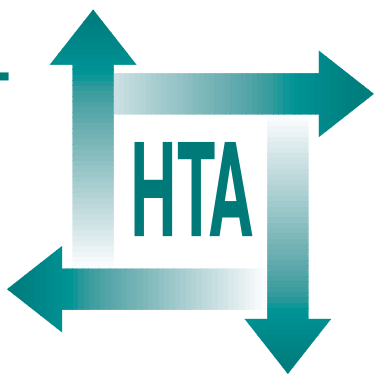
Health Technology Assessment 2002; Vol. 6: No. 11

Screening for gestational diabetes: a systematic review and economic evaluation

DA Scott
E Loveman
L McIntyre
N Waugh

Health Technology Assessment
NHS R&D HTA Programme



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
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
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.

Screening for gestational diabetes: a systematic review and economic evaluation

DA Scott[1]
E Loveman[1]
L McIntyre[2]
N Waugh[1*]

[1] Southampton Health Technology Assessments Centre, University of Southampton, UK
[2] Independent researcher, Stromness, Isle of Orkney, UK

[*] Corresponding author
(From November 2002, Dr Waugh will be based at Department of Public Health, University of Aberdeen, Medical School Buildings, Aberdeen)

Declared competing interests of the authors: none

Published October 2002

This report should be referenced as follows:

Scott DA, Loveman E, McIntyre L, Waugh N. Screening for gestational diabetes: a systematic review and economic evaluation. *Health Technol Assess* 2002;**6**(11).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA website (see opposite).

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

The research reported in this monograph was commissioned by the HTA Programme on behalf of the National Institute for Clinical Excellence (NICE). Technology assessment reports are completed in a limited time to inform the appraisal and guidance development processes managed by NICE. The review brings together evidence on key aspects of the use of the technology concerned. However, appraisals and guidance produced by NICE are informed by a wide range of sources.

The research reported in this monograph was funded as project number 99/09/50.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme, NICE or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned 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 editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

T

Contents

List of abbreviations

ACOG	American College of Obstetricians and Gynecologists
ADA	American Diabetes Association
ARHT	Aberdeen Royal Hospitals Trust
BDA	British Diabetic Association
BG	blood glucose
BMI	body mass index
C&C	Carpenter & Coustan modified GTT criteria
CBG	capillary blood glucose
CI	confidence interval
CPG	capillary plasma glucose
CS	Caesarean section
DM	diabetes mellitus
DTHT	Dundee Teaching Hospitals Trust
FATD	fetal abdominal transverse diameter
FBG	fasting blood glucose
FPG	fasting plasma glucose
GCT	glucose challenge test
GDM	gestational diabetes mellitus
GTT	glucose tolerance test
HbA_{1C}	glycosylated haemoglobin
IGT	impaired glucose tolerance
LGA	large for gestational age
MIMS	*Monthly Index of Medical Specialties*
N/A	not available
NDDG	National Diabetes Data Group (USA)
NPV	negative predictive value
NSC	National Screening Committee
N/S	not specified/stated
NSF	National Service Framework
OGTT	oral glucose tolerance test
OR	odds ratio
PET	pre-eclampsic toxaemia
PG	plasma glucose
PPV	positive predictive value
QALY	quality-adjusted life-year
RBG	random blood glucose
RCT	randomised controlled trial
RF	risk factor
ROC	receiver-operator characteristic
RPG	random plasma glucose
RR	relative risk
SCBU	special care baby unit
SHPIC	Scottish Health Purchasing Information Centre
SOCG	Society of Obstetricians and Gynecologists of Canada
VBG	venous blood glucose
VPG	venous plasma glucose
WHO	World Health Organization

Executive summary

Background

Screening for gestational diabetes mellitus (GDM) has been controversial, with some expert bodies advising universal screening, others selective screening, and yet others advising against screening at all. This has partly been a result of debate about the definition of GDM, and partly because of the profusion of different tests available, both for screening and definite diagnosis. In the UK, there is no national policy on screening, and a variety of practices exist in different parts of the country. There have also been doubts about the treatment of GDM, and particularly about management of minor degrees of glucose elevation, which are better described as glucose intolerance rather than true diabetes.

Objectives

To provide an updated review of current knowledge, to clarify research needs, and to assist with policy making in the interim, pending future research.

Methods

A literature review was carried out, with a particular focus on screening methods and costs, and an appraisal of screening for GDM against the criteria for assessing screening programmes used by the UK National Screening Committee (NSC).

Results

There is still debate about what is meant by GDM – the threshold for diagnosis is not soundly based; the terms GDM and impaired glucose tolerance are not used in a standard fashion in pregnancy; there is almost certainly a continuum of risk to the baby, rather than there being separate normal and abnormal groups; and the key risk factor in most women may be maternal overweight, with glucose intolerance being an associate of that. In addition there are some rare genetic conditions, which affect a few women, such as glucokinase and hepatic nuclear factor disorders.

GDM is usually defined according to divergence from normal glucose levels, but glucose levels are usually raised in pregnancy, and so diagnosis by normal levels in non-pregnant women may misclassify many normal pregnant women as abnormal. This may lead to anxiety and the inconvenience of extra investigations and 'disease' care. The Caesarean section rate appears to be increased by the diagnosis alone.

Ideally, the condition should be defined by the incidence of adverse effects. However, the most common reported complication of GDM is 'macrosomia' in the baby. This is usually defined by arbitrary weight cut-offs (usually a birth weight of 4000 g, but sometimes 4500 g), but such neat thresholds fail to distinguish between larger than average healthy babies and those that have the abnormal growth patterns associated with high insulin levels in the womb.

Screening for GDM fails to meet some of the NSC criteria.

A number of screening tests have been used but some, such as glycosylated haemoglobin and fructosamine, have proved unsatisfactory and can be discarded. Others, such as urine testing or random blood glucose, are far from satisfactory, although they may be cheap to do. There is marked international variation. Risk factors such as weight, age and family history are useful for selective screening but some patients with GDM would be missed. Fasting plasma glucose (FPG) is convenient and reliable, but some pregnant women have normal fasting levels but raised levels of glucose after meals, and would be missed by screening based on FPG alone. Glucose challenge tests (GCTs) are based on glucose levels after a glucose drink, but also have shortcomings. The definitive diagnosis is usually by oral glucose tolerance test (OGTT), but the glucose load and timing vary in different countries; taking a 75 g glucose load is unnatural, makes some women sick, and the reproducibility of the test is poor. More natural methods such as test meals have been used, but not widely.

Conclusions

Interim conclusions

There are clearly some women whose glucose levels rise sufficiently in pregnancy to cause harm to their babies. However, there are also many women with lower levels of glucose intolerance whose babies are not at risk, but who may suffer anxiety and inconvenience as a result of being classed as abnormal. On balance, the present evidence suggests that we should not have universal screening, but a highly selective policy, based on age and overweight.

The best test at present, for those deemed to need testing, is probably the GCT, preferably combined with an FPG. The benefits of a follow-up OGTT are doubtful.

Recommendations for research

The main research needs appear to be:

1. There is a need to better define the 'disease' by documenting the frequency of adverse events, best done by population-based epidemiological surveys. These should include ethnic groups, as risks appear to vary, although this may be partly due to the prevalence of overweight. This work would relate outcomes of pregnancy to maternal blood glucose and other factors, to determine the level of glucose at which outcomes worsened significantly. Data on other factors such as overweight would be used to determine whether glucose intolerance was an independent cause, and if so at what level.
2. If such research showed that there was a continuum of risk, rather than there being distinct normal and abnormal groups, economic analysis should examine the cost-effectiveness of intervention at different levels.
3. Trials of the marginal costs and benefits of different screening tests – for example, FPG versus GCT – and whether if these are positive, a follow-up OGTT is necessary, because it is far from being a gold standard.
4. Trials of intervention in key groups, such as those with normal FPG but elevated postprandial levels.
5. After all these, further analysis of the cost-effectiveness of screening – should it be done, and if so, how selective should it be?

Some research is under way overseas, and it is recommended that the results of the two main trials, the Hyperglycaemia and Pregnancy Outcome Study (HAPO) and the ACHOIS trial (a collaborative trial of treatment for screen-detected GDM) be awaited before further research is commissioned by the Health Technology Assessment Programme.

Chapter 1
Background

Introduction

There were two reasons for carrying out this review. Firstly, screening for gestational diabetes mellitus (GDM) has been given high priority on several occasions by the Population Screening Panel (now part of the merged Diagnostic Technologies and Screening Panel) of the Health Technology Assessment Programme. However, it has not been possible to commission the primary research considered necessary. This might have been partly because of a lack of focus in the commissioning brief, and it was felt that a systematic review and economic evaluation might help clarify the questions that now need to be addressed. However, it is also likely that the number of bids to carry out the research was low, because many clinicians feel that screening has been shown to be necessary, and that doing a trial with a no-screening arm might be unethical. Another reason proposed trials have not been funded, and in fact the reason for the breakdown of one of the outline proposals, is size of population required. To have 100 patients in each of three arms (e.g. no screen, glucose challenge test (GCT), fasting plasma glucose (FPG)) would need 15,000 pregnant women based on a 2% incidence rate. Thus a multicentre trial is essential.

It could be argued that diffusion of screening has already occurred, and that it is too late to do the research. However, a UK survey in 1999 showed that there was still considerable variation in screening practices.[1]

The second reason for this study was that work was under way on the National Service Framework (NSF) for diabetes, and it was hoped that a systematic review and economic evaluation might assist with the work on the draft NSF and then on the implications of implementing its recommendations.

Following an initial introduction to GDM and the current issues surrounding its diagnosis (chapter 1), this report focuses on whether the evidence for GDM screening is compatible with the criteria used by the National Screening Committee (NSC) for a screening test (chapter 2). The report then reviews the evidence of effectiveness (chapter 3), and provides an economic appraisal (chapter 4). Chapter 5 discusses the weaknesses of the research evidence, and gives recommendations for future research and interim conclusions for clinical practice.

Current issues

Screening for GDM has long been controversial.

The American Diabetes Association (ADA) concluded in 1996 that selective screening was inadequate and that all pregnant women should be screened. In 1997 and 1998 they revised these recommendations in favour of selectively screening women satisfying at least one of the following criteria: aged 25 or over; aged under 25 and with body mass index (BMI) over 27; a family history of diabetes; and ethnic groups with a high prevalence of diabetes.[2,3] The American College of Obstetricians and Gynecologists (ACOG) recommended screening pregnant women aged 30 and over, younger if they have historic risk factors.[4]

The US Preventive Services Task Force noted that "no properly controlled trial has examined the benefit of universal screening or selective screening compared to routine care without screening".[5] The Canadian Task Force on the Periodic Health Examination carried out what appeared to be a systematic review in 1992 and concluded that universal screening could not be supported, but that "women have various degrees of glucose intolerance and that a certain proportion will have adverse outcomes and could benefit from screening".[6] However, critics argued that the review failed to include a number of studies on adverse outcomes to the neonate[7] or to fully assess long-term consequences for the mothers.[8]

Critical reviews by Jarrett,[9] Stephenson[10] and Goer[11] concluded that there was no case for routine screening, but these were not fully systematic reviews by the Oxman and Guyatt criteria.[12] A review by Dornhorst and Beard[13] noted that "the obstetric benefits from screening are poorly validated" but seemed more convinced of the benefits arising from prevention of later type 2 diabetes.

A later review by Dornhorst and Chan[14] noted that:

> "In today's world of evidence-based medicine, audit and cost-constraints, the case for screening and treating women with GDM is severely hampered by the lack of a clear definition, agreed diagnostic criteria, and evidence that improving mild disturbances of maternal glycaemia improves pregnancy outcome."

There have been various calls for randomised controlled trials (RCTs) before screening for GDM is introduced.[15,16]

There are probably six main issues:

1. Should there be screening for GDM at all? Does it meet the World Health Organization (WHO) criteria for a screening programme?
2. If there is to be screening, should it be universal or selective?
3. Which screening test should be used? In the USA, the 50-g GCT is the one most often used for screening. Other tests have included questionnaires, urine testing, random blood glucose (RBG), fasting blood glucose (FBG), test meals, fructosamine and glycosylated haemoglobin. The questionnaires are used to collect personal and family history, and this information may then be used as part of two-stage screening – firstly by risk factors, then by biochemical tests such as the 50-g GCT. The decision depends not only on sensitivity and specificity, but also on cost-effectiveness.
4. What is the gold standard definitive test in those who are screen-positive? The oral glucose tolerance test (OGTT) is the choice, but using 75 g or 100 g glucose loads depending on which side of the Atlantic it is used. But there is also debate about the need for this – could the 50-g GCT be sufficient as a guide to treatment on its own?
5. What cut-off levels should be used for the diagnosis? In pregnancy, there is a physiological state of glucose intolerance, with insulin resistance in the third trimester. The cut-offs in the non-pregnant state may therefore be inappropriate.
6. For those treated, what should the target glucose level be, and should fasting or postprandial glucose or both be used?

The aim of a screening programme should be to reduce ill health in mother and/or child. There are three groups of possible benefits:

- to the baby;
- to the mother in pregnancy;
- to the mother in later life.

What is GDM?

For a person to be diagnosed as having a disease, two things are usually necessary. Firstly, there should be a clear distinction between normality and disease. Secondly, the disease should cause harm. The harm usually gives a guide to defining the condition. For example, the ADA definition of diabetes is based on the level above which the risk of microvascular disease rises steeply.[2]

Definitions

GDM is defined as "carbohydrate intolerance of variable severity with onset during pregnancy with return to normal after delivery" (adapted from ADA[17] – the original definition did not include return to normality after delivery, but by definition glucose intolerance that does not resolve is not gestational – it either preceded pregnancy and was only diagnosed during it, or was type 2 diabetes or impaired glucose tolerance (IGT) that came on during pregnancy.)

The first definition, adapted by the National Diabetes Data Group (NDDG) in the USA in 1979[18] from the original O'Sullivan and Mahan[19] version, was based on levels in a 3-hour OGTT that predicted later diabetes in the mother. A later modification of this by Carpenter and Coustan[20] lowered the threshold for diagnosis. This definition was adopted by the Fourth International Workshop-Conference on Gestational Diabetes Mellitus,[21] and there is evidence in support of this threshold from a 1996 Toronto study in which the marginal group (i.e. those diagnosed by Fourth Workshop but not NDDG criteria) had worse outcomes than both women with normal levels, and those who had GDM as defined by the ADA but were treated.[22] However, there were confounding factors in the borderline group, who were older and heavier.[23]

This shift from a diagnosis based on maternal risks to one that incorporates fetal risks increased the prevalence of GDM in the USA from 4% to 7% of (mainly white) pregnant women.

In Europe, diagnosis has been based on the 75-g OGTT. The new WHO criteria[24] define GDM as either diabetes (FPG ≥ 7.0 mmol/l or 2-hour glucose 11.1 mmol/l or over) or IGT (2-hour glucose ≥ 7.8 mmol/l), which represents a lowering

of the FPG threshold (from ≥ 7.8 to ≥ 7.0 mmol/l). In one study comparing new and old criteria,[25] there was little difference in prevalence, mainly because nearly all those with FPG between 7.0 and 7.8 mmol/l already qualify as GDM through 2-hour levels ≥ 7.8 mmol/l. The glucose load affects some results more than others. Weiss and colleagues compared results after 75-g and 100-g OGTTs, and found no difference in the 1-hour levels, but a 0.94 mmol/l difference in the 2-hour ones.[26]

A review of the literature on definitions by Martin[27] found that the different criteria could lead to a difference in diagnosis from 1% to 10% of pregnant women. In one Australian hospital, the standard OGTT dose has been 50 g.[28]

Whereas IGT has an upper limit, GDM does not, and this could lead to the diagnosis of GDM being associated with a spectrum of risk.

A good review of the background to the definitions debate is given in chapter 2 of the book by Dornhorst and Hadden.[29]

Problems

Several difficulties arise with GDM. Firstly, glucose levels in pregnancy form a continuum, rather than showing a bimodal distribution between those with GDM and those who are normal. This is similar to other conditions such as hypertension and hyperlipidaemia, where the risk of adverse consequences is proportional to the degree of elevation. There is no threshold that clearly distinguishes between low-risk and high-risk pregnancies.[30] In the Toronto study, there was a gradient of risk in women with glucose levels below the GDM threshold (all with OGTT results normal), with the proportion with a macrosomic child doubling with every 1 mmol increment in 3-hour blood glucose level.[31]

Secondly, the harms seen, such as the shoulder dystocia, hypoxia and neonatal hypoglycaemia that are associated with macrosomia, are also seen, although less often, in women with lower levels of blood glucose than those at which GDM is usually diagnosed. It may be that blood glucose is a marker of a metabolic state associated with higher risk, rather than the direct cause of complications. Again, this is similar to the situations with blood pressure and hyperlipidaemia where the consequences, such as stroke and myocardial infarction, are not unique to those conditions, just more common.

Thirdly, macrosomia itself is defined by weight at birth, rather than by ill health. An arbitrary cut-off of 4000 g will include many healthy large babies. Refinements such as correcting for gestational age will help, but still do not distinguish between normality and disease. The cut-offs that have been used are 4000 and 4500 g; these are arbitrary values based on the normal distribution, rather than on changes in the incidence of complications. Essel and Opai-Tetteh[32] reported that only 10% of mothers of babies over 4000 g birth weight had GDM; most 'macrosomic' babies are born to non-GDM mothers. They also commented that mothers of macrosomic babies tended to be older, fatter and multiparous, and that weight over 70 kg at delivery was one of the strongest predictors of high birth weight. This study was done in a poor, mainly black area of South Africa. Similarly, a Brazilian study reported in 2000 that only 2% of large babies could be attributed to GDM.[25] The Toronto study had a similar finding – 91% of babies with birth weight over 4000 g were born to mothers with normal glucose levels. Even using the higher cut-off of 4500 g to define macrosomia, the proportion born to mothers with GDM is small. Spellacy and co-workers[33] found that only 5% of neonates with birth weight over 4500 g were born to GDM mothers, giving a relative risk of 3.0 for GDM, which compares to a relative risk of 26 for babies born to women weighing over 90 kg, who had 44% of macrosomic babies.

If we use fetal hyperinsulinaemia (via amniotic fluid or cord blood measurements) as a guide to which babies are larger but unhealthy, then only about 15–20% of babies over 4000 g birth weight are abnormally so.[34,35] However, the correlation between birth weight and cord insulin is weak ($r = 0.22$), as shown in a series of unselected births in Sheffield, where high cord insulins were found in babies of all birth weights.[35]

Ales and Santini reviewed the literature in 1989 and noted that there was evidence of increased neonatal morbidity only for birth weights over 4500 g, with no increase in those between 4000 g and 4500 g.[36]

The key difference between 'macrosomic' and healthy large babies is overgrowth of insulin-sensitive tissues such as adipose tissue, especially around chest and shoulder and abdominally.[14,37] Unfortunately, we have no easy way of distinguishing these groups. However, weight alone shows marked correlation with morbidity, as the data in *Table 1* show.[33]

Sacks[38] has suggested other ways of defining macrosomia. Firstly, he points out that we should refer to

TABLE 1 *Incidence of adverse events by birth weight*

	Birth weight		
Adverse event	**2500–3499 g**	**4500–4999 g**	**≥ 5000 g**
Shoulder dystocia	0.3%	7.3%	14.6%
Birth injury	2.7%	6.5%	11%
Perinatal mortality	3.5 per 1000	8.1 per 1000	24.4 per 1000

weight for gestational age, not just a fixed 4000 g for all ages. Secondly, he suggests a ponderal index of (weight × 100 / length cubed). Thirdly, he suggests an index of 'birth symmetry', this being the ratio of weight to height, with both expressed as ratios of the 50th percentile. He notes that heads tend to be smaller in GDM. Sharp and co-workers[39] found that neonatal BMI, despite the difficulty in measuring length, gave a better guide to infant adiposity than weight alone. Hammami and colleagues reported that babies who were large for gestational age (LGA) and born to mothers with IGT, had more fat than those who were LGA but born of mothers with normal glucose tolerance.[40]

The topic of GDM is already bedevilled by poorly defined terms, but perhaps we need another term to replace macrosomia, in order to indicate more clearly that the problem is not just size, but the abnormal growth pattern seen in infants subjected to high levels of insulin *in utero*. Hyperinsulinaemia-induced macrosomia? 'Diabetes-related macrosomia' would not do, because some children of mothers with GDM will also be large healthy babies.

