHS and Personal Social Services were included in the analysis. The costs and the quantities were reported separately, thus enhancing the reproducibility of the study in other settings. The resource quantities and unit costs appear to have been reliably obtained from several published sources and from the authors' setting. A sensitivity analysis was appropriately conducted to reflect variations in GH dose, cost and annual discounting rate. The price year was reported, which will aid any future reflation exercise.

Other issues
The results are generalisable to the UK NHS. The authors stated that the results of this review require careful interpretation, given that the incorporated parameter values were not necessarily achievable in practice. The authors acknowledged other limitations of their study. These were related to the poor quality and heterogeneity of the included trials and the inability to establish more robust evidence for final height outcomes. The authors made significant efforts to obtain QoL data (utilities) from the literature and by considering a survey among TS sufferers and their parents. However, owing to various constraints, suitable data were not obtainable to formulate a CUA. The authors noted that a cost–utility approach would be more informative to decision-makers since height gain is a condition-specific outcome and a link to health benefits from a gain in final height is not established.

Implications of the study
The authors suggested that, given that only a minority of children with a licensed condition are currently receiving GH treatment in the UK, the budgetary impact of increased prescribing of GH treatment will require close examination. Specific recommendations are made for further large multi-centre RCTs focusing on final height and QoL (provided as utilities) as outcome measures.

4. Growth hormone in children (for growth hormone deficiency, Turner's syndrome, chronic renal failure and idiopathic short stature)

Anthony D, Stevens A. Report to the Development and Evaluation Committee. Report No. 57. Southampton: Development and Evaluation Committee: 1996:1–32.

Health technology
The study examined the use of GH treatment in children with GHD, TS, CRF and ISS.

Disease
Urological and male genital diseases; female genital diseases and pregnancy complications; neonatal diseases and abnormalities; endocrine diseases; musculoskeletal diseases; nervous system diseases.

Type of intervention
Treatment.

Hypothesis/study question
The objective of the study was to carry out a CUA comparing the use of GH treatment with no treatment in children with the specified conditions. The treatment drug appears to have been chosen on the basis of its current use as a standard licensed treatment for GHD, CRF and TS in the UK. The authors did not state the perspective for the economic study, but it appears to have been conducted from the perspective of the NHS.

Economic study type
CUA.

Study population
The study population comprised primarily pre-pubertal boys and girls (age range: 1–13 years) with GHD, CRF, TS or ISS. Children with skeletal dysplasia (achondroplasia, PWS, Noonon, Russell–Silver) and those with intrauterine growth retardation were excluded from the review.

Setting
The setting was not stated, but it was likely to have been primary care or the community. The economic study was carried out in the UK.

Dates to which data relate
The effectiveness and resource use data were derived from published studies dating from 1987 to 1995. The prices were derived from the BNF 1995.

Source of effectiveness data

The effectiveness data were derived from a systematic review of the literature.

Outcomes assessed in the review

Height gain was the primary outcome measurement in this review. A variety of measures were employed in the reporting of this outcome, such as changes in SD, height change and height velocity per year. Other outcomes included psychological, academic and behavioural measures. Adverse events were also considered.

Study designs and other criteria for inclusion in the review

The inclusion criteria were not stated. The included studies comprised RCTs, a controlled trial and case series.

Sources searched to identify primary studies

The authors searched MEDLINE (from 1985 to May 1996), HealthPLAN (from 1985 to December 1995) and GEARS, and also consulted a local expert in order to identify primary studies

Criteria used to ensure the validity of primary studies

The grading and reporting of studies as 'best quality' was based on several criteria. Specifically, the type of trial, the number of patients, patient accountability, focus of the trial and relevance of the outcomes.

Methods used to judge relevance and validity and for extracting data

Not reported.

Number of primary studies included

Twenty-eight studies were included in the review. There were eight RCTs, one controlled trial and 19 case series.

Method of combination of primary studies

The studies were combined in narrative form, according to the respective clinical condition.

Investigation of differences between studies

The authors provided a narrative account of differences between the primary studies. In particular, the difficulty in comparing height gain results from studies using different outcome measures was identified.

Results of the review

The authors reported that height improvements were observed in all treated children (regardless of disorder) when compared with their expected gain without treatment. The results were reported selectively by reference to five studies that were claimed to provide the 'best quality' evidence as follows.

In terms of height gain, children (including pubertal patients) treated for GHD achieved a mean of 2.6 SDs over an average of 5 years. Their initial height ranged from –6.2 SD to –2.9 SD and their final height from –1 SD to –3.3 SDs.