Schwartz and Teramo reviewed the history of the term, and noted that it has replaced two other unsatisfactory terms – large for gestational age in the USA, and heavy for dates in the UK.[41] The authors quote Potter and Craig[42] as recommending that "the term macrosomia should be applied only when organ weight is disproportionately great in relation to body weight".

The underlying cause of pathological macrosomia is thought to be fetal hyperinsulinaemia in response to maternal hyperglycaemia, and amniotic fluid insulin is higher in women with GDM than in those with normal glucose tolerance, although not as high as in those with type 1 or 2 diabetes.[43] Healthy larger than average babies do not have high cord insulin levels.

A study in which amniotic fluid insulin levels were measured and compared with OGTT results suggested that a cut-off of 8.9 mmol/l at 1 hour may identify maternal gestational glucose intolerance sufficient to cause increased fetal insulin production, but the study was done only on women with elevated amniotic insulin levels, and so cannot give specificity.[34]

Current definitions are statistical, not pathophysiological. As Schwartz and Teramo[41] point out, using statistical cut-offs means that boys have a higher rate of macrosomia than girls, because boys are on average 130 g heavier. Also, some countries will have higher rates, and even within countries there may be differences because of, for example, altitude or ethnicity (white infants being on average 140 g heavier than Afro-Caribbean ones).

Defining GDM by adverse consequences

The difficulty of defining conditions by the frequency of adverse events, when these are not unique but only more common, has been referred to already. A Danish study compared outcomes in 143 consecutive GDM pregnancies with 143 controls, matched for age and pre-pregnancy BMI.[44] The frequencies of some outcomes are shown in *Table 2*.

TABLE 2 *Comparison of outcomes in GDM pregnancies and matched controls*

	GDM group	Controls
Maternal hypertension	20%	11%
Macrosomia (> 4500 g)	14%	6%

Other adverse events, such as Caesarean section, induction of labour, admission to neonatal unit, and neonatal hypoglycaemia, were more common in the GDM group, but this may have been a result of the diagnosis itself. The Toronto study reported that Caesarean sections were more common in mothers diagnosed as having GDM, and that this was seen in those with no macrosomia.[22] In the Danish study, most infants of GDM mothers had a blood glucose tested, but few of those of non-GDM mothers.[44] Thirty-eight of the

GDM mothers received tolbutamide or insulin treatment but hypoglycaemia (defined as < 2.0 mmol/l) was almost as common (21%) in the group treated with diet. Perhaps the most notable finding from this study was that treatment of GDM with tolbutamide or insulin had little effect on outcomes such as macrosomia.

One attempt to define IGT of pregnancy by adverse outcomes failed because there were insufficient differences in outcomes between 212 women with IGT in pregnancy (diabetes having been excluded) and 212 controls.[45]

Two studies reported that the relationship between adverse outcomes such as macrosomia and glucose levels shows a continuum.[31,38] Another study found that only 5% of mothers of macrosomic (> 4500 g) infants had GDM – so that if the aim is to reduce the number of infants with birth weight over that level, interventions in GDM could have very little effect.[33] The same study found that 45% of mothers of large babies were over 90 kg in weight, suggesting that maternal overweight is much more important than glucose intolerance.

Similarly, Nordlander and co-workers[46] compared two groups of infants, one healthy and the other with classical morbidity, born to mothers with GDM, and found that there was no difference in third trimester maternal blood glucose, but that maternal weight pre-pregnancy was a significant predictor of neonatal morbidity.

Diagnostic levels

The original diagnosis of GDM was based on the likelihood of progression to later permanent diabetes.[19] However, the current focus in the UK is on outcomes of pregnancy, related to the St Vincent target of getting outcomes in diabetic pregnancies to approximate to those in non-diabetic pregnancies. Hence the original criteria may not be appropriate.

The WHO provisional recommendations[24] defined GDM in terms of the 75-g glucose tolerance test (GTT), but using cut-offs for both diabetes and IGT – in effect an FPG of 7.0 mmol/l or over or a 2-hour OGTT glucose of 7.8 mmol/l or over. So gestational 'diabetes' includes a much larger range than non-gestational diabetes. This is despite the known physiological increase in glucose levels in pregnancy, although this applies only to non-fasting levels; fasting levels are similar, as shown in a population-based study from Sweden that compared OGTT results at 17 and 32 weeks' gestation.[47] The 2-hour mean levels were 6.5 and 8.0 mmol/l at 17 and 32 weeks respectively. Applying the WHO criteria would suggest that 18% of these Swedish women had GDM.

The effects of this have been shown in a recent very large study from Brazil.[25] Over 5000 pregnant women had a 75-g GTT. The authors used the usual cut-off levels for diabetes and IGT outwith pregnancy to subdivide those who, according to the WHO criteria, had GDM. Very few of those with GDM would meet the criteria for diabetes – a range from 0.1% in the age range 20–24, rising to 1.2% of those aged 35 and over. Over all ages, only 0.4% of those with GDM had diabetes. This study also found that GDM as defined by WHO criteria would account for only four of all macrosomic births.

Hence GDM as currently defined is nearly all IGT of pregnancy, but based on non-pregnant glucose levels, without taking into account the physiological increase in non-fasting glucose levels in pregnancy. The risks in the IGT group seem to be low. Nasrat and co-workers[45] found that in their untreated IGT group (FPG ≤ 8 mmol/l, 2-hour glucose 8–11 mmol/l, after a 100-g OGTT), birth weight was a little higher, neonatal blood glucose was a little lower, and haematocrit a little higher (they give data in relation to means and standard deviations, rather than absolute values) but that no difference in adverse outcomes was observed.

It has been argued that it is difficult to disentangle the effects of glucose elevation from the covariants associated with it, such as weight, that glucose intolerance is merely a marker for other underlying conditions, and that it may be these other conditions that cause the morbidity.[48] The authors of this study note that screening 1000 pregnant women with the 75-g GTT would find 76 cases of GDM, but that only four of these would have diabetes.

Using the OGTT, at least two values must be abnormal before GDM is diagnosed. Langer and colleagues[49] noted that women with only one abnormal value had an increased proportion of macrosomic babies, and carried out a trial in this group, randomising half to management with diet and, if necessary, insulin, with the aim of achieving tight glycaemic control. The treated group had fewer macrosomic babies than the untreated group (6% versus 24%). If we define a disease through ability to benefit from treatment (and if we assume that lowering birth weight is of

benefit), then this study suggests that the current criteria are too high.

Naylor[50] has reviewed the history of the diagnosis of GDM. He concludes that the criteria are conceptually flawed by having a dichotomous definition of normal and abnormal, whereas what is needed are data on outcomes across a gradation of fasting and GTT levels.

Green and co-workers[51] found that there was a much stronger link between maternal BMI and infant BMI than there was with maternal glucose levels, until the maternal plasma glucose was over 11.1 mmol/l (after a 50-g GCT), after which there was a steep rise.

Conclusions

GDM as currently defined covers a range of glucose intolerance, most of which is IGT rather than diabetes, although the proportion will vary amongst ethnic groups. There are varying diagnostic thresholds, but there appears to be a continuum of risk, making dichotomous definitions inappropriate. There are also problems defining the harms that GDM may do, and in particular the use of birth weight to define macrosomia, which fails to distinguish between healthy large babies and those who are overweight and unevenly developed as a result of maternal hyperinsulinaemia.

Chapter 2

Does screening for GDM satisfy the NSC criteria?

In 1966, the WHO published their criteria for assessing screening programmes.[52] The UK NSC has adapted and updated these to take account of modern standards for rigour of evidence, and the greater concern about the potential harms that can result from screening for diseases. The NSC criteria are published in its annual reports and on its website.[53]

This chapter applies the 19 criteria to GDM, using the term to also cover IGT in pregnancy.

1. **The condition should be an important health problem.**
   Important can be defined in several ways, including frequency of the condition, severity of sequelae, and economic burden of disease. Only a small proportion of pregnant women has GDM. The exact percentage varies with definition, but in the UK it is usually estimated to be around 2%. Jensen and colleagues[44] compared results in 143 GDM pregnancies with the same number of controls matched for age, parity and BMI, and found higher rates of maternal hypertension (20% versus 11%), macrosomia (14% versus 6%) and neonatal hypoglycaemia (24% versus 0%). They also found increases in induction of labour, Caesarean section and admission to neonatal care, but these could have been a result of the diagnosis itself.

   Hence about 0.5% (based on complications in up to 24% of the 2%) of all infants may suffer adverse consequences – a small proportion but quite a large absolute number (about 3100). However, were there to be universal screening, 622,000 women a year in England and Wales would need to be screened in order to identify 12,400 in whom there would be intervention aimed at reducing complications in about 10% (e.g. from 14% to 6%, based on Jensen and co-workers;[44] about 1240 babies); the interventions would not succeed in all so the number of adverse events prevented would be much less than 1240.

   The health of the mothers also needs to be considered. There is an increased risk of later type 2 diabetes.[19]

   Criterion met? Borderline – low proportion of all births adversely affected, but individual adversity sometimes serious.

2. **The epidemiology and natural history of the condition, including development from latent to developed disease, should be adequately understood and there should be a detectable risk factor, or disease marker, and a latent period or early symptomatic stage.**
   We have some knowledge of the epidemiology. We know the prevalence in pregnant women, although the cut-off is uncertain (discussed further under criterion 5). There are predisposing factors, especially obesity. There are ethnic variations, partly explained by variations in weight. We lack knowledge of natural history in the UK at present, although there are long-term follow-up studies from other countries, albeit in previous decades. We do not understand the mechanism of progression to type 2 diabetes, or why some women progress and others do not, but it is assumed that GDM is the first sign of an insufficiency of insulin production, either absolute or related to increased needs because of insulin resistance.[30]

   There is a latent period between resolution of GDM after delivery, and later type 2 diabetes. There is also an asymptomatic period during pregnancy, when glucose levels are high but before these have any adverse consequences for the baby.

   Criterion met? Yes.

3. **All the cost-effective primary prevention interventions should have been implemented as far as possible.**
   Because one of the main risk factors (perhaps the dominant one) is weight, this criterion implies that there should be measures to promote normal weight amongst women who are going to become pregnant. While this aim should be pursued, we know from research and reviews in obesity that intervention is often ineffective.[54] However, in terms of absolute numbers, most cases arise in women who are

overweight rather than obese, and with less weight to lose, there may be more chance of success. Weight loss to normal levels has to occur before pregnancy, because weight loss during pregnancy is undesirable.[55]

Evidence in support of some degree of preventability comes from the seasonality of macrosomia, which is highest in January to April, drops in spring, and then rises again. This is thought to relate to maternal weight gain during festive and holiday periods.[56]

Criterion met? No. There are health promotion campaigns to encourage people to lose weight, but not targeted at 'pre-pregnant' women.

4. **There should be a simple, safe, precise and validated screening test.**
Screening tests are considered in detail in chapter 3. The tests used, such as FPG or GCT, are fairly simple and laboratory quality assurance is assured. The tests are safe. However, the variety of tests used indicates a weak evidence base on best test, and as outlined in chapter 1, validation is lacking.

Criterion met? Partly – simple and safe, but validation lacking.

5. **The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.**
The distribution of test values is known, as is the physiological elevation of plasma glucose in pregnancy, but the cut-off level is uncertain. Should the non-pregnant thresholds be used, as recommended by the WHO, or should higher cut-offs be used, as argued by Lao and Lee?[57]

If we use later diabetes as a guide to diagnostic thresholds, we should use higher cut-offs in pregnancy, because at any level of raised glucose, non-pregnant women are at higher risk than pregnant ones.[58]

Criterion met? No.

6. **The test should be acceptable to the population.**
Criterion met? Not known. Consent is assumed, and the simplicity of a simple blood test, probably done at a routine visit, suggests acceptability. However no studies of acceptability in women who have been fully informed appear to have been done. Rates of vomiting of between 0.5% and 11% have been reported with the 75-g GTT.[59–62]

Criterion met? Yes? Tests are examined further in chapter 3.

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.**
Further diagnosis in the UK involves an OGTT using 75 g of glucose, an artificial and unphysiological test with poor reproducibility. Some authorities argue in favour of more natural test meals.[63] In other parts of the world, the OGTT is done with different amounts of glucose, such as 100 g (standard in the USA) or 50 g in Australia.[28]

A 1987 study compared the use of NDDG and WHO criteria, and concluded that using the WHO criteria would reduce the number diagnosed as having GDM by about half.[64]

Criterion met? No?

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.**
Diet and insulin. Treatment takes various forms, and depends on the level of blood glucose. Interventions include diet alone, diet plus insulin, and intensive monitoring of plasma glucose levels, usually by finger prick and home testing. Most women with GDM are treated with diet alone, but about 15–20% are thought to need insulin. A meta-analysis of insulin treatment trials shows that the incidence of macrosomia can be reduced considerably, from 17% to 6% [odds ratio (OR) 0.35; 95% confidence interval (CI), 0.24 to 0.52] but that there were no reductions in rates of Caesarean section or birth trauma.[65] However, other studies have reported that macrosomia still occurs even in women with very good diabetic control.[66] Evidence for the efficacy of treatment comes from the Toronto study, in which the group with diagnosed and treated GDM had lower rates of macrosomia than the untreated borderline group – 14% versus 29%, with normal mothers having a rate of 10%.[22]

A Danish case–control study in 2000 concluded that treatment did not seem to affect

outcomes.[44] Persson and co-workers[67] randomised women to diet alone and diet plus insulin, and found no difference in outcomes, but 15% of the diet alone group received insulin, and this may explain the lack of any difference in diabetes control between the groups.

Simmons and Robertson[68] argue that treatment of GDM mothers may benefit the infant not only perinatally, but also in terms of reduced obesity later (follow-up was 2 years), but this study was not an RCT, had low numbers (20) and confounding variables (insulin-treated mothers were older and more obese).

The first-line treatment for all women is diet, although in a Cochrane review, Walkinshaw found no differences in the frequencies of birth weight over 4000 g or Caesarean section between women treated with diet and controls who received no dietary treatment.[65] Currently there is uncertainty about the best method for deciding which women need insulin. Most clinicians use maternal blood glucose levels, for example, Langer and co-workers[69] recommended that insulin be considered for mothers with FPG of 4.8 mmol/l. However, the optimum blood glucose levels are not known for certain. Should the target for treatment be levels as seen in non-pregnant women, or the higher levels seen in normal pregnancies? An alternative to blood glucose is to monitor fetal development, either by non-invasive ultrasound[70,71] or by more intrusive methods such as amniotic insulin measurement (reviewed by Kjos and Buchanan[30]). Unfortunately ultrasound is probably no better than clinical assessment at predicting macrosomia.[72]

Interestingly, the elevation of blood glucose levels seen in pregnancy in Western countries, and which is believed to be a result of diminished insulin sensitivity, is not seen if pregnant women adopt a Third World diet high in fibre, suggesting that diet could have more potential for preventing the changes in glucose tolerance in pregnancy than previously thought.[73]

Insulin is now not the only option for those women with glucose levels not controlled by diet. The oral hypoglycaemic agent glyburide is a large molecule that does not cross the placenta, and is therefore an alternative to insulin. Langer and co-workers[74] report that most women can be treated with glyburide, with only 8 out of 201 needing insulin.

Monitoring. Intensive glucose monitoring has been shown to reduce macrosomia, partly through increasing the use of insulin. Goldberg and co-workers[75] in a before and after comparison, with well-matched groups, found that home glucose monitoring, fasting and postprandial, reduced the incidence of macrosomia from 24% to 9%, mainly by detection of high glucose levels and subsequent treatment with insulin; 50% of the home monitoring group used insulin compared to 21% of the historical controls. Rey[76] found that in women with plasma glucose over 7.8 mmol/l 1 hour after a standard meal, tight glucose monitoring (at home) produced better outcomes than fortnightly glucose monitoring in an outpatient clinic.

Exercise. In addition to diet and insulin, there have been a few studies of exercise. Bung and colleagues[77] randomised Hispanic women with fasting hyperglycaemia to diet and exercise (recumbent cycling) versus diet and insulin and concluded that exercise did not harm the fetus. Jovanovic-Peterson and Peterson[78] found that adding exercise to diet improved FPG by about 1 mmol/l after 4 weeks, but that upper body exercises were less likely to trigger uterine contractions than some lower body ones; rapid walking on a treadmill was fine but jogging and cycling were associated with contractions. Avery and co-workers[79] carried out an RCT of exercise. It was not blinded, the mode of randomisation was not clear, and there were selection effects from the high refusal rate, with only 29 of 144 being recruited. Postprandial glucose levels were up to 10 mg/dl lower in exercisers but no differences in outcomes were seen, perhaps because of small numbers. Lesser and colleagues[80] reported no difference in postprandial glucose levels, but this study had very small numbers (six women with GDM) and only a single bout of exercise.

A questionnaire survey by Dye and co-workers[81] in a very large population-based cohort in 1997 found no difference in the prevalence of GDM amongst regular exercisers except in those with BMI over 33, in whom the prevalences of GDM were 10.3% in those who took no exercise and 5.7% in those who did.

Criterion met? Uncertain, because the benefits reported have been in terms of reduction in macrosomia rather than in adverse outcomes such as birth trauma or Caesarean section.

It has also been stated that screening for GDM can have long-term benefits by reducing later type 2 diabetes[82] but Wein and colleagues[83] found that intensive dietary advice made no difference.

Gregory and colleagues[84] suggested that "continuation of dietary and behavioural changes initiated during pregnancy theoretically could delay or prevent progression to overt diabetes". Unfortunately there is no scientific literature to validate these claims, which assume that behavioural change is long term, achieved via counselling at "semiannual contraceptive or yearly gynecologic evaluations" rather than special visits. All women are assumed to see their doctor for dietary advice and glucose monitoring. However, many women may not routinely attend. In addition, dietary advice may be better given by a dietician rather than a general practitioner. Compliance after pregnancy might be low, especially in the absence of adequate monitoring and follow-up. Regular follow-up visits to dieticians or clinics would have to be arranged at additional cost to maximise compliance with dietary advice.

9. **There should be agreed evidence-based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.**
   There is partial agreement over treatment of women with the highest glucose levels (by OGTT, as defined by ADA), by diet first, then insulin. But there is disagreement about the value of treatment in the groups below the diagnostic threshold, despite evidence of harm when untreated[22] and benefit from treatment.[85]

   Criterion met? Partially.

10. **Clinical management of the condition and patient outcomes should be optimised by all healthcare providers prior to participation in a screening programme.**
    The treatments – diet, insulin and intensive monitoring – are standard ones.

    Criterion met? Yes.

11. **There must be evidence from high-quality RCTs that the screening programme is effective in reducing mortality or morbidity.**
    As has been pointed out by both the US Preventive Services Task Force[5] and the Canadian Task Force on the Periodic Health Examination,[6] there are no high-quality RCTs of screening with morbidity as an outcome.

    Screening versus no screening. There has been a natural experiment in Ontario, where screening policies differed markedly between the Hamilton catchment area (around the evidence-based centre in McMaster) and the rest of the province.[86] There was universal screening everywhere except Hamilton, but no screening there. There was a steep rise in reported GDM in the rest of Ontario, and a decline in Hamilton. Yet the incidence of macrosomia was identical, at 12.7% and 12.5%. The authors conclude that there is no evidence that universal screening has had any beneficial effects on outcomes of pregnancy, and they call for randomised trials of screening.