Children treated for TS achieved a mean gain of 8.1 cm (3.2 in) over original projections over a 5-year treatment period. Their original height was 143.8 cm (4 ft 8 in) and the mean final height was 151.9 cm (5 ft).

Children treated for CRF achieved a mean gain of 1.48 SDs in comparison with no treatment over a 2-year period. The initial height of the children was –2.94 SDs (control group –2.82 SDs) and the final height was –1.55 SDS (control group –2.91 SDs).

Treated children with ISS gained a mean of 1.1 SDs over a 3-year period. Their initial height was –2.7 SDs and their final height was –1.6 SDs.

No significant differences were found for psychological benefits when a treated group (of children with the entire range of disorders) was compared with a control group of ISS patients.

Other studies of psychological outcomes reported conflicting results.

Adverse events were not viewed to be a substantial consequence of the treatment

Measure of benefits used in the economic analysis

The measure of benefit used was QALYs. The IHQL was used in the valuation of utilities for children with the target disorders. The authors stated that the greatest QALY gain from treatment (within the treatment period) would be approximately 0.1 and the least gain would be a loss of QALY of 0.1. Given that treatment is given from an average age of 10 years, for approximately 4–6 years, the authors proposed that children would gain 0.5 QALYs in the best scenario and would lose 0.5 QALYs in the worst scenario.

161

Direct costs

The UK NHS cost was included in the analysis. The projected annual cost of growth hormone treatment for a 9-year-old child with GHD was calculated from a published source, using a resource use dosage rate of 15 units/m^2 per week at 1995 prices. Children with TS, CRF and ISS were reported to require double this dose, and all patients would require increased dosages at the onset of puberty. The resource quantities and the costs were not reported separately, but the cost and resource use per patient were recorded. Discounting was not reported.

Indirect costs

In line with the perspective of the economic analysis, the indirect costs were not considered.

Currency

UK pounds sterling (£).

Statistical analysis of costs

The cost data were deterministic.

Sensitivity analysis

There was no formal sensitivity analysis. However, the potential cost of treating all children with GHD and CRF, along with the prospect of supplying treatment on demand to all children with short stature in the south and west regions of the UK, was explored.

Estimated benefits used in the economic analysis

The authors reported that the best scenario in terms of benefit was 1.5 more QALYs when assuming 15 years of benefit (5 years of which were in the treatment period), and 5.5 more QALYS when assuming 55 years of benefit (5 years of which were in the treatment period).

Cost results

The annual average cost of treating a 9-year-old child with GHD was reported to be approximately £7000, based on the lower dose regimen.

The approximate cost for a 9-year-old child with TS, CRF or ISS would be £14,000.

Synthesis of costs and benefits

The (best scenario) cost per QALY for children with GHD was reported to be between £5700 and £20,800.

The cost per QALY for children with TS, CRF and ISS was between £11,400 and £41,700 (assuming between 15 and 55 years of benefit).

Authors' conclusions

GH treatment should be recommended for children with short stature associated with GHD, TS and CRF. There is currently insufficient evidence to support the use of this treatment in children with ISS. The few studies examining psychological benefits of the treatment presented conflicting results.

CRD commentary
Selection of comparators

Although no explicit justification was provided for the study of the GH drug, it would appear to represent the standard licensed treatment for GHD, CRF and TS in the UK. The comparison with no treatment permitted the evaluation of the active value of the treatment drug. You should decide if this represents a widely used technology in your own setting.

Validity of estimate of measure of effectiveness

Although the authors cited a search strategy for the review of the literature, it was unclear whether the remainder of the process was conducted systematically so as to minimise potential biases. The 'best quality' evidence was derived on the basis of grading the studies according to specific criteria. The usefulness of these criteria in terms of ascertaining validity is likely to be limited. Although some RCTs were included, the majority of the studies were of weaker designs. Given that some studies lacked a control group, the authors acknowledged difficulties in assessing a direct causal relationship between treatment and outcome in some cases. In addition, the authors also acknowledged that the comparison of height gain results was hampered by differences in the outcome measures used across the studies. All of the above represent substantial threats to the reliability of the results.

Validity of estimate of measure of benefit

The measure of benefit used for the utility analysis was the IHQL, based on an estimate. Although no further references were cited in relation to the validity of this measure, the authors quantified the methods and parameters employed in the analysis of the target disorders.

Validity of estimate of costs

The cost analysis was based on a conservative estimate, which included only the average cost of GHD treatment from the perspective of the NHS. Monitoring costs were not included, nor was the extra cost of GH treatment at puberty. An appropriate extrapolation of the costs (based on the higher projected resource use) was presented

for patients with CRF and TS. The resource quantities and the costs were not reported separately, although the costs and resource use per patient (based on a reliable source) were recorded. This might not allow the analysis to be easily reworked for other settings. The price year was reported, which will aid any future reflation exercise. However, there was no sensitivity analysis to explore any variation in the costs. Discounting was not reported, although it was potentially relevant given the time frame of some of the studies.

Other issues
The results are generalisable to the NHS. However, the authors did not directly compare their findings with other studies, nor did they directly address the issue of generalisability to other settings beyond the south and west regions.

Implications of the study
The authors suggested that, given the high cost of GH treatment, future good-quality controlled trials with longer follow-up are needed to determine reliably the benefits of such treatment. In addition, more research is required to address the motivation for, and expected benefits arising from, the use of GH treatment.

Health Technology Assessment Programme

Director,
Professor Tom Walley,
Director, NHS HTA Programme,
Department of Pharmacology &
Therapeutics,
University of Liverpool

Deputy Director,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield,
School of Health and Related
Research

Prioritisation Strategy Group

Members

Chair,
Professor Tom Walley,
Director, NHS HTA Programme,
Department of Pharmacology &
Therapeutics,
University of Liverpool

Professor Bruce Campbell,
Consultant Vascular & General
Surgeon, Royal Devon & Exeter
Hospital

Professor Robin E Ferner,
Consultant Physician and
Director, West Midlands Centre
for Adverse Drug Reactions,
City Hospital NHS Trust,
Birmingham

Dr Edmund Jessop, Medical
Adviser, National Specialist,
Commissioning Advisory Group
(NSCAG), Department of
Health, London

Professor Jon Nicholl, Director,
Medical Care Research Unit,
University of Sheffield,
School of Health and
Related Research

Dr Ron Zimmern, Director,
Public Health Genetics Unit,
Strangeways Research
Laboratories, Cambridge

HTA Commissioning Board

Members

Programme Director,
Professor Tom Walley,
Director, NHS HTA Programme,
Department of Pharmacology &
Therapeutics,
University of Liverpool

Chair,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield,
School of Health and Related
Research

Deputy Chair,
Dr Andrew Farmer,
University Lecturer in General
Practice, Department of
Primary Health Care,
University of Oxford

Dr Jeffrey Aronson,
Reader in Clinical
Pharmacology, Department of
Clinical Pharmacology,
Radcliffe Infirmary, Oxford

Professor Deborah Ashby,
Professor of Medical Statistics,
Department of Environmental
and Preventative Medicine,
Queen Mary University of
London

Professor Ann Bowling,
Professor of Health Services
Research, Primary Care and
Population Studies,
University College London

Professor John Cairns,
Professor of Health Economics,
Public Health Policy,
London School of Hygiene
and Tropical Medicine,
London

Professor Nicky Cullum,
Director of Centre for Evidence
Based Nursing, Department of
Health Sciences, University of
York

Professor Jon Deeks,
Professor of Health Statistics,
University of Birmingham

Professor Jenny Donovan,
Professor of Social Medicine,
Department of Social Medicine,
University of Bristol

Professor Freddie Hamdy,
Professor of Urology,
University of Sheffield

Professor Allan House,
Professor of Liaison Psychiatry,
University of Leeds

Professor Sallie Lamb, Director,
Warwick Clinical Trials Unit,
University of Warwick

Professor Stuart Logan,
Director of Health & Social
Care Research, The Peninsula
Medical School, Universities of
Exeter & Plymouth

Professor Miranda Mugford,
Professor of Health Economics,
University of East Anglia

Dr Linda Patterson,
Consultant Physician,
Department of Medicine,
Burnley General Hospital

Professor Ian Roberts,
Professor of Epidemiology &
Public Health, Intervention
Research Unit, London School
of Hygiene and Tropical
Medicine

Professor Mark Sculpher,
Professor of Health Economics,
Centre for Health Economics,
Institute for Research in the
Social Services,
University of York

Professor Kate Thomas,
Professor of Complementary
and Alternative Medicine,
University of Leeds

Professor David John Torgerson,
Director of York Trial Unit,
Department of Health Sciences,
University of York

Professor Hywel Williams,
Professor of
Dermato-Epidemiology,
University of Nottingham