    The harms of screening and treatment also need to be considered. Langer and co-workers[85] noted that the group with tightest control had more small for gestational age babies. The proportion was higher in those treated with insulin, but was also increased in those on diet alone. The authors conclude that both high and low levels of insulin in the fetal circulation can do harm.

    Criterion met? No.

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.**
    There is a lack of published evidence on acceptability, but screening is standard in many areas (although tests used are not consistent[1]) without apparent dissent, and perhaps this indicates that it is acceptable. However, the extent to which consent is fully informed is not known. Are the women being screened given details of possible treatment, such as insulin injections?

    Criterion met? Probably.

13. **The benefit from the screening programme should outweigh the physical and psychological harm caused by the test, diagnostic procedures and treatment.**
    The uncertainties surrounding benefits have been outlined earlier. There are also harms, such as anxiety following the diagnosis, and the increased rate of operative delivery. The

Toronto study[22] found that the diagnosis itself increased the rate of Caesarean delivery even in the absence of macrosomia, to 33% in GDM compared to 20% in non-diabetic mothers. Caesarean section is a major operation, with inevitable surgical and anaesthetic risks. Not all centres have observed the same effect on Caesarean section rates. A New South Wales study reported a Caesarean section rate of 20% in women with GDM, compared to 16% in non-diabetic women, and the difference was partly explained by age and parity.[87]

A Cochrane review also found that reduction in macrosomia is not necessarily followed by significant differences in the rates of sections, forceps deliveries or birth trauma.[65]

Santini and Ales[88] have calculated that to prevent one case of macrosomia, 3716 women would have to be screened; 250 would require further tests such as ultrasound, and 134 more women would have Caesarean sections. It should be noted that only 20–30% of babies of women with GDM have macrosomia.[30]

The benefit to harm ratio would be increased by more selective screening. A study in 1987 showed that an obese woman aged 25–29 with a family history of diabetes has a 1 in 200 chance of GDM, and that a non-obese woman has a chance of 1 in 500.[89]

One of the costs, both human and financial, is the high rate of Caesarean section in women classified as having GDM.[22] Rouse and co-workers[90] wondered if ultrasound could be used to promote a more selective section rate, but found that 443 sections (and US$930,000) would be needed to prevent one permanent brachial plexus injury.

Another possible cost is that the effect of treatment may be to shift the whole spectrum of birth weight downwards, so that reducing the number of large babies means increasing the number of very small ones.[85]

There is also the problem of who benefits? Some authorities advocate very tight control of maternal blood glucose levels in order to secure best outcomes for the baby,[91] but this may expose the mother to the dangers of hypoglycaemia (reviewed by Fraser[66]). Maternal hyperinsulinaemia also exposes the baby to the risk of neonatal hypoglycaemia.[92] Again, selective treatment of mothers based on assessment of fetal risk according to maternal glucose level[67] or ultrasound screening[71] may target intervention more effectively at the 15% of infants who are at risk.

Criterion met? Uncertain.

14. **The opportunity cost of the programme (including testing, diagnosis, treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e. value for money).**
Testing requires no special equipment, and could be done in primary care. Treatment with diet or insulin requires only staff time. Intensive monitoring is usually done with the aid of glucose meters, but these are inexpensive and widely available. The costs of the programme can be calculated, but the benefits are less certain. We do not have any common currency expression of benefit such as cost per quality-adjusted life year (QALY), or cost per adverse event avoided. If some adverse events such as Caesarean sections and admission to a special care baby unit (SCBU) can be avoided, the costs of screening will be offset.

    This report will endeavour to produce some estimates in a later chapter.

    Criterion met? Uncertain.

15. **There must be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.**
Criterion met? No. There is no national screening programme, and hence no national standards or quality assurance.

16. **Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be made available prior to the commencement of the screening programme.**
Criterion met? Uncertain. The facilities for testing exist, in both hospital clinics and primary care, but we lack data on the availability of dietetic and specialist nurse support, both of which are in short supply in some areas.

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 provided or current interventions increased within the resources available.
One other option has been considered – that instead of screening and early intervention, there should be ultrasound monitoring for macrosomia, and intervention once that was detected. However, the poor predictive capacity of ultrasound has already been referred to, and the other objection to this approach would be that by the time macrosomia had been detected, damage would have been done. So there does not seem to be a late detection/late treatment option.

Criterion met? Yes.

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.**

    It is not known whether all women have access to accurate information.

    Criterion met? Uncertain.

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.**
    Not applicable at present.

Conclusion

Screening for GDM meets only some of the NSC criteria.

Chapter 3

Evidence of effectiveness

Methods used for the review

Preliminary searches showed that very few RCTs of screening practices existed and so the review of the literature was unrestricted to study design. All primary studies that investigated any method of screening for GDM were included. The search included the authors' personal reference collections and searches of MEDLINE, EMBASE and the Cochrane library. The search strategy included a series of keywords including gestation*, diabet* and screen* and were exploded for clinical trial, cohort studies, case–control studies and research design. The full search strategy is given in appendix 1. Citations of retrieved references were also searched. Authors of studies reviewed were not contacted.

Selection of papers for review

The search identified a large number of potential papers for review. The titles and abstracts were inspected independently by two of the authors to assess their relevance to the focus of the review. The majority of studies identified by the searches were case series, but a number of quasi-experimental observational studies were identified. We did not include studies that evaluated the effects of antecedent diabetes on pregnant women, or studies that did not evaluate screening for GDM in some way. Only English language studies were identified. In total, 135 studies were included at this stage (appendix 2).

Relevant papers were retrieved and reviewed by two of the authors. Data extracted from these studies included tests and thresholds used for screening and diagnosis, incidence of GDM, sensitivity, specificity and positive predictive value (PPV) of the tests, country of study, time of testing and fasting status. The studies' calculations of sensitivity, specificity and PPV were checked. Cases were noted where data were not available in the original report to check calculations.

Diagnostic tests

The GTT is regarded as the 'gold standard' for diagnosis of GDM after a positive screening result, although it is deemed "too time-consuming and expensive for routine screening".[5] However, the quantity of the glucose load and the thresholds for diagnosis are issues of contention. The GTTs most commonly used are the 3-hour test with a 100 g glucose load and the 2-hour test with 75 g glucose. The former is used in the USA, the latter in Europe.

The 3-hour 100-g GTT

The 3-hour 100-g GTT uses a 100 g glucose load taken orally after fasting, with blood glucose being measured before glucose administration and then after 1, 2 and 3 hours. Most guidelines recommend that two of the four values meet or exceed the criteria for an abnormal GTT in order to diagnose GDM. Owen and co-workers[93] found that of 280 practising resident doctors in the USA, 96% thought that two abnormal GTT values defined GDM, whereas only 4% considered one abnormal value was sufficient.

The original classification for GDM through a positive GTT was set in 1964 by O'Sullivan and Mahan[19] and based on whole blood samples. These were converted for plasma glucose determination by the NDDG in 1979.[18] However, because the conversion algorithm (plasma glucose = 1.14 × blood glucose) was felt to lack accuracy (O'Sullivan and Mahan used Somogyi-Nelson and current practice uses enzymatic methods, which give lower values), Carpenter and Coustan (C&C)[20] modified the criteria in 1982 to plasma glucose = ((whole blood glucose – 0.2775 mmol/l) × 1.14). Magee and co-workers[94] tested both the NDDG and C&C's modified criteria and supported the use of the more inclusive modified criteria. They demonstrated that approximately 50% more women were classed as GDM with the C&C modified criteria, with similar incidences of perinatal complications. However, there are still variations in the adoption of these criteria and both are commonly referred to in the published literature.[95–99] Details of the criteria are given in *Table 3*.

Blood or plasma levels are presented hereafter in mmol/l, with the conversion rate of 1 mg/dl = 0.0555 mmol/l having been used. Early papers, such as O'Sullivan and co-workers,[100] used the Somogyi-Nelson method of glucose measurement. Modern methods now give lower results. A 7.2 mmol/l (130 mg/dl) threshold used by

TABLE 3 *The NDDG/O'Sullivan & Mahan and C&C thresholds for diagnosis of GDM*

	NDDG/O'Sullivan & Mahan (mmol/l)	C&C (mmol/l)
Fasting	≥ 5.8	≥ 5.3
1-hour	≥ 10.5	≥ 10.0
2-hour	≥ 9.1	≥ 8.6
3-hour	≥ 8.0	≥ 7.8

these authors equates to 7.9 mmol/l (142 mg/dl) by today's methods.

The result of having different sets of thresholds for the 100-g GTT is varying levels of prevalence depending on which set are used, with the C&C criteria always giving a higher prevalence of GDM. In a high prevalence, predominantly African–American population, Bobrowski and co-workers[101] found a rate of 29% using the NDDG criteria compared to 38% using the C&C criteria. A 1999 Italian study by Corrado and colleagues[102] found prevalences of 3.4% (34/1000) using NDDG and 4.6% (46/1000) using C&C. A US study including predominantly black, Hispanic and Indian women detected prevalences of 3.9% and 4.4% using the NDDG and C&C criteria, respectively.[103] Perucchini and colleagues[104] found that omitting the 3-hour value would have missed 7.6% of women with GDM.

This would naturally affect the sensitivity and specificity of a screening test. As Sermer and co-workers[31] demonstrated, in a population aged ≥ 24, of whom 90% had a GTT, a 1-hour 50-g GCT (≥ 140 mg/dl) had 77% sensitivity, 82% specificity and a PPV of 14% using the NDDG classification of GDM. However, when the C&C classification was used, sensitivity was 68%, specificity 84% and PPV 23%.

Reporting of reproducibility has been reasonable. Harlass and colleagues[105] found 78% reproducibility in a population of army dependents (O'Sullivan and Mahan criteria) at 1–2 weeks. Neiger and Coustan[106] found 34% (36/106) of women with only one abnormal value had two abnormal values on a repeat test and thus were classified as having GDM (O'Sullivan and Mahan criteria).

Perucchini and co-workers[104] put the incidence of vomiting with the 100 g load at 6.8%, and Fachnie and co-workers[107] reported 4%.

The 2-hour 75-g GTT

As with the 100-g GTT, varying classifications for the 75 g version have also been reported. The WHO classification of GDM requires either a pre-glucose or 2-hour test to be equalled or exceeded (*Table 4*). The WHO criteria of 1980[108] (fasting ≥ 8.0 mmol/l, 2-hour ≥ 11.0 mmol/l) was revised to that shown in *Table 4* in 1999.[24]

TABLE 4 *WHO criteria*

	Plasma (mmol/l)	Blood (mmol/l)
Fasting	≥ 7.0	≥ 6.1
2-hour	≥ 11.1	≥ 10.0

It is interesting that the WHO has used the original NDDG conversion algorithm rather than the modified C&C one.

In a Danish population, all of who had a 75-g GTT, Kvetny and co-workers[109] found an incidence of 3.6% using the WHO classification, but these were women with risk factors.

Hatem and colleagues[62] found a reference value for the 75-g GTT (determined by 97.5th percentile) at 2 hours of 9.6 mmol/l, during the third trimester. In 1995, Martin and co-workers[60] found the 95th percentile was 5.1 mmol/l fasting and 7.8 mmol/l at 2 hours. Incidence of GDM, expectedly, varied by criteria used: 4.2% using a 2-hour value of ≥ 8.0 mmol/l; 5.2% using a 2-hour value of ≥ 7.8 mmol/l; and 5.5% using a fasting value of ≥ 5.5 mmol/l and 2-hour value of ≥ 8.0 mmol/l.

Reported rates of vomiting have varied from 0.5% to 11%.[59–62] Hatem and co-workers[62] suggested that the WHO-recommended 75 g glucose load should be used to reduce the incidence of vomiting seen with the 100 g load. There is no evidence supporting the effectiveness of this strategy.

The 3-hour 100-g GTT versus the 2-hour 75-g GTT

A study in Thai women by Deerochanawong and co-workers[110] compared NDDG (3-hour 100-g GTT) criteria with WHO (2-hour 75-g GTT with ≥ 7.8 mmol/l at 2 hours) and found prevalences of GDM of 1.4% (10/709) with NDDG and 15.6% with WHO. All the women in the NDDG group

were identified by the WHO criteria. This illustrates the consequences of choosing criteria for a GTT.

Li and colleagues[64] carried out 75-g GTT tests in 216 women who had been diagnosed as having GDM by NDDG criteria following the 100-g test. By WHO criteria, 51% had normal glucose tolerance, and when randomised to treatment or no treatment, there was no difference. This implies that using the WHO criteria based on a 75-g test would not result in any harm, but would save patients' time and healthcare funds.

Other diagnostic tests

A 50-g GTT has occasionally been used, but not evaluated in the literature. Hanson and Kallner,[111] in a population with risk factors, found 98% agreement between 2-hour and 3-hour 50-g GTTs. Fachnie and co-workers[107] advised 50 g or 75 g loads for patients incapable of tolerating the 100 g glucose load.

Screening tests

A number of screening tests for GDM have been devised. These include risk factor screening, urine testing, various blood tests, and combinations thereof. The majority of the papers used in this review have been observational studies, with only a few incorporating any control groups. Very few studies have used an RCT design.

Since O'Sullivan and co-workers' pioneering work with the GCT,[100] the majority of the subsequent work has assumed 100% sensitivity of this screening test and forwarded only positive screens for a diagnostic test. Thereby, arguably, many women with GDM may have been missed and uncertainty remains as to the true incidence of GDM and the sensitivity, specificity and value of the various screening tests, and their suggested cut-off points. This assumption has been the focus of debate in the literature[103] but has not been substantiated empirically. Incremental rises in incidence of GDM cases above various cut-offs on the GCT were demonstrated by Landy and colleagues.[112] Although the authors restricted use of a diagnostic GTT to women with a 1-hour GCT above 7.8 mmol/l, incidence of GDM ranged from 18% at 7.8 mmol/l to 100% at > 12.5 mmol/l. It may be inferred from this that a proportion of GDM cases may have been identified at lower thresholds on the GCT.

Hence the sensitivity of the screening method is often assumed rather than demonstrated to be 100%. This is more likely to be true the lower the screening threshold applied. (For example, we could expect a threshold of ≥ 7.2 mmol/l would yield closer to 100% sensitivity than ≥ 8.3 mmol/l.) O'Sullivan and co-workers[100] tested all women with a GTT regardless of their screening result, which meant that all GDM women in the population were identified. Although some studies have followed this approach, the majority have not.

A distinction must be made when considering the method of glucose measurement, as noted above. Plasma glucose gives higher readings than blood glucose measurement. The former can be converted to the latter by the following formula:

plasma glucose = whole blood glucose × 1.14

This formula, rather than that adjusted by C&C, has been retained here because this is the one used by WHO. Most modern studies have used plasma values. However, where whole blood glucose values have been used, their plasma equivalents are also given. This has been necessary predominantly in the case of random glucose tests.

UK current practice

A survey of UK obstetric units in 1999[1] found that the majority (89%) screened for GDM but with little consensus on the appropriate method both between and within centres. For example, 84% of units used more than one screening method. Risk factor screening was used by 67% of practices to select those for the diagnostic GTT; 57% used urine testing with the GTT; 36% used risk factor screening in combination with a biochemical test to determine those who should have a GTT; and 30% used a blood test as a prelude to the GTT. Of the diagnostic tests used, 79% used a 75-g GTT, 14% used a 50-g GTT while 1% used a test meal. Of the blood tests used, 43% used the RBG, 11% used RPG, 10% used the 50-g GCT, 8% used another test, 25% used a combination of the tests and 5% didn't indicate which they used. An earlier survey found a variety of screening practices for GDM in one District Health Authority in England.[113] Only 8 out of 18 hospitals operated a screening policy. Six did an RBG, one a fasting blood glucose (FBG) and one a GCT. The authors noted that the "diversity of approaches as to who, when and how to screen for gestational diabetes in our health region" was attributable to uncertainty as to whether it was worth screening for GDM at all. They further noted that the GCT "is the most thoroughly evaluated method of screening for gestational diabetes ... Although one centre in NE Thames has recently introduced this method, most are put off by its expense and inconvenience".

An Italian survey of 283 gynaecologists reported 53% (151/283) carried out screening with a glucose load. Of these, 36% (55/151) gave a 50-g GCT to all women, 17% (26/151) a 100-g GCT to all women and 40% (60/151) restricted the test to women with risk factors.[114] In a US survey, 98.5% of clinicians used the 50-g GCT.[90]

Risk factor screening

There have been several studies of subgroups of women at higher risk of GDM, with a view to having selective rather than universal screening. Common risk factors for GDM include advanced maternal age (women > 40 years reported to have twice the risk of those aged 25–29 years), family history of diabetes, non-white ethnic origin, obesity, increased weight gain in early adulthood and current smoker.[89] Dornhorst and co-workers[115] found the following rates of GDM by ethnicity: white 0.4%; black 1.5%; South-East Asian 3.5%; and Indian 4.4%.

Using risk factors alone as a screening test, low sensitivities and specificities have been reported. For example, Marquette and co-workers[116] reported 50% sensitivity, 58% specificity and a PPV of only 3%. Helton and colleagues[117] reported 69% sensitivity, 68% specificity and PPV of 5%. Sacks and co-workers[89] found a prevalence in patients with risk factors of 4.2% (including > 25 years) compared to 0.4% in those without risk factors. (GDM was defined by GCT ≥ 7.5 mmol/l and positive GTT.) Weeks and co-workers[118] found that selective screening of patients using risk factors of obesity, glycosuria, family history of diabetes, or previous macrosomic, stillborn or anomalous fetus would have missed 43% of women with GDM. An Australian study in which all women were given the 75 g GTT (n = 1185), found 39% of women with GDM had no historic risk factors and would therefore have been missed by selective screening.[119] Three studies concluded that using historic and clinical risk factors to detect those at risk from GDM would mean that only about 50% of women with GDM would be identified.[100,116,120]

Advancing maternal age increases the risk of GDM. Rodriguez and colleagues[121] found a prevalence of nil under 18 years (based on 1-hour 50-g GCT with screens ≥ 7.8 mmol/l having a GTT; n = 437). In a mixed population (32% white; 39% Mexican–American; 22% black; 7% others), Truscello and co-workers[122] found a prevalence of 1.4% (2/137) in teenagers (12–18 years). The same population had a high incidence of macrosomia (8.7%) and 54% had excessive maternal weight gain. A 1.7% prevalence in adolescents < 19 years was reported by Khine and co-workers.[95] Coustan and colleagues[4] found a prevalence of 0.7% under 20 years of age. Court and co-workers[123] found a prevalence of 1.9% in women aged 15–24 years. O'Sullivan and co-workers[100] found a prevalence of GDM of 4.3% above age 25 while in those below age 25 the prevalence was 0.4%; 88% of their gestational diabetics were aged 25 or older. Marquette and colleagues[124] found a prevalence of 0.8% below 24 years compared with 5.5% at age 24 and above. A 1985 study[116] found a prevalence of GDM of 4.4% aged 24 or over compared with 0.77% below age 24; Sacks and colleagues[89] found that only 12% of women with GDM were aged under 25 years.

However, of the above studies, only the 1973 O'Sullivan study gave all women a diagnostic test.[100] The authors found a sensitivity of 63% and specificity of 56% using clinical history (birth of a baby ≥ 9 lbs, history in two or more pregnancies of fetal death, neonatal death, congenital anomaly, prematurity, excessive weight gain, hypertension or proteinuria, or family history of diabetes) as a screening test. When age > 25 years was added to risk factors, sensitivity was 69% and specificity was 35%.

BMI in early pregnancy is associated with infant birth weight and maternal fasting glucose in the third trimester,[125] and with cord insulin levels. The same correlation with cord insulin was seen with fat mass as estimated by skinfold measurements.

Risk factors are common. A study in Spain in 2000 by Jimenez-Moleon and co-workers[126] reported that 54% of women had one risk factor. Davey and Hamblin[127] found that four risk factors (age, BMI, ethnic group and family history) gave most of the information, and that adding other items such as a previous history of GDM and glycosuria added little – selective screening using the four criteria would have missed only two cases. Risk factors, including age, have also been used in combination with blood tests (see below).

The risk factors mentioned above are all prenatal ones. More recently, Edwards and co-workers[55] have studied the effect of weight gain during pregnancy, and found that birth weight of the baby correlates with maternal weight gain during pregnancy. Weight gain of more than 16 kg carried a higher risk of having an infant weighing over 4000 g. Obese (BMI over 29) women had a GDM prevalence of 17%, compared to 3% in those of normal (BMI 20–26) weight. Sixteen per cent of babies born to obese mothers had birth weights over 4000 g compared to 3% of those born to non-obese mothers.

Onyeije and Divon[128] found that early weight gain in pregnancy increased the risk of macrosomia. Women with GDM who had infants with birth weights under 4000 g gained less weight in each trimester than GDM mothers who had macrosomic infants (e.g. in first trimester 4.8 kg versus 8.2 kg).

A Brazilian study in 1997 by Branchtein and colleagues[129] found that waist-to-hip ratio and waist circumference were associated with glucose intolerance. The authors measured these at 21–28 weeks and concluded that central fat distribution predicts IGT. Skinfold thickness was also a predictor, suggesting that both obesity and its distribution are important. Newman[130] noted that maternal BMI was a strong predictor – no macrosomic infants were born to mothers with a BMI < 25.

Studies from previous decades now need to be interpreted with caution, because the prevalence of obesity and overweight has been increasing, particularly in children. A UK study in 2001 by Bundred and co-workers[131] found that the proportion of overweight children in the Wirral Health Authority area in north-west England rose from 15% in 1989 to 24% in 1998; the proportions who were obese rose from 5.4% to 9.2%. As these children reach reproductive ages, the prevalence of GDM in younger age groups may rise.

Risk factors may provide an indication of abnormal maternal metabolism even in women whose subsequent OGTTs are normal. Ford and colleagues[132] carried out OGTTs in women who had one or more risk factors for GDM, such as obesity, a previous baby over 4000 g or a family history of diabetes, and then studied outcomes in the group with normal OGTTs compared to women with none of these risk factors. The OGTT group had larger babies than the reference group, although this was seen more in the overweight ones.

Urine testing

Testing for the presence of glucose in the urine (or glycosuria) as a predictor of GDM has been superseded by the use of blood tests.[133] However, urine is routinely collected during pregnancy for purposes other than the detection of GDM, and testing for glycosuria continues. It is likely that this convenience and the associated very small marginal cost are the reasons for the continued use of urine testing, rather than its perceived efficacy. Where urine testing is recommended it is usually in association with blood testing.[133]

Hooper[134] used a threshold of ≥ 1.7 mmol/l (tested at routine antenatal visits) and reported 36% sensitivity, 98% specificity and PPV of 27%. Watson[135] found that 73% (16/22) of women with GDM did not have glycosuria and sensitivity was low at 27%, while specificity was 84% and PPV 7%. Gribble and colleagues[133] had an even lower sensitivity of 7%, with specificity 98% and PPV 13%. (Both these studies used a GCT ≥ 7.8 mmol/l + GTT as comparator.) Glycosuria is common in pregnant women unaffected by GDM.[136]

Blood tests

These include the measurement of glucose in the blood or plasma with or without administration of a glucose load, and measurement of fructosamine and HbA_{1c} levels. Thresholds of the various tests are debated, as are the amount and constituent of any glucose load and whether or not glucose testing should be preceded by a period of fasting.

A study in 2000 examined the change in blood glucose levels over time from sampling to testing.[137] One group had their plasma analysed immediately while the other had the usual delay for testing, during which the sample was kept at room temperature. Mean readings were 6.75 mmol/l for the former and 6.5 mmol/l for the latter. The authors found 22% (35/158) of women screened positive after the first sample were misclassified as negative after the second sample.

RPG

The RPG measures non-fasting glucose level. Strictly speaking it is not random, and some suggest that it should be called a casual blood glucose. No glucose load is given, and it is not taken at any fixed time after meals. Analysis can be either plasma glucose (RPG) or whole blood glucose (RBG). (The distinction is made more confusing by studies that call their test RBG when in fact it is plasma glucose that is measured.) According to the survey reported above, the RBG is the most popular biochemical test in practice in the UK. Arguments in its favour are cost and convenience compared to a glucose-loaded test. Some clinics take into account time since last meal.

As with urine testing, antenatal blood tests are also routinely required during pregnancy, hence as Lind[138] found, the cost of the needle and the syringe was shared with several other determinations.

Neilson and co-workers[139] did not recommend use of the RBG after a threshold ≥ 6.1 mmol/l detected only one in six cases of GDM, with sensitivity 17%, specificity 99% and PPV of 4.5%.

Although the paper lacks the data to confirm these calculations, it is reported that all women with an RBG ≥ 4.0 mmol/l were given a GTT.

Jowett and colleagues[140] found a wide variation in sensitivities with venous RPG at different times during the day, with the highest sensitivities recorded at 15.00 hours. Using a threshold of 5.6 mmol/l, sensitivity and specificity ranged from 29% to 80% and 74% to 80%, respectively. At a higher threshold of 6.0 mmol/l, sensitivity and specificity ranged from 41% to 58% and 74% to 96%, respectively. (Sensitivities and specificities have been recalculated because the authors' calculations were based on incorrect formulae.) Nasrat and co-workers[61] used the RPG together with a threshold of ≥ 7.0 mmol/l if women had eaten within 2 hours and a threshold of ≥ 6.4 mmol/l if they had eaten more than 2 hours previously. Sensitivity was 16%, specificity 96% and PPV 47%. No data are available in the study to confirm these figures but 91% of the women were given a 2-hour 75-g GTT. Using RPG, McElduff and co-workers[141] found a sensitivity of 46% and PPV of 12% for the RPG at a threshold of 6.1 mmol/l.

Each of these studies drew a distinction between women who had eaten within 2 hours of the test and adjusted the threshold upward accordingly (with the exception of McElduff and co-workers, who excluded women who had not eaten in the last 2 hours).

Table 5 illustrates equivalent plasma and whole blood levels used in these studies. Plasma (or equivalent) values ranged from 6.0 to 7.3 mmol/l if women had eaten within 2 hours and 5.6 to 6.6 mmol/l otherwise.

FPG

This should be measured after a period of fasting, usually overnight. The required period of fasting is nevertheless not stated in any of the following studies. The FPG has either been used on its own as a screening test or in conjunction with risk factors.

Agarwal and co-workers[142] found the FPG using a threshold of ≥ 5.3 mmol/l had 48% sensitivity and 97.5% specificity. A lower threshold (≥ 4.3 mmol/l) had 93% sensitivity and 38.5% specificity. However, these values are probably only applicable to the local (United Arab Emirates) population.

A Brazilian study by Reichelt and co-workers[143] examined a range of thresholds (4.5–4.9 mmol/l in 0.1 mmol/l intervals) for the FPG and found that a value of ≥ 4.9 mmol/l maximised both sensitivity at 88% and specificity at 78%, but PPV was only 1.3%. Although all women were given the diagnostic (2-hour 75-g) GTT, the population consisted of women ≥ 20 years, 45% white, 41% mixed and 14% black. Perucchini and co-workers[104] found a threshold of ≥ 4.8 mmol/l gave a sensitivity of 81% and a specificity of 76%.

The advantages of FPG are reported to include that it is not affected by gestational age, is similar in different ethnic groups, and that it has less variability and better reproducibility.[104]

In Edinburgh, Jones and Walker[144] found that to identify all women with abnormal 2-hour OGTT levels, the threshold for the FPG would need to be 4.2 mmol/l in early pregnancy and 3.9 mmol/l at 28 weeks. Perucchini and colleagues[104] concluded that the best cut-off for FPG was 4.8 mmol/l, which gave a sensitivity of 81% and a specificity of 76%

TABLE 5 *Equivalent blood and plasma levels*

	Plasma value (mmol/l)	**Equivalent whole blood (mmol/l)**
Jowett *et al.*, 1987[140]	5.6 (> 2 hours eating) 6.0 (< 2 hours eating)	4.9 5.3
Nasrat *et al.*, 1988[61]	6.4 (> 2 hours eating) 7.0 (< 2 hours eating)	5.6 6.1
McElduff *et al.*, 1994[141]	6.1 (< 2 hours eating) > 2 hours eating	5.3 N/A
	Whole blood (mmol/l)	**Equivalent plasma (mmol/l)**
Lind, 1985[138]	5.8 (> 2 hours eating) 6.4 (< 2 hours eating)	6.6 7.3
Nielsen *et al.*, 1988[139]	5.6 (> 2 hours eating) 6.1 (< 2 hours eating)	6.4 7.0

(using a 3-hour 100-g OGTT as gold standard). Rey[76] argued that the FPG had advantages of ease of administration, low cost, reliability and reproducibility, but that it needed to be validated in different populations, and that it had to be shown that screening by FPG lowered adverse outcomes.

FPG and risk factors

A 1992 study by Sacks and colleagues[145] in a population with one or more risk factors (> 25 years, with family history of diabetes, or obesity), found the FPG, using a constructed receiver operating characteristic (ROC) curve (which assumed all GDM detected by GCT ≥ 7.5 mmol/l + GTT), had 80% sensitivity for a threshold of ≥ 4.9 mmol/l.

The 1-hour 50-g GCT

The 1-hour oral GCT developed by O'Sullivan and colleagues[100] is the most evaluated screening test in the literature. A 50 g glucose load is taken orally with 150 ml of fluid, with blood glucose measured after 1 hour. The most frequent and effective timing of this screening test is between 24 and 28 weeks' gestation. The methodological problem with most studies noted above is the assumption that the GCT is 100% sensitive, especially at lower (~7.2 mmol/l) threshold levels. A distinction between whole blood and plasma levels was not deemed necessary because virtually all studies used plasma values.

The seminal work by O'Sullivan and colleagues in 1973,[100] one of a handful of studies that applied the diagnostic GTT to all patients and not merely the positive screenees, found a sensitivity of 79%, specificity of 87% and PPV of 15%. The threshold used was equivalent to ≥ 7.9 mmol/l using modern glucose measuring methods. McElduff and co-workers,[141] in a side-by-side comparison of the GCT and FPG, found the GCT (threshold ≥ 7.8 mmol/l) to be less than 100% sensitive at 86%. Four women, diagnosed from positive FPGs, had negative GCT screens.

Those studies assuming 100% sensitivity, using a threshold of ≥ 7.2 mmol/l, have reported specificities of 73–74% and PPVs of 10–22%.[146] With assumed sensitivity of 100%, using a threshold of ≥ 7.8 mmol/l, specificities of 83–95% and PPVs of 15–22% have been reported.[147]

Marquette and co-workers[124] found a threshold of ≥ 8.3 mmol/l gave a sensitivity of 96% and PPV of 24% based on screens ≥ 7.2 mmol/l having a GTT. Lavin[148] reported the same PPV at ≥ 8.3 mmol/l. A study in a high prevalence group (European and Asian, 6.6%) by Yalcin and Zorlu,[146] in which women with screens ≥ 7.2 mmol/l were given a GTT, reported sensitivities of 98% and 88%, specificities of 44% and 85%, and PPVs of 32% and 9%, using thresholds of ≥ 7.5 and ≥ 7.8 mmol/l, respectively.

Meriggi and co-workers,[149] in a selected population, contrasted the GCT using different blood and plasma thresholds. The plasma levels were lower but yielded similar results. At the plasma threshold of ≥ 7.5 mmol/l, sensitivity was 100%, specificity 80% and PPV 21%. For a blood glucose level to deliver 100% sensitivity, 85% specificity and 27% PPV, the threshold had to be 8.6 mmol/l (*Table 6*). ROC analysis was used to determine these optimal cut-offs, although the model may have been skewed in that only positive screenees received a GTT. (ROC analysis is used to trade off sensitivity with specificity.)

TABLE 6 *Data on comparative plasma and blood thresholds*[144]

	Sensitivity	Specificity	PPV
Plasma			
≥ 7.5 mmol/l	100	80	21
≥ 7.8 mmol/l	96	84	25
≥ 8.0 mmol/l	82	88	27
≥ 8.3 mmol/l	74	90	29
Blood			
≥ 8.6 mmol/l	100	85	27
≥ 8.9 mmol/l	89	87	26
≥ 9.1 mmol/l	82	90	31
≥ 9.4 mmol/l	70	91	31

The majority of GCT studies, apart from a few of the early ones, used plasma values, but the above illustrates similar results could be obtained through whole blood samples with higher values.

The effect of fasting status on the GCT

The Third International Workshop-Conference on Gestational Diabetes Mellitus recommended that no regard for time of last meal on the 50-g GCT was necessary.[150] However, demonstrations that time since last meal does influence the response to a GCT have raised concerns about this recommendation.[151] Lewis and co-workers[152] examined the effect of prior meal on the glucose response to the GCT. In a controlled clinical trial, 10 women with GDM and 12 controls underwent three 50-g GCTs over a 2-week period. Screening was undertaken in the fasting state, 1 hour after a test meal (500 kcal with 50% carbohydrate), and 2 hours after the test meal. All women subsequently underwent a 3-hour 100-g GTT. In the control group, the 1-hour plasma glucose level was significantly higher in the fasting state than in the 1- or 2-hour state, leading to a 58% false-positive rate in the fasting state. In the

GDM women, the 1-hour plasma glucose level was similar in the fasting and 1-hour test, but significantly lower than the 2-hour test. Despite the use of a small sample, the results suggest that the GCT is affected by the time of last meal.

Two further studies support this conclusion. A study of a Turkish population by Cetin and Cetin[153] grouped women according to time since last meal: < 2 hours, 2–3 hours, > 3 hours. Using a threshold of 7.8 mmol/l as a reference their 'suggested' values (all above 7.8 mmol/l), which varied by time since last meal, were said to improve efficiency by reducing the number of GTTs required. (A methodology that would depend on the cost of a GTT.) Details of the basis for selection of 'suggested' values were not stated, nor how the balance between sensitivity and specificity was achieved. Higher specificity has not sacrificed sensitivity in two of the groups but in the third, < 2 hours since last meal, 63% sensitivity and 91% specificity has been judged superior to 75% sensitivity and 86% specificity.

Sermer and co-workers[31] followed the same procedure but applied it more robustly using ROC analysis. Again, the authors used the 7.8 mmol/l value as the reference point at each of four time periods since the last meal. In each case, they found higher thresholds maximised the area under the ROC curve (areas under ROC curve ranged from 0.74 to 0.81). In addition, 90% of women in this study were given a GTT.

Table 7 shows the trade-offs between sensitivity and specificity made to optimise the area under the ROC curve.

Intra-day timing for GCT screening

Kirkpatrick and co-workers,[154] who tested their patients (n = 1511) throughout the day, found a significant increase in blood glucose levels after 11 am (5.9 vs 5.5 mmol/l, $p < 0.0001$).

Optimal gestation for GCT screening

Watson[155] conducted screening tests at three intervals during pregnancy. Assuming all women with GDM were detected (GCT ≥ 7.8 mmol/l), specificity was 79% and PPV 20% when screening at 20, 28 and 34 weeks. Screening at 20 weeks gave 33% sensitivity, 95% specificity and a PPV of 26%. Screening at 20 and 28 weeks improved sensitivity to 89% and specificity to 87% with a PPV of 26%. In total, 33% (9/27) of identified cases of GDM were detected by 20-week GCT, 56% (15/27) at 28 weeks after negative screen at 20 weeks, and 11% (3/27) at 34 weeks after negative screen at 28 weeks. Berkowitz and colleagues[99] detected 29% of women with GDM prior to 24 weeks and 71% after 24 weeks.

Jovanovic and Peterson[156] tested women at three time periods: 9–20 weeks, 27–31 weeks and 33–36 weeks. They assumed that all women with GDM were identified during this process, in which negative screens (using a ≥ 8.3 mmol/l threshold) were not given a GTT. The 33–36-week period therefore had 100% sensitivity. Using this period as comparator, the 9–20-week period produced 5% sensitivity and PPV 11% while the 9–20 and 27–31-week periods together produced 67% sensitivity and PPV of 6.9%.

Nahum and Huffaker[157] concluded that patients with a screen of ≥ 7.8 mmol/l during the first trimester are at a high risk. The PPV of a first trimester screen ≥ 7.5 mmol/l leading to a third trimester screen of ≥ 7.5 mmol/l was 65%. The conclusion cannot be supported by the data available in the paper. From a sample of 208 predominantly white women, the authors detected 9 cases of GDM in the first trimester, 8 in the second and 65 in the third.

A study in a 'high-risk' fasting population (7.9% prevalence), all of whom were given a GTT, found

TABLE 7 *Optimal thresholds to optimise the area under the ROC curve [reproduced (in part) from Sermer* et al., *1994[31]]*

Time since last meal	Threshold (mmol/l)	Area under ROC curve	Sensitivity (%)	Specificity (%)	PPV (%)
< 1 hour	7.8	0.82	81.6	82.7	15.7
	8.2	0.84	78.9	88.1	20.8
1–2 hours	7.8	0.76	65.8	85.6	15.2
	8.2	0.78	65.8	91.0	22.3
2–3 hours	7.8	0.79	73.7	84.5	15.8
	7.9	0.80	73.7	86.5	17.7
> 3 hours	7.8	0.82	89.5	73.6	11.8
	8.3	0.83	84.2	81.6	15.3

that 63% of women with GDM at third trimester (27+ weeks) would have gone undetected if screening was only performed at second trimester (14–26 weeks).[158] During the third trimester sensitivities were both 88%, and specificities 73% and 82% at ≥ 7.8 and ≥ 8.3 mmol/l for the second and third trimesters, respectively.

In another high-risk population with ≥ 1 risk factor (20% prevalence), Super and co-workers[159] found 91% sensitivity and 88% specificity during the first trimester (using a ≥ 7.2 mmol/l threshold) compared to 85% sensitivity and 79% specificity (using a ≥ 6.86 mmol/l threshold) in the second trimester (optimal values calculated by ROC analysis). However, all women were not tested again in the third trimester as a comparator.

Meyer and colleagues[160] compared early (mean 15 weeks) versus late (28 weeks) screening (≥ 7.5 mmol/l) in a majority Hispanic population (52% Hispanic, 38% black, 10% white) and found a prevalence of 2.4% early (8 GDM) and 3.7% late (a further 12 GDM).

At 16–20 weeks, using a threshold of ≥ 7.5 mmol/l, Mello and co-workers[96] reported 64% sensitivity, 87% specificity and PPV of 40% in a high-risk (11.7%) Italian population.

GCT thresholds for the diagnosis of GDM

Several studies have attempted to determine diagnosis of GDM through an elevated value for the GCT screen. Landy and co-workers[112] found a threshold of ≥ 10.3 mmol/l yielded 38% sensitivity, 96% specificity and PPV of 79%. All women (n = 16) with screen thresholds ≥ 12.5 mmol/l were diagnosed to have GDM. ROC analysis was used to predict this optimal threshold point but only positive screens were given a GTT. Bobrowski and co-workers[101] recommended that a threshold of ≥ 12.2 mmol/l be diagnostic. The highest positive screen in their study with a negative diagnosis was 12 mmol/l, while all 27 patients ≥ 12.2 mmol/l were diagnosed with GDM. (The composition of this study was 67% African–American with an incidence of 29–33%.)

Shivvers and Lucas[161] found that a GCT screening threshold of ≥ 11.1 mmol/l was not always confirmed as GDM on OGTT. Of 59 evaluable patients, 11 were normal on the GTT and 48 had GDM. In a predominantly black, Hispanic and Indian population, Hong and co-workers[103] found all 16 women with a screen above 11.1 mmol/l to have GDM.

Inconvenience and harms of the GCT

Menon and co-workers[162] reported the GCT inconvenient. A study in Toronto by Kerbel and co-workers[163] found false-positives had a decreased perception of their own health.