Current and past membership details of all HTA 'committees' are available from the HTA website (www.hta.ac.uk)

Diagnostic Technologies & Screening Panel

Members

Chair,
Dr Ron Zimmern, Director of the Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

Ms Norma Armston, Freelance Consumer Advocate, Bolton

Professor Max Bachmann, Professor of Health Care Interfaces, Department of Health Policy and Practice, University of East Anglia

Professor Rudy Bilous Professor of Clinical Medicine & Consultant Physician, The Academic Centre, South Tees Hospitals NHS Trust

Ms Dea Birkett, Service User Representative, London

Dr Paul Cockcroft, Consultant Medical Microbiologist and Clinical Director of Pathology, Department of Clinical Microbiology, St Mary's Hospital, Portsmouth

Professor Adrian K Dixon, Professor of Radiology, University Department of Radiology, University of Cambridge Clinical School

Dr David Elliman, Consultant in Community Child Health, Islington PCT & Great Ormond Street Hospital, London

Professor Glyn Elwyn, Research Chair, Centre for Health Sciences Research, Cardiff University, Department of General Practice, Cardiff

Professor Paul Glasziou, Director, Centre for Evidence-Based Practice, University of Oxford

Dr Jennifer J Kurinczuk, Consultant Clinical Epidemiologist, National Perinatal Epidemiology Unit, Oxford

Dr Susanne M Ludgate, Clinical Director, Medicines & Healthcare Products Regulatory Agency, London

Mr Stephen Pilling, Director, Centre for Outcomes, Research & Effectiveness, Joint Director, National Collaborating Centre for Mental Health, University College London

Mrs Una Rennard, Service User Representative, Oxford

Dr Phil Shackley, Senior Lecturer in Health Economics, Academic Vascular Unit, University of Sheffield

Dr Margaret Somerville, Director of Public Health Learning, Peninsula Medical School, University of Plymouth

Dr Graham Taylor, Scientific Director & Senior Lecturer, Regional DNA Laboratory, The Leeds Teaching Hospitals

Professor Lindsay Wilson Turnbull, Scientific Director, Centre for MR Investigations & YCR Professor of Radiology, University of Hull

Professor Martin J Whittle, Clinical Co-director, National Co-ordinating Centre for Women's and Childhealth

Dr Dennis Wright, Consultant Biochemist & Clinical Director, The North West London Hospitals NHS Trust, Middlesex

Pharmaceuticals Panel

Members

Chair,
Professor Robin Ferner, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham

Ms Anne Baileff, Consultant Nurse in First Contact Care, Southampton City Primary Care Trust, University of Southampton

Professor Imti Choonara, Professor in Child Health, Academic Division of Child Health, University of Nottingham

Professor John Geddes, Professor of Epidemiological Psychiatry, University of Oxford

Mrs Barbara Greggains, Non-Executive Director, Greggains Management Ltd

Dr Bill Gutteridge, Medical Adviser, National Specialist Commissioning Advisory Group (NSCAG), London

Mrs Sharon Hart, Consultant Pharmaceutical Adviser, Reading

Dr Jonathan Karnon, Senior Research Fellow, Health Economics and Decision Science, University of Sheffield

Dr Yoon Loke, Senior Lecturer in Clinical Pharmacology, University of East Anglia

Ms Barbara Meredith, Lay Member, Epsom

Dr Andrew Prentice, Senior Lecturer and Consultant Obstetrician & Gynaecologist, Department of Obstetrics & Gynaecology, University of Cambridge

Dr Frances Rotblat, CPMP Delegate, Medicines & Healthcare Products Regulatory Agency, London

Dr Martin Shelly, General Practitioner, Leeds

Mrs Katrina Simister, Assistant Director New Medicines, National Prescribing Centre, Liverpool

Dr Richard Tiner, Medical Director, Medical Department, Association of the British Pharmaceutical Industry, London

Current and past membership details of all HTA 'committees' are available from the HTA website (www.hta.ac.uk)

Therapeutic Procedures Panel

Members

Chair,
Professor Bruce Campbell,
Consultant Vascular and
General Surgeon, Department
of Surgery, Royal Devon &
Exeter Hospital

Dr Mahmood Adil, Deputy
Regional Director of Public
Health, Department of Health,
Manchester

Dr Aileen Clarke,
Consultant in Public Health,
Public Health Resource Unit,
Oxford

Professor Matthew Cooke,
Professor of Emergency
Medicine, Warwick Emergency
Care and Rehabilitation,
University of Warwick