A more physiologic solution

Schwartz and co-workers[97] proposed a modified GCT in 450 ml water as opposed to the standard 150 ml. This more palatable, diluted solution was at least as sensitive and specific as the standard solution – there was only one significant difference in ten comparisons. Nausea or vomiting was significantly more common in the standard mix, 11% versus 3% ($p < 0.05$); including other side-effects of dizziness and sweatiness gave 14% versus 3%, and 90% of the women preferred the taste of the modified solution and none complained about its volume. (This study composed a high-risk population; 81% were Mexican–Indians in which the incidence was 10.6%.)

GCT combined with risk factors (including maternal age)

In a study of Oriental women with at least one risk factor, and in which all women were given the GTT, Jirapinyo and colleagues[164] found a sensitivity of 86% and specificity of 65% leading to a PPV of 23% (GCT threshold of ≥ 7.8 mmol/l). Raising the threshold reduced sensitivity to 83% but increased specificity to 79% and PPV to 32% (≥ 8.3 mmol/l).

Several studies have incorporated maternal age as a screening factor prior to administering the GCT. Only one of these gave everyone a GTT.[100] In a population aged ≥ 25 years, using a threshold (recalculated) of ≥ 7.9 mmol/l, sensitivity was 88%, specificity 82% and PPV 19%.

Two studies by Marquette and co-workers[116,124] found sensitivities of 92% and 88%, respectively, for women aged ≥ 24 years using a threshold of 7.2 mmol/l. PPVs were 14% and 5%. With a threshold of ≥ 8.3 mmol/l, sensitivity was 88% and PPV 31%.[124] For the same age threshold, another study reported sensitivity of 95%, while sensitivity was 86% for women aged ≥ 28 years and 62% for women aged ≥ 30 years (all ≥ 8.3 mmol/l, non-fasting).[165]

Combining age thresholds and other risk factors has given reasonable sensitivities. O'Sullivan and colleagues[100] also considered clinical history together with maternal age and the GCT. They found a sensitivity of 53% and specificity of 93% in their population of women with a clinical history who were given a GCT (≥ 7.9 mmol/l).

When screening was confined to women aged ≥ 25 years, sensitivity rose to 62% and specificity fell to 80% with a PPV of 19%.

Coustan and co-workers[4] used age as their most significant risk factor but included younger women with risk factors. Using a GCT threshold of 7.2 mmol/l for women aged ≥ 25 years (or younger if they had a positive risk factor), sensitivity was 85% with 7.2 mmol/l cut-off and 95% with 7.8 mmol/l cut-off. For women aged ≥ 30 years it was 65%.

Reliability

Sacks and co-workers[166] tested (predominantly Hispanic) women on two consecutive days using a threshold of ≥ 7.5 mmol/l. They found reliance on a single test would have missed 27% (8/30) women with GDM and concluded that the 1-hour 50-g GCT was only "moderately reproducible".

The 2-hour 50-g GCT

Weiner and co-workers[167] examined the efficacy of a 2-hour GCT (tested at 2 hours post-glucose load rather than the normal 1 hour). Adjusting the 2-hour threshold to identify all women with GDM, they found sensitivities of 83% and 100%, specificities of 87% and 85% and PPVs of 16% and 15% for the 1-hour and 2-hour tests, respectively. (Diagnosis was based on women with 1-hour GCT ≥ 7.8 mmol/l or 2-hour GCT ≥ 6.5 mmol/l receiving a GTT.)

In a separate sample (n = 185), designed to validate the above result, the authors found sensitivities of 75% and 75%, specificities of 82% and 92% and PPVs of 8% and 17% for the 1-hour and 2-hour tests, respectively. They continued the ≥ 7.8 mmol/l threshold for the 1-hour test and applied the ≥ 6.5 mmol/l threshold for the 2-hour test calculated as above.

Glucose polymer load

Several studies have examined the use of glucose polymers rather than the standard monomer glucose load. A study by Court and co-workers[123] in which all women received a 3-hour 100-g GTT found a 100 g glucose polymer (threshold ≥ 8.0 mmol/l) load to have sensitivity of 88%, specificity 80% and PPV 25%.

Studies that have directly compared polymer against monomer glucose load are considered on page 24.[168]

Jelly beans

Lamar and co-workers[169] used 50 g of jelly beans as an alternative to the 50-g GCT. Women were given the GTT and the results are summarised in the comparisons section on page 23.

Breakfast test meal

Coustan and colleagues[170] evaluated screening with the breakfast test meal. The test meal outperformed the GCT in ROC curve analysis. Using a threshold of ≥ 6.7 mmol/l, sensitivity was 75% and specificity 94%. Lowering the threshold to ≥ 5.6 mmol/l led to 96% sensitivity and 74% specificity. However, the content of the meal would be considered nutritionally unsatisfactory by today's standards, as 47% of the calories came from fat.

Fructosamine

A number of studies have examined the level of fructosamine in the blood as a screening test, although many of these were in local ethnic groups.

Despite the attention in the literature, the fructosamine test is insensitive and therefore of little or no value as a screening test. Sensitivity was 0% and specificity 100%. Studies reported fructosamine levels of women with GDM to be within normal limits.[171]

Only one study, in Saudi Arabia, found that 3/6 of the women with GDM identified by a 100-g GTT had fructosamine levels above normal limits.[172] Sensitivity was 50% and specificity 90%.

Fructosamine corrected for protein

Hughes and colleagues[173] used a 'second generation' fructosamine test corrected for total protein. They found 79% sensitivity and 77% specificity (based on those with GDM detected by 50-g GCT ≥ 7.8 mmol/l). The population comprised 67% Arab and 23% Indo-Pakistan. An uncorrected fructosamine test on the same population would have been "less good". However, a study by Bor and co-workers[174] in Turkey produced conflicting results. Of 12/54 women diagnosed with GDM (by 100-g GTT), only 1/12 had a fructosamine (corrected for protein) above normal range (≥ 2.8 mmol/l).

Glycosylated haemoglobin

Glycosylated haemoglobin (HbA_{1C}) is used in diabetes as an index of long-term diabetic control. It has been tested as a potential screen for GDM. Reports show, however, that it is insensitive.[161]

Other tests

Fetal abdominal transverse diameter

Grandjean and co-workers[175] measured fetal abdominal transverse diameter (FATD) to attempt to predict GDM. In a population of women with traditional risk factors, all of whom had a 100-g GTT, using an FATD above the 95th percentile produced 57% sensitivity, 77% specificity and a

PPV of 36%. Screening by FATD would therefore have missed 40% of women with GDM.

However, in women already known to have GDM but with fasting glucose level under 5.8 mmol/l, ultrasound may be of use in identifying those in whom insulin treatment can reduce macrosomia.[70]

Direct comparisons between screening tests

Jelly beans versus GCT

A 1999 study gave women both a 50 g jelly bean load and a 50-g GCT.[169] All women went on to a 3-hour 100-g GTT (excluding 24 dropouts) regardless of screening result and 5/112 (3.7%) women were found to have GDM. Of these, 3/5 were identified by the GCT only, one by jelly beans only and one by both tests. No significant difference in screening performance was reported, probably because of the small numbers.

This study is distinctive for its high proportion of side-effects. Nausea, dizziness or headaches were suffered by 38% of women having the GCT and 20% given jelly beans. The authors noted that women preferred the jelly beans.[169]

Risk factors versus GCT

As noted in the main text, O'Sullivan and co-workers[100] found lower sensitivity using clinical history and women aged ≥ 25 years than with the GCT alone. When the GCT was given only to those with clinical history and aged ≥ 25, sensitivity was further reduced because all women with GDM in this group were missed.

Jovanovic and Peterson[155] reported only 1 of 19 women with GDM would have been detected by risk factor screening rather than the GCT.

Griffin and co-workers[176] found a higher prevalence using the GCT (threshold ≥ 7.8 mmol/l) rather than risk factors as a screening test. Incidence of GDM when using risk factors was 1.45% compared with 2.7% when using the GCT. All positive screens – those with risk factors or a GCT ≥ 7.8 mmol/l – were given a GTT.

A study in 1987, using a split sample (n = 2000), found double the prevalence in a risk factor screened group (4.2% versus 2.1%) compared to a GCT screened group.[177] However, the GCT threshold used was ≥ 8.3 mmol/l. These were middle/upper class patients, and no baseline comparison of characteristics between the two groups was given.

GCT versus 100 g carbohydrate load

A direct comparison in 1998 (Rust and co-workers[178]) found the 1-hour 50-g GCT to be superior to a 2-hour 100 g carbohydrate load. Assuming 100% sensitivity with a ≥ 7.8 mmol/l threshold for a GTT, the GCT had sensitivity of 100%, specificity of 95% and PPV of 43%. This compares with the carbohydrate load having roughly similar specificity of 98% and PPV of 40%, but a much lower sensitivity of 25%. The study population was 82% black, 16% white.

An observational study in 1999 by Bevier and colleagues[179] noted that women who had GDM by GCT but not by 100-g OGTT nevertheless had higher infant morbidity, reducible by intervention with dietary advice and counselling (although the study was not analysed by intention to treat – there was some crossover from control to intervention, with an end of study numerical imbalance).

Routine versus selective screening

A 1999 study in British Columbia by Bebbington and co-workers[180] randomised low-risk women to routine GCT or to testing for 'selected indicators'. There were no differences in birth weights or proportions with macrosomia (12% versus 11.5%). Similarly, a study in Texas by Casey and colleagues[181] saw no difference before and after a shift from selective screening (31% of women got GCTs) to universal (88% had GCTs). This was a large study with 13,000 deliveries in each period. Macrosomic percentages were 8.3% and 8.6%.

Fructosamine versus GCT

Two studies have compared a fructosamine test with a 50-g GCT. A study in Turkey gave all women a GTT but had only a small number of patients (n = 42).[182] This must also be considered a high-risk group because prevalence of GDM was 33% (14/42). Fructosamine identified 10/14 while the GCT identified 11/14, a non-significant difference. Sensitivity of fructosamine was 71%, specificity 44% and PPV 40%. The other study, in Spain,[183] reported fructosamine to have 8% (4/48) sensitivity and 100% specificity while the GCT had 100% sensitivity. However, the reference point was the GCT, because only those with a GCT threshold ≥ 7.8 mmol/l went on to have a GTT.

RPG versus GCT

McElduff and co-workers[141] concluded after a side-by-side comparison of the RPG and GCT that the GCT detected significantly more of the women with GDM. In their study the GCT (threshold GCT ≥ 7.8 mmol/l, RPG ≥ 6.1 mmol/l) detected significantly more of the women with GDM (24 out of

28 versus 13 out of 28 for the RPG). The GCT had an 8.8% false-positive rate and the RPG 13.4%.

Mathai and co-workers[184] studied an Indian population (prevalence 5%) and found neither the RPG nor GCT were able to provide high sensitivity or specificity in a study in which all women were given a GTT. The RPG (threshold ≥ 5.0 mmol/l) had 63% sensitivity and 66% specificity while the GCT (≥ 6.4 mmol/l) had similarly 63% sensitivity and 55% specificity.

FPG versus GCT

Perucchini and co-workers[104] gave all women a GTT in a comparison of the FPG with a 50-g GCT. Using a ≥ 4.8 mmol/l threshold, the FPG had 81% sensitivity and 76% specificity. Lowering the threshold to ≥ 4.4 mmol/l gave 100% sensitivity and 39% specificity. At best the GCT had 68% sensitivity and 82% specificity (using a threshold of ≥ 7.0 mmol/l). However, after a 2-hour fast (threshold ≥ 7.0 mmol/l) the GCT had 100% sensitivity and 71% specificity.

Fuhrmann[185] used whole blood rather than plasma in comparing FBG with a 1-hour and 2-hour GCT in a low incidence (1.1%) population in Germany. All women had a GTT and 28 women were diagnosed with GDM. Of the thresholds and tests examined, only the FBG ≥ 4.4 mmol/l identified all 28 cases of GDM (*Table 8*). The plasma equivalents of whole blood levels used throughout this study are calculated in the table.

A study in Hong Kong by Tam and co-workers[186] compared the 1-hour GCT with FPG and concluded that FPG has as good a predictive value as GCT, with similar areas under the ROC curves, but was much more convenient and should be preferred.

Polymer versus standard glucose

Reece and co-workers[187] found a high level of agreement between the two tests: five patients with positive screens on both tests had positive GTTs. Another five patients had inconsistent responses to the two tests (all positive screens, either polymer or glucose, were given a GTT). Three had positive polymer screens and negative glucose screens, two had positive glucose screens and negative polymer screens. All were negative on the GTT. The authors reported 40% nausea with the GCT and 10% with the polymer.

A double-blind study in 1992 by Bergus and Murphy,[168] in which all women were given a 3-hour 100-g GTT, compared glucose monomer with glucose polymer. Sensitivity data were not individually reported but the polymer was better tolerated: 17/33 monomer and 9/33 polymer had symptoms including feelings of nausea, sickness, dizziness, tiredness, feeling bloated, headaches, vomiting, and abdominal discomfort.

Murphy and co-workers[188] compared 50 g polymer, 50-g GCT and a 50 g chocolate bar. (The results for the chocolate bar were not reported individually, however.) All patients (bar 16 unable to complete because of vomiting) were given a GTT and the polymer produced superior results to the GCT. The polymer had 100% sensitivity, 93% specificity and PPV of 49% compared to 33%, 74% and 9%, respectively for the GCT (≥ 7.5 mmol/l). However, these results are qualified by the fact that there were only three cases of GDM detected (and only one screened positive on the GCT).

Discussion and conclusions

The choice of criteria for a 100-g GTT will influence the prevalence, with the C&C recommendations giving higher results. However, the WHO-recommended 75 g is the usual diagnostic test adopted in the UK. Generally manageable rates of vomiting have been reported with either the 100 g or 75 g loads but it has been recom-

TABLE 8 *Data from Fuhrmann (1989)[184]*

Test	Whole blood threshold (mmol/l)	Equivalent plasma threshold (mmol/l)	Sensitivity (%) (no. of GDMs detected)	Specificity (%)	PPV (%)
FBG	4.40	4.70	100 (28/28)	74	4
	4.72	5.06	93 (26/28)	83	6
	5.00	5.38	86 (24/28)	89	7.5
	5.50	5.95	75 (21/28)	95	15
1-hour GCT	7.22	7.91	97 (27/28)	80	5
	7.77	8.54	97 (27/28)	90	10
2-hour GCT	6.38	6.96	75 (21/28)	93	10
	6.94	7.60	75 (21/28)	96	18

mended a 50-g GTT be substituted when necessary. However, no studies have looked at the magnitude of the loss in sensitivity resulting from this reduced glucose load.

The inconsistency of screening practices reported in the UK is evidence enough that either better evidence of best option, or guidelines to promote use of the best option, are needed. Some screening practices are easy to rule out on account of their poor sensitivity, while others warrant more careful consideration. The use of risk factors (including maternal age) has led to large numbers of diagnostic tests being performed but high proportions (up to 50%) of women with GDM being missed. A screening strategy may, however, depend on the ethnic mix of the local population; this may explain some of the reported inconsistent treatment across regional boundaries in the UK. Low incidences of GDM in women under 25 years have been generally reported in a number of studies. Age is proven as a risk factor, but used without a biochemical screening test has delivered disappointing results in individual studies as well as side-by-side comparisons. Urine testing has low sensitivity and specificity and is a poor screening test for GDM. The only argument for its continual use is the very low marginal cost. Jelly beans, which may cause less nausea, are unproven. A 100 g carbohydrate load is unnecessary.

Fructosamine and HbA_{1c} appear to have little value as screening tests. Breakfast test meals may be a realistic alternative to a GCT or random/fasting glucose test but need to have carefully defined content; at present it would appear these are designated by the particular centre at which they are used. FATD has only been used in one study and did not perform well, and would depend on whether ultrasound testing was done specially, or was routine (and if so at what gestation).

The RPG/FPG and GCT biochemical tests have delivered the highest sensitivities and specificities in the literature. There is no evidence that whole blood or plasma testing is superior, as long as the threshold levels appropriate to each are used. The papers on RPG generally report poorer sensitivities. Studies on the FPG are more mixed, although in mixed ethnic populations. Individual studies on the GCT have reported high sensitivities and specificities, although many may be biased by the failure to give negative screenees a GTT.

The common difficulty when faced with comparing the efficacy of screening tests is having to trade off between sensitivity and specificity. Often one test or threshold has a higher sensitivity, the other higher specificity. Any trade-off depends on the cost or undesirability of errors (consequences of false-positives and false-negatives) and is increasingly difficult at the margin. One way of selecting the best threshold or test is ROC analysis, but unfortunately this has not been widely used in this topic area.

ROC curves plot one-minus-specificity against sensitivity and are commonly used for direct statistical comparison of two tests. Diagnostic accuracy is expressed as the area under the ROC curve (a value between 0.5 and 1.0). The higher the value, the closer the ROC plot is to perfect (= 1.0). For example, a value of 0.8 would mean that a randomly selected woman with GDM would be detected by that particular test 80% of the time. Unfortunately, the use of ROC analysis in the GDM literature has been limited in terms of comparisons of alternative tests. Where this has been attempted, results have been weakened by an assumption of 100% detection of women with GDM (i.e. only positive screens were given a GTT) and hence the comparison of one test with another is assumed to be 100% sensitive (usually the 1-hour 50-g GCT). Predictably, the comparator (the FPG in this case) has performed less well than the GCT.

A few studies have used ROC analysis to determine optimal thresholds for a particular biochemical screening test.[189] This is achieved through the identification of an operating point on the ROC plot at a specificity/sensitivity combination that maximises the function. However, to calculate this point ideally requires cost data, and probability of false-positives and false-negatives.

There is a substantial literature on the GCT, but even there uncertainties remain. Intra-day variations have been reported to influence glucose levels, as has delay between sampling and analysis. If further research were to acknowledge these factors as significant, neither would be easily remedied. A more dilute solution is advocated to overcome nausea. Reproducibility has been declared moderate and use in combination with risk factors (including maternal age) has delivered mixed results. Third trimester testing is clearly the best choice for the GCT but studies have shown the success of repeat testing during trimesters, although this is a more expensive and inconvenient option. Mixed results have been reported in comparisons between 1-hour and 2-hour GCTs. Polymer loads have seemed better than glucose but there are few studies available.

It is of concern that while GCT testing on the whole is done without regard to fasting status, several studies have suggested that time since last meal does affect glucose levels. Some RPG/RBG studies have enquired as to the time since last meal and adjusted their thresholds respectively.

Of the five studies directly comparing random or fasting glucose with a GCT, all but one gave all women a diagnostic GTT. Nonetheless, results are inconclusive. Only one paper showed the GCT to be clearly superior, detecting significantly more of the women with GDM than an RPG, but this study only gave those women with a positive GCT a diagnostic test.[142] Mathai and colleages[184] found very similar sensitivities for both GCT and RPG using ROC analysis, albeit in an Indian population. The other two studies compared fasting glucose. Fuhrmann[185] reported high sensitivity with the FBG but generally lower sensitivities in a number of comparisons with both 1-hour and 2-hour GCTs. Similarly, Perucchini and co-workers[104] reported higher sensitivity with the FPG but lower specificity than the 1-hour GCT. A 2-hour GCT on the other hand had higher sensitivity and lower specificity. The Mathai study,[184] which was inconclusive, and the Tam study,[186] which found equivalence, used ROC analysis to compare the two tests.

The optimal thresholds are difficult to determine from the literature. However, available evidence suggests a threshold for the FPG may have to be as low as 4.7 mmol/l.[104] The most commonly cited levels for the GCT are 7.2 and 7.8 mmol/l. ROC analysis by Meriggi and co-workers[149] has suggested a 7.5 mmol/l threshold while Sermer and co-workers[31] found values of 7.9–8.3 mmol/l, depending on time since last meal.

Any trial comparing FPG and GCT should consider these values. The trial should be multicentre, should use ROC analysis to directly compare screening tests and should include subgroup analyses especially relating to ethnicity.

Current guidelines

Guidelines produced by Diabetes UK (formerly the British Diabetic Association) for GDM screening suggest that:

- Urine should be tested for glucose at every antenatal visit (to ensure any women developing type 1 diabetes are not missed).
- RPG measurements should be made at booking, at 28 weeks' gestation and in the presence of glycosuria (however no threshold is given for the random glucose).
- A 75-g GTT should be performed if RPG concentrations are > 6.1 mmol/l in the fasting state or > 7.0 mmol/l within 2 hours of food.

However, Diabetes UK suggest that as there is no consensus on the screening for GDM, these guidelines attempt only to take a balanced approach that can be adapted locally and may change as new information becomes available. Using these screening guidelines will increase the cost of screening because of the repeat testing at both booking and 28 weeks' gestation. It is also interesting to note that despite the reported inadequacy of testing for glycosuria, urine testing is included in these guidelines, although this is more to ensure that women with type 1 diabetes are not missed.

The clinical practice recommendations of the ADA[190] recommend selective screening of pregnant women with one or more of the following:

- age > 25 years
- overweight before pregnancy
- member of an ethnic group with a high prevalence of GDM
- diabetes in first-degree relatives
- history of abnormal glucose tolerance
- history of poor obstetric outcome

Screening is either undertaken by the 50-g GCT (threshold 7.2 mmol/l) and then the 100-g GTT, or directly with the 100-g GTT. Using these guidelines reduces the cost compared to universal screening, but some women with GDM are likely to be missed.

The Scottish Intercollegiate Guidance Network recommend urine and RBG are tested at each routine antenatal visit. There have been no studies to determine the sensitivity and specificity of this RBG and urine testing combination. Again, the repeat testing is likely to reduce the cost-effectiveness of such a strategy.

Chapter 4

Economics of screening for GDM

Review of the literature

This review identified ten published papers on the costs of screening for GDM, but none from the UK: nine are from the USA, and one from Australia. These are presented in appendix 3.

Direct costs

Marquette and co-workers,[124] Lavin[148] and Swinker[191] produced similar results, giving the cost of a GCT at about one-quarter that of a GTT (US$11–18). A 1986 study calculated the GCT at US$7.25 and the GTT at US$64, including direct and indirect costs.[167] In 1991, a GCT was said to cost US$17.75 compared to a GTT at US$59.15, although it was unclear exactly what was included in these figures.[165] Most recent US data costs the GCT at US$25 and a GTT at US$50 in terms of laboratory costs.[192]

Outcome measurement

The only cost-effectiveness measurement was cost per case of GDM detected. (No study has attempted to estimate the cost-effectiveness of GDM screening taking into account adverse events.) The first estimate comes from a 1983 paper that only included materials and laboratory costs.[191] Cost per case of GDM detected is calculated as US$173, consisting of the 1-hour 50-g GCT and a 100-g GTT to those with a screen ≥ 130 mg/dl. A 1989 paper[4] using 1985 costs,[124] found the same scenario cost US$249. At an elevated cut-off point of 140 mg/dl or above, cost per GDM detected was US$866 and US$1215 in 1986 (two separate samples in the same paper; including direct and indirect costs), or US$222 in 1989 (1985 direct costs only). Raising the threshold again to 150 mg/dl, the cost per case of GDM detected was US$328.96 in 1985, US$699 (including direct and indirect costs) in 1986, US$722 in 1989 (1985 costs) and US$722.31 in 1991. Using a 2-hour threshold of 118 mg/dl or above with the GCT generated cost per case of GDM detected of US$622 and US$831 in 1986. When diagnosis utilised risk factor screening, cost per case of GDM detected was estimated at US$1805. Further estimates using age-based screening in combination with the GCT have been examined. Screening only those patients aged 25 years and over and using a threshold of 130 mg/dl and above gave a cost per case of GDM detected of US$215 or US$192 using a threshold of 140 mg/dl or above in 1989 (using 1985 costs). In an adolescent population (< 20 years), the cost per case of GDM detected using a GCT threshold of 140 mg/dl was US$2733. In this population, all women with GDM had a family history of diabetes, thus lowering the cost per case of GDM detected by risk factor screening to US$1258.

Conclusions

None of the above give UK costs, and a straight dollar to pound conversion would be unsafe. However, the results are still useful as a guide to relative costs. The main weakness is the lack of an outcome-based cost-effectiveness measure, such as cost per macrosomia avoided.

Cost of screening for GDM in the UK

Direct costs

These include the cost of materials used in the various screening and diagnostic tests, staff time in taking the tests, biochemical analysis, and counselling and dietary advice to those screened positive.

In the absence of any UK costing published studies or locally available data, the construction of the model has relied on costs data generated by the Scottish Health Purchasing Information Centre (SHPIC) through their costing unit at Dundee Teaching Hospitals Trust (DTHT), for an unpublished report on diabetic pregnancies. Where used, these costs have been inflated to current prices using the most recent indices available in Netten and Curtis.[193]

Local costs for materials, administration time and biochemical analysis were obtained by SHPIC from interviews with NHS staff at Aberdeen Royal Hospitals Trust (ARHT) in 1996/97. These (inflated to current prices) are shown in *Table 9*. The current price of a glucose load (Calsip) was obtained from *MIMS (Monthly Index of Medical Specialties)*.[194] This was £0.93 for 200 ml, consisting of 50 g carbohydrate, the correct load for a 50-g GCT. A 75-g and 100-g GTT are then calculated at £1.40 and £1.86,

TABLE 9 *Breakdown of costs associated with screening/diagnostic tests*

	Baseline	Source
Cost of screening tests: materials, administration and analysis		
Grade of nurse assumed to carry out test administration	**D**	ARHT
Nurse hourly rate (average of scale including NI and superannuation)	**£10.06**	Salary scales
50-g GCT		
Glucose load	**£0.93**	*MIMS*[194]
Biochemical analysis	**£5.43**	ARHT (inflated)
Nurse time (minutes)	**10**	ARHT
Nursing cost of administration	**£1.68**	Calculated
Total cost of 50-g GCT	**£8.04**	Calculated
RPG/FPG		
Biochemical analysis	**£3.26**	ARHT (inflated)
Nurse time (minutes)	**5**	ARHT
Nursing cost of administration	**£0.84**	Calculated
Total cost of FPG	**£4.10**	Calculated
Cost of diagnostic tests: material, administration and analysis		
Grade of nurse assumed to carry out test administration	**D**	ARHT
Nurse hourly rate (average of scale including NI and superannuation)	**£10.06**	Salary scales
75-g GTT		
Glucose load	**£1.40**	*MIMS*[194]
Biochemical analysis	**£10.86**	ARHT (inflated)
Nurse time (minutes)	**45**	ARHT
Nursing cost of administration	**£7.54**	Calculated
Total cost of 75-g GTT	**£19.80**	Calculated

respectively. Salary costs have been taken from current pay scales with the times for analysis, administration and dietary advice calculated by SHPIC having been retained. Averages of the scales were taken and National Insurance and superannuation were added. All calculations are based on a working week of 37.5 hours.

Other intangible patient costs, including the value of information,[195] anxiety[196] and discomfort or sickness on glucose load administration, have been excluded from the model.

Selection of efficacy data for use in the model

In selecting sensitivity and specificity data from the literature to use in the model, consideration was only given to those studies in which all patients were given a diagnostic GTT and not merely the positive screenees. The data from these studies are presented in appendix 4.

The majority of studies used a case series design. Typically there were no comparison or control groups used; the studies were a descriptive measure of the intervention and the outcome of screening for GDM. A small number of studies described their study design as randomised trials, however, on examination these are more quasi-experimental. The randomisation procedure was not described in any of these studies, nor was there evidence of any concealment of allocation, or blinding (e.g. Murphy and co-workers[188]). Because of the variance in study design and other methodological weaknesses (discussed below) in the study we were unable to synthesise the data in a meta-analysis. A number of the studies used fructosamine. It was subsequently discovered, on examination of these studies, that their focus was on patients explicitly referred for a GTT, thereby being a highly select and high incidence group. Furthermore, criteria recommending these women for the diagnostic test was not mentioned.[197]

GCT

The GCT was the screening test of choice in ten of these studies.[104] The study by Uncu and co-workers[182] was eliminated because of preselection. In another study no cases of GDM were identified.[168] Court and colleagues[123] used an unusual 100-g GCT.

Despite a range of cut-off values for the GCT being investigated across the studies, all but one study (Mathai and colleagues[184]) in a high incidence population used screening cut-off thresholds in the range 7.7–8.2 mmol/l. The range of sensitivity was demonstrated to be quite wide (between 33% and 96%), specificity 52–91% and PPV 9–49%. High incidences of GDM were witnessed in several studies,[104,153,158,164] and particularly so in Schwartz and co-workers,[97] although all of these could be expected because of the use of high incidence ethnic populations, with the exception of the Perucchini study.[104] This does raise the suspicion, especially given the high proportion of European women, of whether that study used preselection testing. Two of these studies included at least some women with risk factors (including age) in their study sample.[100] The majority of the subset did not pay any regard to the women's fasting status, while two studies fasted their participants. Several studies presented incomplete data in terms of sensitivity, specificity and PPV (necessary to construct our model).[97,104,123,158,184] The authors' calculations of these data were able to be confirmed in some cases.[100,153,164,188] The methodology of these four studies nevertheless contains flaws: the O'Sullivan study[100] is now dated and provides no information on fasting status; Jirapinyo and co-workers[164] has a wide spectrum for the women's screening interval (8–36 weeks); the Murphy study[188] used a small sample size; and Cetin and Cetin[153] repeated the GTT in patients with only one abnormal value but did not differentiate these patients from the patients who had two or more.

In selecting efficacy data for the cost-effectiveness model, of these four studies (*Table 10*), data reported by Cetin and Cetin[153] and O'Sullivan and co-workers[100] were used. The Murphy study[188] was considered unsuitable because the sample size was only 41 patients, likewise Jirapinyo and co-workers'[164] spectrum of screening times (appendix 4) ruled it unrealistic.

RPG/FPG

RPG tests comprise three studies in appendix 4.[140] The reliability of the data in Nasrat and co-workers[61] must be questioned. The figure of 47% for PPV frankly does not make sense and cannot be checked from the data presented in the paper. Mathai and co-workers[184] did not present complete data nor could their data be confirmed. In an interesting study design, the Jowett[140] study admitted participants for 24 hours and obtained samples every 3 hours, demonstrating a range of sensitivities for two cut-off points. However, these women were a preselected high-risk group, and the authors used an incorrect sensitivity and specificity calculation. Amongst these a range of cut-off values were utilised, which may account to some extent for the differences shown in sensitivity and specificity (16–66% and 66–96%, respectively).

TABLE 10 *Extracted data from 'robust' GCT studies*

50-g GCT with no regard to time of last meal in all women								
Threshold		**Sensitivity**	**Specificity**	**PPV**	***N***	**Diagnostic test**	**Incidence**	**Study**	**Data confirmed**
7.5 mmol/l		33%	74%	9%	41	100-g GTT (NDDG)	7.3%	Murphy, 1994[188]	Yes
7.8 mmol/l		65%	88%	27%	274	100-g GTT (NDDG)	6.2%	Cetin, 1997[153]	No
8.2 mmol/l		79%	87%	15%	752	100-g GTT (NDDG)	2.5%	O'Sullivan, 1973[100]	Yes
50-g GCT in women with risk factors, non-fasting								
Factor	**Threshold**	**Sensitivity**	**Specificity**	**PPV**	***N***	**Diagnostic test**	**Incidence**	**Study**	**Data confirmed**
Clinical risk factors	7.8 mmol/l	86%	65%	23%	396	100-g GTT (NDDG)	10.6%	Jirapinyo, 1993[164]	Yes
Age ≥ 25	8.2 mmol/l	88%	82%	19%	361	100-g GTT (NDDG)	4.6%	O'Sullivan, 1973[100]	Yes
Clinical risk factors and age ≥ 25	8.2 mmol/l	63%	80%	19%	235	100-g GTT (NDDG)	6.8%	O'Sullivan, 1973[100]	Yes
Clinical risk factors	8.3 mmol/l	83%	79%	32%	396	100-g GTT (NDDG)	10.6%	Jirapinyo, 1993[164]	Yes

The FPG was used in four studies.[185] A further study by Agarwal and co-workers[142] preselected women for a diagnostic test and was thus eliminated. The deficiencies with the Perucchini study[104] data are noted above. Both Reichelt and co-workers[143] and Fuhrmann[185] used large sample sizes and although the Fuhrmann data have been confirmed from the study, it uses a non-standardised (50 g) diagnostic test, which may not have picked up all cases of GDM. The Reichelt[143] data on the other hand could not be confirmed from the paper and the incidence of GDM, particularly in a sample with a large non-white population, seems very low. Once again, a range of cut-off values was used and consequently a range of sensitivity and specificities was demonstrated. Three of these studies included women with some risk factors.

Three studies presented side-by-side comparisons of RPG/FPG and GCT but results are inconclusive. The Mathai study[184] found very similar sensitivities for both GCT and RPG using ROC analysis, in an Indian population. The other two studies compared FPG. Fuhrmann[185] reported high sensitivity with the FBG but generally lower sensitivities in a number of comparisons with both 1-hour and 2-hour GCTs. Similarly, the Perucchini study[104] reported higher sensitivity with the FPG but lower specificity than the 1-hour GCT. A 2-hour GCT, on the other hand, had higher sensitivity and lower specificity than the FPG. One theme that is enduring through the majority of studies, regardless of screening test used, is that the testing period was between 24 and 28 weeks (second trimester). Thus these studies are assuming that all cases of GDM will be identified by the time of this cut-off.

Therefore, for the RPG and FPG, only the Reichelt study[143] was considered suitable (*Table 11*). Fuhrmann[185] was disqualified by its chosen (unproven) diagnostic test. The reasons for the selection of this option for the FPG were firstly that this was the authors' chosen threshold from those examined and secondly, despite the inability to confirm the data from the information presented in the study, the authors did use a standard GTT test, which Fuhrmann[185] failed to do. (The use of a 50-g GTT has raised doubts in the present authors' minds regarding the 100% sensitivity of the Fuhrmann 5.0 mmol/l threshold reported.)

Within the range of thresholds considered by the above studies, the following strategies and data have been used in the model: (1) Reichelt and co-workers[143] used ROC analysis to determine the optimum threshold at 4.9 mmol/l from their range of thresholds for the FPG; (2) Cetin and Cetin[153] presented only one strategy for the GCT using a threshold of 7.8 mmol/l; (3) O'Sullivan and co-workers[100] used an 8.2 mmol/l threshold; (4) O'Sullivan and co-workers[100] used a threshold of 8.2 mmol/l for women aged 25 years and over; (5) O'Sullivan and co-workers[100] again used an 8.2 mmol/l threshold for women aged 25 and over with risk factors. O'Sullivan also made errors in presenting data for this strategy (specificity of 87% rather than the correct 80%) but we were able to check this from the data presented in the paper. Strategy (6), giving all women a GTT (assuming 100% sensitivity), has also been added for reference. In practice the efficacy of the GTT is likely to vary as the glucose loads and thresholds are varied.

TABLE 11 *Extracted data from 'robust' RPG and FPG studies*

Random plasma glucose in women eaten within 2 hours								
Threshold	**Sensitivity**	**Specificity**	**PPV**	***N***	**Diagnostic test**	**Incidence**	**Study**	**Data confirmed**
7.0 mmol/l	16%	96%	47%	276	75-g GTT (WHO)	1.2%	Nasrat *et al.*, 1988[61]	No
Fasting plasma glucose								
Threshold	**Sensitivity**	**Specificity**	**PPV**	***N***	**Diagnostic test**	**Incidence**	**Study**	**Data confirmed**
4.5 mmol/l	94%	51%	0.6%	5010	75-g GTT (WHO)	0.3%	Reichelt *et al.*, 1988[143]	No
4.6 mmol/l	94%	58%	0.7%	5010	75-g GTT (WHO)	0.3%	Reichelt *et al.*, 1988[143]	No
4.7 mmol/l	94%	66%	0.9%	5010	75-g GTT (WHO)	0.3%	Reichelt *et al.*, 1988[143]	No
4.8 mmol/l	88%	72%	1.0%	5010	75-g GTT (WHO)	0.3%	Reichelt *et al.*, 1988[143]	No
4.9 mmol/l	88%	78%	1.3%	5010	75-g GTT (WHO)	0.3%	Reichelt *et al.*, 1988[143]	No

Results

Table 12 presents these strategies in terms of number of screens required, true-positives, false-positives, number of GTTs required, total cost, average cost, added benefit (in terms of cases of GDM detected), added cost and marginal cost. A marginal analysis approach (in order of cost) has determined the order in which the strategies appear. The data below the table show the underlying assumptions.

Most recent data have put the number of annual deliveries in England and Wales at 621,900.[198] Prevalences of 1, 2 and 3% would yield approximately 6219, 12,438 and 18,657 annual cases of GDM, respectively. Initially we assumed an incidence of GDM of 2%. As can be seen, strategies (2), (3) and (4) have been dominated by (1), the FPG strategy. The reasons for selecting this option for the FPG were firstly that this was the authors' chosen threshold from those examined and secondly, that despite the inability to confirm the data from the information presented in the study, the authors did use a standard GTT test, which Fuhrmann[185] failed to do.

The marginal costs per case of GDM detected of strategies (5) and (1) are £488 and £489; strategy (1) detects 6009 more of the women with GDM but at an additional cost of £2,940,824.

It should be noted that these results are based on the data from single studies.

Conclusion and discussion

If screening is to be deployed on the assumptions in *Table 12,* the choice would appear to be between a universal FPG and giving the GCT to those aged 25 and over with risk factors. That said, use of the FPG on selected populations has not been considered in the literature and thus not here in the model either, but should be the subject of further research. The FPG strategy detects 6000 more of the women with GDM but at an additional cost of £3 million. If these additional women could be shown to be at lower risk of adverse events then strategists could justify choosing a GCT on the restricted population. The FPG strategy also leads to an additional 90,000 false-positives. Kerbel and co-workers[163] observed that "pregnant women with false-positive GDM screening results experience a significant decline in their perception of their own health". No cost of this has been represented in the model and the inconvenience/anxiety caused to these patients should be considered in further research and incorporated into decision-making. In addition, the O'Sullivan[100] data for screening women aged ≥ 25 years is based on research published in 1973, and an increase in the proportion of women having later pregnancies in the present times would probably increase screening costs and thereby lower its cost-effectiveness. Based on these data, a comprehensive programme of GDM detection could cost around £5–6 million per year. Marginal cost per case of GDM detected falls to £333 when incidence is 3% but rises to £959 when incidence is 1%.

Current practice in the UK, from the Mires survey, shows much variation in screening techniques and the associated consumption of resources.[1] This might be justified in terms of marginal cost if the different screening strategies were based on differences in prevalence rates of local ethnic populations. UK-reported prevalences are relatively consistent, ranging from 0 to 2%, even in studies with large ethnic groups. Dornhurst and co-workers[115] found an incidence of 1.5% in a multiracial group (Indian subcontinent 10.5%, Afro-Caribbean 17.7%, South East Asian 5%, Middle Eastern and other non-white 11.4%) as did Maresh and co-workers[199] with a 75% non-white population. A number of UK-based studies did not define the ethnic mix of the study sample, and there were variances in inclusion and exclusion criteria, making comparisons difficult. Similarly, the use of different diagnostic tests and thresholds may reflect the prevalences reported.

However, any alternative to universal screening may be problematic, as Cousins and colleagues[200] observed:

> "Adding complexity to test administration and interpreting lead to patient and physician error and non-compliance ... Each decision point increases the possibility of error and takes the physician's time to consider the risks and to explain to each patient why she does or does not need screening. It would be more efficient and less error prone to use universal screening, in which a clerk would schedule the test automatically at 24 to 28 weeks gestation."