Mr Mark Emberton, Senior
Lecturer in Oncological
Urology, Institute of Urology,
University College Hospital

Professor Paul Gregg,
Professor of Orthopaedic
Surgical Science, Department of
General Practice and Primary
Care, South Tees Hospital NHS
Trust, Middlesbrough

Ms Maryann L Hardy,
Lecturer, Division of
Radiography, University of
Bradford

Dr Simon de Lusignan,
Senior Lecturer, Primary Care
Informatics, Department of
Community Health Sciences,
St George's Hospital Medical
School, London

Dr Peter Martin, Consultant
Neurologist, Addenbrooke's
Hospital, Cambridge

Professor Neil McIntosh,
Edward Clark Professor of Child
Life & Health, Department of
Child Life & Health, University
of Edinburgh

Professor Jim Neilson,
Professor of Obstetrics and
Gynaecology, Department of
Obstetrics and Gynaecology,
University of Liverpool

Dr John C Pounsford,
Consultant Physician,
Directorate of Medical Services,
North Bristol NHS Trust

Dr Karen Roberts, Nurse
Consultant, Queen Elizabeth
Hospital, Gateshead

Dr Vimal Sharma, Consultant
Psychiatrist/Hon. Senior
Lecturer, Mental Health
Resource Centre, Cheshire and
Wirral Partnership NHS Trust,
Wallasey

Professor Scott Weich,
Professor of Psychiatry,
Division of Health in the
Community, University of
Warwick

Disease Prevention Panel

Members

Chair,
Dr Edmund Jessop, Medical
Adviser, National Specialist
Commissioning Advisory Group
(NSCAG), London

Mrs Sheila Clark, Chief
Executive, St James's Hospital,
Portsmouth

Mr Richard Copeland,
Lead Pharmacist: Clinical
Economy/Interface,
Wansbeck General Hospital,
Northumberland

Dr Elizabeth Fellow-Smith,
Medical Director,
West London Mental Health
Trust, Middlesex

Mr Ian Flack, Director PPI
Forum Support, Council of
Ethnic Minority Voluntary
Sector Organisations,
Stratford

Dr John Jackson,
General Practitioner,
Newcastle upon Tyne

Mrs Veronica James, Chief
Officer, Horsham District Age
Concern, Horsham

Professor Mike Kelly,
Director, Centre for Public
Health Excellence,
National Institute for Health
and Clinical Excellence,
London

Professor Yi Mien Koh,
Director of Public Health and
Medical Director, London
NHS (North West London
Strategic Health Authority),
London

Ms Jeanett Martin,
Director of Clinical Leadership
& Quality, Lewisham PCT,
London

Dr Chris McCall, General
Practitioner, Dorset

Dr David Pencheon, Director,
Eastern Region Public Health
Observatory, Cambridge

Dr Ken Stein, Senior Clinical
Lecturer in Public Health,
Director, Peninsula Technology
Assessment Group,
University of Exeter,
Exeter

Dr Carol Tannahill, Director,
Glasgow Centre for Population
Health, Glasgow

Professor Margaret Thorogood,
Professor of Epidemiology,
University of Warwick,
Coventry

Dr Ewan Wilkinson,
Consultant in Public Health,
Royal Liverpool University
Hospital, Liverpool

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in
Medicine, Centre for Statistics
in Medicine, University of
Oxford

Professor John Bond,
Director, Centre for Health
Services Research, University of
Newcastle upon Tyne, School of
Population & Health Sciences,
Newcastle upon Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery,
Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway
Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive,
Regulation and Improvement
Authority, Belfast

Ms Tracy Bury,
Project Manager, World
Confederation for Physical
Therapy, London

Professor Iain T Cameron,
Professor of Obstetrics and
Gynaecology and Head of the
School of Medicine,
University of Southampton

Dr Christine Clark,
Medical Writer & Consultant
Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing & Head of
Research, School of Health
Sciences, University of
Birmingham, Edgbaston,
Birmingham

Professor Barry Cookson,
Director, Laboratory of
Healthcare Associated Infection,
Health Protection Agency,
London

Dr Carl Counsell, Clinical
Senior Lecturer in Neurology,
Department of Medicine &
Therapeutics, University of
Aberdeen

Professor Howard Cuckle,
Professor of Reproductive
Epidemiology, Department of
Paediatrics, Obstetrics &
Gynaecology, University of
Leeds