The primary motive for determining the presence of GDM in women is the belief that these women have a greater risk of adverse pregnancy outcomes such as hypertension, and Caesarean sections because of macrosomic infants. However, it needs to be established if the higher rate of these complications in women with GDM is a result of the diagnosis alone or other risk factors such as age and obesity. The raised glucose level may be a

TABLE 12 *Analysis of selected screening strategies – marginal and average cost per case of GDM detected*

Strategy	No. of screens	True-positives	False-negatives	Missed	Cost of GCT/ FBGs	No. of GTTs	Cost of GTTs	Total cost	Average cost per case of GDM detected	Added benefit (extra no. of cases detected)	Added cost	Marginal cost per case of GDM detected
(5) GCT ≥ 8.2 mmol/l to age ≥ 25 and risk factors	192,789	4,937	38,558	4,602	£1,549,233	43,494	£861,257	£2,410,491	£488	N/A	N/A	£488
(1) FPG ≥ 4.9 mmol/l to all	621,900	10,945	130,599	1,555	£2,548,515	141,544	£2,802,800	£5,351,315	£489	6,009	£2,940,824	£489
(4) GCT ≥ 8.2 mmol/l to age ≥ 25	460,206	9,632	82,837	1,493	£3,698,170	92,469	£1,831,031	£5,529,201	£574	Dominates		
(2) GCT ≥ 7.8 mmol/l to all	621,900	5,255	74,628	4,353	£4,997,527	79,883	£1,581,809	£6,579,335	£1,252	Dominates		
(3) GCT ≥ 8.2 mmol/l to all	621,900	7,763	80,847	2,612	£4,997,527	88,610	£1,754,607	£6,752,133	£870	Dominates		
(6) GTT to all	621,900	12,438	0	0	£0	621,900	£12,314,586	£12,314,586	£990	1,493	£6,963,271	£4,665

Assumptions:		Source
Annual number of deliveries in England	621,900	*Health Statistics Quarterly*, Winter 2000[198]
Incidence of GDM	0.02	Estimated
Annual number of cases of GDM	12,438	Calculated
Sensitivity of GTT	1.00	Assumption
Sensitivity of GCT at ≥ 7.8 mmol/l	0.65	Cetin and Cetin, 1997[153]
Sensitivity of GCT at ≥ 8.2 mmol/l	0.79	O'Sullivan *et al.*, 1973[100]
Sensitivity of GCT at ≥ 8.2 mmol/l to women ≥ 25 years	0.88	O'Sullivan *et al.*, 1973[100]
Sensitivity of GCT at ≥ 8.2 mmol/l to women ≥ 25 years and risk factors	0.63	O'Sullivan *et al.*, 1973[100]
Sensitivity of FPG at 4.9 mmol/l	0.88	Reichelt *et al.*, 1998[143]
False-positive rate of GCT at ≥ 7.8 mmol/l	0.12	Cetin and Cetin, 1997[153]
False-positive rate of GCT at ≥ 8.2 mmol/l	0.13	O'Sullivan *et al.*, 1973[100]
False-positive rate of GCT at ≥ 8.2 mmol/l to women ≥ 25 years	0.18	O'Sullivan *et al.*, 1973[100]
False-positive rate of GCT at ≥ 8.2 mmol/l to women ≥ 25 years and risk factors	0.20	O'Sullivan *et al.*, 1973[100]
False-positive rate of FPG at 4.9 mmol/l	0.21	Reichelt *et al.*, 1998[143]
Proportion of GDM women with GCT < 7.8 mmol/l	0.35	Cetin and Cetin, 1997[153]
Proportion of GDM women with GCT < 8.2 mmol/l	0.21	O'Sullivan *et al.*, 1973[100]
Proportion of GDM women aged ≥ 25 with GCT < 8.2 mmol/l	0.12	O'Sullivan *et al.*, 1973[100]
Proportion of GDM women aged ≥ 25 and risk factors with GCT < 8.2 mmol/l	0.37	O'Sullivan *et al.*, 1973[100]
Proportion of GDM women with an FPG < 4.9 mmol/l	0.125	Reichelt *et al.*, 1998[143]
Proportion of deliveries to women aged ≥ 25 years	0.74	*Health Statistics Quarterly*, Winter 2000[198]
Proportion of deliveries to women aged ≥ 25 years with risk factors	0.31	O'Sullivan *et al.*, 1973[100]
Health service cost of a 50-g GCT	£8.04	ARHT (inflated)
Health service cost of a 75-g GTT	£19.80	ARHT (inflated)
Health service cost of an RBG	£4.10	ARHT (inflated)

correlate rather than a cause. In his Cochrane review, Walkinshaw[65] appraised studies that investigated the effect of dietary therapy in women with GDM on neonatal outcomes and fetal growth. No differences in birth weight or Caesarean deliveries were observed between those women who were given dietary treatment and those who were not.

Long-term consequences

Women contracting GDM are at increased risk of developing diabetes later in life.[5] It was estimated in 1992 by Howard[201] that as many as one-third of women with GDM will develop diabetes in later life. A 1993 study by Gregory and co-workers[84] estimated that women with GDM have a 50% risk of developing non-insulin-dependent diabetes within 10 years. A study in 1992 by Damm[202] found rates of 3–65% of diabetes after GDM had been reported. Children whose mothers have had GDM are also at increased risk.[203] However, this raises two questions – should we screen for type 2 diabetes, and if so would screening for GDM (which may or may not be followed by type 2 diabetes years later) be an efficient method?

Diabetes is a major burden on health service resources and can lead to many complications, such as retinopathy, renal failure and nephropathy, as well as being associated with increased risk of disease such as cardiovascular disease. Managing complications of diabetes has been reported to account for 5% of healthcare expenditure in the UK.[204] A more recent study from Cardiff estimated that it might be higher.[205] A US economic model attempted to quantify costs saved if the progression to diabetes in GDM women is delayed. Based on the (far from safe) assumption that education and counselling can delay or prevent disease progression, the model demonstrated that in a population of 125,370, if the rate of delay is just 10%, this would save in the region of US$70,000 a year, with a saving over 10 years of US$32,000. A 50% reduction could save approximately US$330,000.[84]

So what is the incidence of diabetes post-GDM and can this be predicted? A number of studies have investigated the incidence of diabetes following a diagnosis of GDM and factors that may be predictive of such a diagnosis. Damm and co-workers[202] studied women with previous dietary-controlled GDM over an 11-year period. They reported that 17.4% developed diabetes, with the median interval for diagnosis being 48 months. Significant associations were demonstrated between diabetes development and parity > 2, early diagnosis of GDM, fasting glucose > 5 mmol/l at diagnosis or 2 months post-partum, non-white race, preterm delivery and macrosomic infants. No association was found with maternal weight, age, family history or pregnancy complications. However, when these factors were assessed in a logistic regression analysis, only three variables remained significant predictors: a high fasting glucose level at diagnosis of GDM, preterm delivery and GDM coming on in the first trimester.

An Australian study by Grant and colleagues,[206] using a 75-g GTT with WHO criteria, found an incidence of diabetes of 11%. Again the authors found that raised glucose levels in the early post-partum period were a predictor of later diabetes. Family history was not a predictive factor, but 73.5% of women with diabetes were overweight compared to 44% without diabetes. Kjos and co-workers[207] found an incidence of diabetes of 55% over a 6-year period in high-risk Latino women. These women were not found to be diabetic at post-partum testing. The only significant predictive values were the area under the glucose curve of the post-partum GTT, gestational age at time of diagnosis of GDM, the area under the ante-partum glucose tolerance curve, and a high fasting glucose concentration during pregnancy. Parity, previous GDM, previous infant with anomalies, family history, BMI or hypertension were not predictors of later diabetes.

As can be seen, reported incidence rates of diabetes following GDM vary quite widely. O'Sullivan[208] points out that a number of factors may need to be taken into account when considering these rates, for example the use of different diagnostic criteria (WHO versus NDDG), no account of those not attending, different observational periods, and the use of self-reported measures. However, despite these differences, all of the aforementioned studies have demonstrated that a significant proportion of women with GDM go on to develop diabetes, although the predictors for this are less clear. The early diagnosis of diabetes may, as predicted by Gregory and co-workers,[84] have cost saving consequences.

If it could be assumed that dietary manipulation reduced macrosomia and therefore the need for Caesarean section, there would be health service savings from a screening programme. Naylor and co-workers[22] assumed the Caesarean rate is 20.2% in normal pregnancy and 33% in pregnancies complicated by GDM. In addition to a more complicated delivery procedure, Caesarean sections require a longer length of stay, 6 days on average, compared with 3 days for a normal birth. The cost for a Caesarean birth was calculated at £2010, compared with £909 for a normal birth

(SHPIC costing unit, obtained from DTHT finance department costing spreadsheet, 1997/98 prices). Therefore, for every Caesarean birth prevented, the cost of screening and treatment would have been offset by £1101, although this assumes that the savings could be realised. Because the main cost of hospital activities is staff, savings are only realised if staff numbers are reduced, but there would be time freed up for other duties.

However, we are a long way from being able to be certain that screening for GDM and putting into place interventions for the mother after delivery would reduce any of these risks.

Chapter 5
Discussion

Weakness of research evidence

Coustan[209] remarked in 1996 that "GDM became a clinical entity without the carefully collected epidemiological data we now require". In a more recent editorial, Dornhorst and Chan[16] conclude that:

> "In today's world of evidence-based medicine, audit and cost constraints, the case for screening and treating women with GDM is severely hampered by the lack of a clear definition, agreed diagnostic criteria, and evidence that improving mild disturbances of maternal glycaemia improves pregnancy outcome."

Best test

Screening by FPG has attractions, and a case could be made for using it as the definitive test.[76] However, there are two concerns with this approach. Firstly, there will be a group of women with normal FPG but abnormal postprandial levels. Secondly, there is some evidence that fetal weight may be more closely linked to postprandial maternal glucose levels than to fasting ones.[210] This may be because maternal glucose intolerance leads to high levels after meals, with peaks of diffusion of glucose across the placenta, leading to stimulation of fetal insulin production, and in turn to overgrowth of insulin-sensitive tissues.[66] Research is needed to identify and study women with normal fasting but high postprandial levels, before we can rely on screening with FPG alone. Detection of this group is not worthwhile unless treatment can affect outcomes, and one study found that postprandial monitoring of glucose levels was no better than preprandial alone.[211] This may not apply to all women with GDM, because de Veciana and co-workers[212] found that in women who required insulin, postprandial monitoring improved HbA_{1c} and outcomes such as birth weight over 4000 g, compared to preprandial monitoring. Both groups monitored FPG. This study appears to support the views of Fraser[66] above – that in at least some women, postprandial peaks in glucose levels are sufficient to cause overgrowth of some fetal tissues.

Economic evidence

As shown in chapter 4, there is a reasonable body of evidence on efficiency of screening, in terms of cost per case of GDM found. For cost-effectiveness we need data on the extent to which the entire screening process (i.e. screening + diagnosis + treatment) reduces adverse outcomes (Caesarean sections and other birth trauma; morbidity such as hypoglycaemia amongst neonates; later diabetes amongst mothers), and the net cost per adverse event prevented. The diversity of events makes it difficult to produce a single incremental cost-effectiveness ratio, unless all could be converted to a single common currency such as the QALY. It might be possible to overcome this if one adverse event dominated the rest. A cost per event prevented could then be compared with the cost-effectiveness of other uses of healthcare funds, and if it appeared cost-effective, screening could be justified and the other benefits regarded as incidental bonuses. If preventing the primary adverse outcome was borderline in cost-effectiveness terms, then simple enumeration of other outcomes avoided might be sufficient.

The commonest adverse outcome of GDM is Caesarean section, but unfortunately there is conflicting advice as to whether this is a necessary consequence or not. Macrosomia is usually regarded as an adverse outcome, but is an arbitrary classification based on cut-offs of 4000 g or 4500 g, and this does not distinguish between large healthy babies and large less healthy ones. Furthermore, the majority of 'macrosomic' babies are born to non-diabetic mothers. We need a better way of identifying the unhealthy babies.

Admission to SCBU could be used as a proxy for neonatal ill health, but as with Caesarean section rates, the diagnosis itself may trigger special care.

Perinatal mortality is too rare an event to be useful.

Interim conclusions on best policy

We need better evidence, and any conclusions need to be interim ones pending further research. However, we need some interim national policies in the meantime, and the following are suggested for debate:

- There is insufficient evidence to justify universal screening for GDM. The condition is poorly

defined, being based too often on arbitrary cut-offs rather than levels of morbidity. There appears to be a continuum of risk, rather than two distinct groups of normal and GDM women.
- Screening for GDM does not fully meet the criteria of the NSC.
- Decisions on screening need to take into account the inconvenience and harms to the very large numbers of false-positives with any of the tests, and the harms such as the high Caesarean section rate that follows the diagnosis alone.
- A high diagnostic threshold should be used, because lower levels of glucose intolerance give low risk of adverse outcomes of pregnancy, and a low diagnostic threshold greatly increases the proportion of mothers labelled as 'abnormal', with all the inconvenience and anxiety that may cause.
- Selective screening should be based mainly on overweight, because adverse outcomes are more closely related to weight than glucose level.
- FPG is useful, but will miss the group that has normal fasting but elevated postprandial glucose levels. However, using an FPG cut-off of < 4.8 mmol/l would exclude GDM.
- A compromise interim solution might be two-stage selective screening based firstly on risk factors, principally overweight and age, but also ethnicity, and secondly on GCT, using 8.2 mmol/l as the GCT threshold. Women under 25 could be tested only if they were both overweight and had a family history, and women over 25 years only if they had at least one risk factor other than age. If GCT was positive, dietary advice would be given, but no OGTT should be done. If glucose levels remained high, further testing could be done and more intensive treatment considered. One problem with this solution is that many clinics are already short of dietetic input.

Research needs – main questions to be answered

The main research needs are as follows:

- There is still a need to further define the 'disease', through documenting the frequency of adverse events. This should be done by observational, population-based epidemiological surveys, similar to the Swedish one by Agardh and colleagues,[47] but in the UK perhaps paying particular attention to adequate numbers from ethnic groups, in whom risks may be different.
- The survey should relate outcomes of pregnancy to maternal blood glucose levels, to see at what level of blood glucose the outcomes worsened significantly. This would then form the basis for diagnosis. Data on confounding variables such as obesity would be collected, and all data analysed to assess the extent to which glucose levels are a significant additional predictor.
- It is anticipated that such surveys would confirm a continuum of risk and glucose levels. It will then be necessary to decide at what level(s) intervention is justified. Economic appraisal should help here, through examination of cost per adverse event avoided, and relation of those to the usual costs of NHS activities. Neonatal health and Caesarean section rates could be the main outcomes, with maternal anxiety and other disbenefits being secondary ones. The cost basis needs to be improved.
- Trials of different screening tests – FPG versus 50-g GCT as screening test; research into whether a follow-up OGTT is worthwhile, because it is far from being a perfect gold standard. Particular attention should be given to women with normal FPGs but abnormal GCTs, and their absolute risk of adverse infant outcomes. A multicentre trial is essential given the large numbers of pregnant women, to give reasonable numbers of those with GDM in each arm. As stated earlier, 15,000 women divided into three arms (GCT, FPG, no screening) could, at a 2% incidence rate, be expected to deliver 100 women with GDM in each of the three arms.
- A randomised trial of treatment in the subgroup with normal FPG but abnormal GCT, stratified by glucose level, is indicated.
- Is screening cost-effective?

Known research studies currently ongoing

A 5-year international study, the HAPO study, funded by the National Institutes of Health, USA, is under way. Approximately 25,000 women, from the USA, Canada, Europe, Asia and Australia will be recruited by 2003. Belfast is one of the centres. Its primary hypothesis is that hyperglycaemia, less severe than overt diabetes, is associated with an increased risk of adverse maternal, fetal and neonatal outcomes that are related to the degree of metabolic disturbance.

A multicentre international trial (ACHOIS) is currently investigating the effect of screening and management of glucose intolerance of

pregnancy in approximately 1000 women. Consenting women with a raised GTT are either informed of their diagnosis and actively managed or are not informed of their diagnosis and receive routine antenatal care. Comparators include maternal and infant physical morbidity, costs of, and the psychological consequences of GDM. This study aims to finish within the next 2–3 years.

A study designed to investigate the long-term implications of a diagnosis of gestational IGT, including the level of risk of developing diabetes, is currently under way in Aberdeen. The long-term effects, including body weight, cardiovascular risk factors, blood pressure, gynaecological outcome and renal function of women recruited to this retrospective study are to be assessed 18 years post-pregnancy.

Acknowledgements

This study was commissioned by the NHS R&D HTA Programme.

We are grateful to the advisory panel, who provided expert advice and comments on the research protocol and/or a draft of this report, but we absolve them from any responsibility for the final product, responsibility for which rests with SHTAC:

Dr Anne Dornhorst, Hammersmith Hospital, London

Professor RJ Jarrett, London

Dr R Fraser, Royal Hallamshire Hospital, Sheffield

Sinead Dunne, Diabetes UK.

We would also like to thank Dr Pam Royle and Ms Liz Hodson at the Information Service, Southampton Health Technology Centre.

References

1. Mires GJ, Williams FLR, Harper V. Screening practices for gestational diabetes mellitus in UK obstetric units. *Diabet Med* 1999;**16**:138–41.
2. American Diabetes Association. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 1997;**20**:1183–7.
3. American Diabetes Association. Gestational diabetes mellitus. *Diabetes Care* 1998;**21** Suppl 1.
4. Coustan DR, Nelson C, Carpenter MW, Carr SR, Rotondo L, Widness JA. Maternal age and screening for gestational diabetes: a population-based study. *Obstet Gynecol* 1989;**73**:557–61.
5. US Preventive Services Task Force. Guide to Clinical Preventive Services: Report of the US Preventive Services Task Force, 2nd edn. Baltimore, USA: Williams and Wilkins; 1996. p. 193–208.
6. Canadian Task Force on the Periodic Health Examination. The periodic health examination, 1992 update: 1. Screening for gestational diabetes mellitus. *Can Med Assoc J* 1992;**147**:435–43.
7. Thompson DM, Wylie B, Bryer-Ash M, Creed M, Jenkins L, Kozak S, *et al.* Gestational diabetes: to screen or not to screen [letter]. *Can Med Assoc J* 1993;**148**:13.
8. Ryan EA. To screen or not to screen? [letter]. *Can Med Assoc J* 1993;**148**:11.
9. Jarrett RJ. Should we screen for gestational diabetes? The concept of gestational diabetes was popularised before considerations of evidence based medicine came on the scene. *BMJ* 1997;**315**:736–7.
10. Stephenson MJ. Screening for gestational diabetes mellitus: a critical review. *J Fam Pract* 1993;**37**:277–83.
11. Goer H. Gestational diabetes: the emperor has no clothes. *Birth Gazette* 1996;**12**:32–5.
12. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;**44**:1271–8.
13. Dornhorst A, Beard RW. Gestational diabetes: a challenge for the future. *Diabet Med* 1993;**10**:897–905.
14. Dornhorst A, Chan SP. The elusive diagnosis of gestational diabetes. *Diabet Med* 1998;**15**:7–10.
15. Griffiths M. Debate over screening for gestational diabetes. *BMJ* 1998;**316**:861.
16. Ray J. Debate over screening for gestational diabetes – evidence from randomised controlled trial is needed. *BMJ* 1998;**316**:861.
17. American Diabetes Association. Position statement on gestational diabetes mellitus. *Diabetes Care* 1989;**9**:430–1.
18. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979;**28**:1039–57.
19. O'Sullivan JB, Mahan CM. Criteria for the oral glucose tolerance test in pregnancy. *Diabetes* 1964;**13**:278–85.
20. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. *Am J Obstet Gynecol* 1982;**144**:768–73.
21. Metzger BE, Coustan DR. Summary and recommendations of the Fourth International Workshop-Conference on Gestational Diabetes Mellitus. *Diabetes Care* 1998;**21** Suppl 2:16–17.
22. Naylor CD, Sermer M, Chen E, Sykora K. Cesarean delivery in relation to birth weight and gestational glucose tolerance: pathophysiology or practice style? Toronto Trihospital Gestational Diabetes Investigators. *JAMA* 1996;**275**:1165–70.
23. Nord E, Hanson U, Persson B. Blood glucose limits in the diagnosis of impaired glucose tolerance during pregnancy. *Acta Obstet Gynecol Scand* 1995;**74**:589–93.
24. Alberti KGMM, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diabetic Med* 1998;**15**:539–53.
25. Schmidt MI, Matos MCG, Reichelt AJ, Costa Forti A, de Limas L, Duncan BB. Prevalence of gestational diabetes mellitus – do the new WHO criteria make a difference? Brazilian Gestational Diabetes Study Group. *Diabetic Med* 2000;**17**:376–80.
26. Weiss PAM, Haeusler MCH, Kainer F, Purstner P, Haas J. Towards universal criteria for gestational diabetes: relationships between 75 and 100 g glucose loads and between capillary and venous glucose concentrations. *Am J Obstet Gynecol* 1998;**178**:830–5.
27. Martin FIR. The diagnosis of gestational diabetes. *Med J Aust* 1991;**155**:112.
28. Oats JN, Beischer NA. Gestational diabetes. *Aust N Z J Obstet Gynaecol* 1986;**26**:2–10.