Dr Katherine Darton,
Information Unit, MIND –
The Mental Health Charity,
London

Professor Carol Dezateux,
Professor of Paediatric
Epidemiology, London

Dr Keith Dodd, Consultant
Paediatrician, Derby

Mr John Dunning,
Consultant Cardiothoracic
Surgeon, Cardiothoracic
Surgical Unit, Papworth
Hospital NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon,
Gloucestershire Royal Hospital,
Gloucester

Professor Martin Eccles,
Professor of Clinical
Effectiveness, Centre for Health
Services Research, University of
Newcastle upon Tyne

Professor Pam Enderby,
Professor of Community
Rehabilitation, Institute of
General Practice and Primary
Care, University of Sheffield

Professor Gene Feder, Professor
of Primary Care Research &
Development, Centre for Health
Sciences, Barts & The London
Queen Mary's School of
Medicine & Dentistry, London

Mr Leonard R Fenwick,
Chief Executive, Newcastle
upon Tyne Hospitals NHS Trust

Mrs Gillian Fletcher,
Antenatal Teacher & Tutor and
President, National Childbirth
Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine,
Department of Medicine,
University of Birmingham,
Queen Elizabeth Hospital,
Edgbaston, Birmingham

Dr Neville Goodman,
Consultant Anaesthetist,
Southmead Hospital, Bristol

Professor Robert E Hawkins,
CRC Professor and Director of
Medical Oncology, Christie CRC
Research Centre, Christie
Hospital NHS Trust, Manchester

Professor Allen Hutchinson,
Director of Public Health &
Deputy Dean of ScHARR,
Department of Public Health,
University of Sheffield

Professor Peter Jones, Professor
of Psychiatry, University of
Cambridge, Cambridge

Professor Stan Kaye, Cancer
Research UK Professor of
Medical Oncology, Section of
Medicine, Royal Marsden
Hospital & Institute of Cancer
Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch
& Ptnrs), The Health Centre,
Thame

Dr Donna Lamping,
Research Degrees Programme
Director & Reader in Psychology,
Health Services Research Unit,
London School of Hygiene and
Tropical Medicine, London

Mr George Levvy,
Chief Executive, Motor
Neurone Disease Association,
Northampton

Professor James Lindesay,
Professor of Psychiatry for the
Elderly, University of Leicester,
Leicester General Hospital

Professor Julian Little,
Professor of Human Genome
Epidemiology, Department of
Epidemiology & Community
Medicine, University of Ottawa

Professor Rajan Madhok,
Consultant in Public Health,
South Manchester Primary
Care Trust, Manchester

Professor Alexander Markham,
Director, Molecular Medicine
Unit, St James's University
Hospital, Leeds

Professor Alistaire McGuire,
Professor of Health Economics,
London School of Economics

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore, Public
Health Director, Southampton
City Primary Care Trust,
Southampton

Dr Sue Moss, Associate Director,
Cancer Screening Evaluation
Unit, Institute of Cancer
Research, Sutton

Mrs Julietta Patnick,
Director, NHS Cancer Screening
Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Professor Chris Price,
Visiting Professor in Clinical
Biochemistry, University of
Oxford

Professor William Rosenberg,
Professor of Hepatology and
Consultant Physician, University
of Southampton, Southampton

Professor Peter Sandercock,
Professor of Medical Neurology,
Department of Clinical
Neurosciences, University of
Edinburgh

Dr Susan Schonfield, Consultant
in Public Health, Hillingdon
PCT, Middlesex

Dr Eamonn Sheridan,
Consultant in Clinical Genetics,
Genetics Department,
St James's University Hospital,
Leeds

Professor Sarah Stewart-Brown,
Professor of Public Health,
University of Warwick,
Division of Health in the
Community Warwick Medical
School, LWMS, Coventry

Professor Ala Szczepura,
Professor of Health Service
Research, Centre for Health
Services Studies, University of
Warwick

Dr Ross Taylor,
Senior Lecturer, Department of
General Practice and Primary
Care, University of Aberdeen

Mrs Joan Webster,
Consumer member, HTA –
Expert Advisory Network

Current and past membership details of all HTA 'committees' are available from the HTA website (www.hta.ac.uk)

Feedback

The HTA Programme and the authors would like to know
your views about this report.

The Correspondence Page on the HTA website
(http://www.hta.ac.uk) is a convenient way to publish
your comments. If you prefer, you can send your comments
to the address below, telling us whether you would like
us to transfer them to the website.

We look forward to hearing from you.