29. Dornhorst A, Hadden D. Diabetes and pregnancy: an international approach to diagnosis and management. Chichester: John Wiley & Sons; 1996.

30. Kjos SL, Buchanan TA. Gestational diabetes mellitus. *N Engl J Med* 1999;**341**:1749–56.

31. Sermer M, Naylor CD, Gare DJ, Kenshole AB, Ritchie JW, Farine D, *et al.* Impact of time since last meal on the gestational glucose challenge test. The Toronto Tri-Hospital Gestational Diabetes Project. *Am J Obstet Gynecol* 1994;**171**:607–16.

32. Essel JK, Opai-Tetteh ET. Macrosomia – maternal and fetal risk factors. *S Afr Med J* 1995;**85**:43–6.

33. Spellacy WN, Miller S, Winegar A, Peterson PQ. Macrosomia – maternal characteristics and infant complications. *Obstet Gynecol* 1985;**66**:158–61.

34. Weiss PAM, Haeusler MCH, Tamussino K, Haas J. Can glucose tolerance test predict fetal hyperinsulinism? *Br J Obstet Gynaecol* 2000;**107**:1480–5.

35. Stanley KP, Fraser RB. Cord insulin and C-peptide – distribution in an unselected population. *Br J Obstet Gynaecol* 1992;**99**:512–15.

36. Ales KL, Santini DL. Should all pregnant women be screened for gestational glucose intolerance? *Lancet* 1989;**i**:1187–91.

37. Modanlou HD, Komatsu G, Dorchester W, Freeman RK, Bosv SK. Large for gestational age neonates: anthropometric reasons for shoulder dystocia. *Obstet Gynecol* 1982;**60**:417–23.

38. Sacks DA. Fetal macrosomia and gestational diabetes: what's the problem? *Obstet Gynecol* 1993;**81**:775–81.

39. Sharp PS, O'Connor R, Young S, Wooton J, Frost P. Gestational diabetes mellitus. *BMJ* 1993;**306**:581.

40. Hammami M, Walters JC, Hockman EM, Koo WW. Disproportionate alterations in body composition of large for gestational age neonates. *J Pediatr* 2001;**138**:817–21.

41. Schwartz R, Teramo KA. What is the significance of macrosomia? *Diabetes Care* 1999;**22**:1201–5.

42. Potter EL, Craig JM, editors. Pathology of the fetus and the infant. Chicago: Year Book, Medical Publishers Inc.; 1975.

43. Fraser RB, Bruce C. Amniotic fluid insulin levels identify the fetus at risk of neonatal hypoglycaemia. *Diabet Med* 1999;**16**:568–72.

44. Jensen DM, Sorensen B, Feilberg-Jorgensen N, Westergaard JG, Beck-Nielsen H. Maternal and perinatal outcomes in 143 Danish women with gestational diabetes mellitus and 143 controls with a similar risk profile. *Diabet Med* 2000;**17**:281–6.

45. Nasrat HA, Augensen K, Abushal M, Shalhoub JT. The outcome of pregnancy following untreated impaired glucose tolerance. *Int J Gynaecol Obstet* 1994;**47**:1–6.

46. Nordlander E, Hanson U, Persson B. Factors influencing neonatal morbidity in gestational diabetic pregnancy. *Br J Obstet Gynaecol* 1989;**96**:671–8.

47. Agardh CD, Aberg A, Norden NE. Glucose levels and insulin secretion during a 75 g glucose challenge test in normal pregnancy. *J Intern Med* 1996;**240**:303–9.

48. Hunter DJS, Keirse MJNC. Gestational diabetes. In: Chalmers I, Enkin M, Keirse MJNC, editors. Effective care in pregnancy and childbirth. Oxford: Oxford University Press; 1989. p. 403–10.

49. Langer O, Anyaegbunam A, Brustman L, Dixon M. Management of women with one abnormal oral glucose tolerance test value reduces adverse outcome in pregnancy. *Am J Obstet Gynecol* 1989;**161**:593–9.

50. Naylor CD. Diagnosing gestational diabetes mellitus: is the gold standard valid? *Diabetes Care* 1989;**12**:565–72.

51. Green JR, Schumacher LB, Pawson IG, Partridge JC, Kretchmer N. Influence of maternal body habitus and glucose tolerance on birth weight. *Obstet Gynecol* 1991;**78**:225–40.

52. Wilson J, Junger G. Principles and practice of screening for disease. WHO Public Health Paper 34. Geneva: WHO; 1968.

53. The UK National Screening Committee's criteria for appraising the viability, effectiveness and appropriateness of a screening programme. 2001. URL: http://www.nsc.nhs.uk

54. NHS Centre for Reviews and Dissemination. The prevention and treatment of obesity. *Effective Health Care* 1997;3.

55. Edwards LE, Hellerstedt WL, Alton IR, Story M, Himes JH. Pregnancy complications and birth outcomes in obese and normal-weight women: effects of gestational weight change. *Obstet Gynecol* 1996;**87**:389–94.

56. Newman RB, Calhoun BC, Maganna E. Seasonal effects on the rate of fetal macrosomia. *Am J Obstet Gynecol* 1999;**180**:S35.

57. Lao TT, Lee CP. Gestational 'impaired glucose tolerance': should the cut-off be raised to 9 mmol l^{-1}? *Diabet Med* 1998;**15**:25–9.

58. Pettitt DJ, Narayan KM, Hanson RL, Knowler WC. Incidence of diabetes mellitus in women following impaired glucose tolerance in pregnancy is lower than following impaired glucose tolerance in the non-pregnant state. *Diabetologia* 1996;**39**:1334.

59. Roberts RN, McManus J, Dobbs S, Hadden D. A standardised breakfast tolerance test in pregnancy: comparison with the 75 g oral glucose tolerance test in unselected mothers and in those with impaired glucose tolerance. *Ulster Med J* 1997;**66**:18–23.

60. Martin FIR, Ratnaike S, Wootton A, Condos P, Sutter PEN. The 75 g oral glucose tolerance test in pregnancy. *Diabetes Res Clin Pract* 1995;**27**:147–51.

61. Nasrat HA, Johnstone FD, Hasan SAM. Is random plasma glucose an efficient screening test for abnormal glucose tolerance in pregnancy? *Br J Obstet Gynaecol* 1988;**95**:855–60.

62. Hatem M, Anthony F, Hogston P, Rowe DJF, Dennis KJ. Reference values for 75 g glucose tolerance test in pregnancy. *BMJ* 1988;**296**:676–8.

63. Sutherland HW, Pearson DWM, Lean MEJ, Campbell DM. Breakfast tolerance test in pregnancy. In: Sutherland HW, Stowers JM, Pearson DWM, editors. Carbohydrate metabolism in pregnancy and the newborn IV. London: Springer-Verlag; 1989. p. 267–75.

64. Li DFH, Wong VCW, O'Hoy KM, Ma HK. Evaluation of the WHO criteria for the 75 g oral glucose tolerance test in pregnancy. *Br J Obstet Gynaecol* 1987;**94**:847–50.

65. Walkinshaw SA. Dietary regulation for 'gestational diabetes' (Cochrane Review). In: The Cochrane Library. Issue 3. Oxford: Update Software; 2000.

66. Fraser R. Diabetic control in pregnancy and intrauterine growth of the fetus. *Br J Obstet Gynaecol* 1995;**102**:275–7.

67. Persson B, Stangenberg M, Hansson U, Nordlander E. Gestational Diabetes Mellitus (GDM): comparative evaluation of two treatment regimens, diet versus insulin and diet. *Diabetes* 1985;**34**:101–4.

68. Simmons D, Robertson S. Influence of maternal insulin treatment on the infants of women with gestational diabetes. *Diabet Med* 1997;**14**:762–5.

69. Langer O, Rodriguez DA, Xenakis EM, McFarland KF, Berkus MD, Arredondo F. Intensified versus conventional management of gestational diabetes. *Am J Obstet Gynecol* 1994;**170**:1036–47.

70. Buchanan TA, Kjos SL, Montoro MN, Wu HP, Madrilejo NG, Gonzalez M, *et al.* Use of fetal ultrasound to select metabolic therapy for pregnancies complicated by mild gestational diabetes. *Diabetes Care* 1994;**17**:275–83.

71. Bochner CJ, Medearis AL, Williams J, Castro L, Hobel CJ, Wade ME. Early third-trimester ultrasound screening in gestational diabetes to determine the risk of macrosomia and labor dystocia at term. *Am J Obstet Gynecol* 1987;**157**:703–8.

72. Johnstone FD, Prescott RJ, Steel JM, Mao JH, Chambers S, Muir N. Clinical and ultrasound prediction of macrosomia in diabetic pregnancy. *Br J Obstet Gynaecol* 1996;**103**:747–54.

73. Fraser RB, Ford FA, Lawrence GF. Insulin sensitivity in third trimester pregnancy. A randomised study of dietary effects. *Br J Obstet Gynaecol* 1988;**95**:223–9.

74. Langer O, Conway DL, Berkus MD, Xenakis EM-J, Gonzales O. Glyburide was as safe and effective as insulin in gestational diabetes. *N Engl J Med* 2000;**343**:1134–8.

75. Goldberg JD, Franklin B, Lasser D, Jornsay DL, Hausknecht RU, Ginsberg-Fellner F, *et al.* Gestational diabetes: impact of home glucose monitoring on neonatal birth weight. *Am J Obstet Gynecol* 1986;**154**:546–50.

76. Rey E. Screening for gestational diabetes mellitus. A simple test may make it easier to study whether screening is worthwhile. *BMJ* 1999;**319**:798–9.

77. Bung P, Artal R, Khodiguian N, Kjos SL. Exercise in gestational diabetes. An optional therapeutic approach? *Diabetes* 1991;**40**:182–5.

78. Jovanovic-Peterson L, Peterson CM. Review of gestational diabetes mellitus and low-calorie diet and physical exercise as therapy. *Diabetes Metab Rev* 1996;**12**:287–308.

79. Avery MD, Leon AS, Kopher RA. Effects of a partially home-based exercise program for women with gestational diabetes. *Obstet Gynecol* 1997;**89**:10–15.

80. Lesser KB, Gruppuso PA, Terry RB, Carpenter M. Exercise fails to improve postprandial glycemic excursion in women with gestational diabetes. *J Matern Fet Med* 1996;**5**:211–17.

81. Dye TD, Knox KL, Artal R, Aubry RH, Wojtowycz MA. Physical activity, obesity, and diabetes during pregnancy. *Am J Epidemiol* 1997;**146**:961–5.

82. Stowers JM, Sutherland HW, Kerridge DF. Long-range implications for the mother: the Aberdeen experience. *Diabetes* 1985;**34**:106–10.

83. Wein P, Beischer NA, Sheedy MT. Studies of postnatal diabetes mellitus in women who had gestational diabetes. Part II. Prevalence and predictors of diabetes mellitus after delivery. *Aust N Z J Obstet Gynaecol* 1997;**37**:420–3.

84. Gregory K, Kjos SL, Peters RK. Cost of non insulin dependent diabetes in women with a history of gestational diabetes: implications for prevention. *Obstet Gynecol* 1993;**81**:782–6.

85. Langer O, Levy JC, Brustman L, Anyaegbunam A, Merkatz R, Divon M. Glycemic control in gestational diabetes mellitus – how tight is tight enough: small for gestational age versus large for gestational age? *Am J Obstet Gynecol* 1989;**161**:646–53.

86. Wen SW, Liu S, Kramer MS, Joseph KS, Levitt C, Marcoux S, *et al.* Impact of prenatal glucose screening on the diagnosis of gestational diabetes and on pregnancy outcomes. *Am J Epidemiol* 2000;**152**:1009–14.

87. Moses RG. Gestational diabetes: is a higher Cesarean section rate inevitable? *Diabetes Care* 2000;**23**:15–17.

88. Santini DL, Ales KL. The impact of universal screening for gestational glucose intolerance on outcome of pregnancy. *Surg Gynecol Obstet* 1990;**170**:427–36.

89. Sacks DA, Abu FS, Karten GJ, Forsythe AB, Hackett JR. Screening for gestational diabetes with the one-hour 50-g glucose test. *Obstet Gynecol* 1987;**70**:89–93.

90. Rouse DJ, Owen J, Goldenberg RL, Cliver SP. The effectiveness and costs of elective caesarean delivery for fetal macrosomia diagnosed by ultrasound. *JAMA* 1996;**276**:1480–6.

91. Roversi GD, Gargiulo M, Nicolini U, Pedretti E, Marini A, Barbarani V, *et al.* A new approach to the treatment of diabetic pregnant women. *Am J Obstet Gynecol* 1979;**135**:567–76.

92. Stenninger E, Schollin J, Aman J. Early postnatal hypoglycaemia in newborn infants of diabetic mothers. *Acta Paediatr* 1997;**86**:1374–6.

93. Owen J, Phelan ST, Landon MB, Gabbe SG. Gestational diabetes survey. *Am J Obstet Gynecol* 1995;**172**:615–20.

94. Magee MS, Walden CE, Benedetti TJ, Knopp RH. Influence of diagnostic criteria on the incidence of gestational diabetes and perinatal morbidity. *JAMA* 1993;**269**:609–15.

95. Khine ML, Winklestein A, Copel JA. Selective screening for gestational diabetes mellitus in adolescent pregnancies. *Obstet Gynecol* 1999;**93**:738–42.

96. Mello G, Parretti E, Mecacci F, Lucchetti R, Lagazio C, Pratesi M, *et al.* Risk factors for fetal macrosomia: the importance of a positive oral glucose challenge test. *Eur J Endocrinol* 1997;**137**:27–33.

97. Schwartz JG, Phillips WT, Blumhardt MR, Langer O. Use of a more physiologic oral glucose solution during screening for gestational diabetes mellitus. *Am J Obstet Gynecol* 1994;**171**:685–91.

98. Tardioli MC, Massi Benedetti M, Angeli G, Damiani F, Frascarelli A, Gigliarelli D, *et al.* A pilot study for the screening of gestational diabetes mellitus. *Diabetes Nutr Metab* 1993;**6**:377–80.

99. Berkowitz GS, Roman SH, Lapinski RH, Alvarez M. Maternal characteristics, neonatal outcome, and the time of diagnosis of gestational diabetes. *Am J Obstet Gynecol* 1992;**167**:976–82.

100. O'Sullivan JB, Mahan CM, Charles D, Danrow RV. Screening criteria for high risk gestational diabetic patients. *Am J Obstet Gynecol* 1973;**116**:895–900.

101. Bobrowski RA, Bottoms SF, Micallef JA, Dombrowski MP. Is the 50-gram glucose screening test ever diagnostic? *J Matern Fetal Med* 1996;**5**:317–20.

102. Corrado F, Stella NC, Mancuso A, Triolo O, Bruno L, Artenisio AC. Screening for gestational diabetes in Sicily. *J Reprod Med Obstet Gynecol* 1999;**44**:875–8.

103. Hong PL, Benjamin F, Deutsch S. First prenatal visit glucose screening. *Am J Perinatol* 1989;**6**:433–6.

104. Perucchini D, Fischer U, Spinas GA, Huch R, Huch A, Lehmann R. Using fasting plasma glucose concentrations to screen for gestational diabetes mellitus: prospective population based study. *BMJ* 1999;**319**:812–15.

105. Harlass FE, Brady K, Read JA. Reproducibility of the oral glucose tolerance test in pregnancy. *Am J Obstet Gynecol* 1991;**164**:564–8.

106. Neiger R, Coustan DR. The role of repeat glucose tolerance tests in the diagnosis of gestational diabetes. *Am J Obstet Gynecol* 1991;**165**:787–90.

107. Fachnie JD, Whitehouse FW, McGrath Z. Vomiting during OGTT in third trimester pregnancy. *Diabetes Care* 1988;**11**:18.

108. WHO Expert Committee on Diabetes Mellitus. Technical Report Series 646. Geneva: World Health Organization; 1980.

109. Kvetny J, Poulsen HF, Damgaard DW. Results from screening for gestational diabetes mellitus in a Danish county. *Dan Med Bull* 1999;**46**:57–9.

110. Deerochanawong C, Putiyanun C, Wongsuryrat M, Serirat S, Jinayon P. Comparison of National Diabetes Data Group and World Health Organization criteria for detecting gestational diabetes mellitus. *Diabetologia* 1996;**39**:1070–3.

111. Hanson U, Kallner A. Oral glucose tolerance test in pregnancy. Evaluation of a simplified procedure. *Acta Obstet Gynecol Scand* 1984;**63**:249–52.

112. Landy HJ, Gomez MO, O'Sullivan MJ. Diagnosing gestational diabetes mellitus: use of a glucose screen without administering the glucose tolerance test. *Obstet Gynecol* 1996;**87**:395–400.

113. Nelson-Piercy C, Gale EAM. Do we know how to screen for gestational diabetes – current practice in one regional health authority. *Diabet Med* 1994;**11**:493–8.

114. Chiaffarino F, Parazzini F, Bortolotti A, Benzi G. Scientific uncertainty is mirrored in clinical practice in Italy. *BMJ* 1998;**316**:861.

115. Dornhorst A, Paterson CM, Nicholls JSD, Wadsworth J, Chiu DC, Elkeles RS, *et al.* High prevalence of gestational diabetes in women from ethnic minority groups. *Diabet Med* 1992;**9**:820–5.

116. Marquette GP, Klein VR, Niebyl JR. Efficacy of screening for gestational diabetes. *Am J Perinatol* 1985;**2**:7–9.

117. Helton MR, Arndt J, Kebede M, King M. Do low-risk prenatal patients really need a screening glucose challenge test? *J Fam Pract* 1997;**44**:556–61.

118. Weeks JW, Major CA, de Veciana M, Morgan MA. Gestational diabetes: does the presence of risk factors influence perinatal outcome? *Am J Obstet Gynecol* 1994;**171**:1003–7.

119. Moses RG, Griffiths R, Davies W. Gestational diabetes: do all women need to be tested? *Aust N Z J Obstet Gynaecol* 1995;**35**:387–9.

120. Coustan DR. Methods of screening for and diagnosing of gestational diabetes. *Clin Perinatol* 1993;**20**:593–602.

121. Rodriguez H, Neiger R, Thompson S, List M, Krohn H. Screening adolescent gravidas for gestational diabetes. *Adolesc Pediatr Gynecol* 1995;**8**:125–7.

122. Truscello AM, Hollingsworth DR, Felice ME, Shragg P. Routine screening for gestational diabetes in white, black, and Mexican–American teenagers. *J Adolesc Health Care* 1988;**9**:150–5.

123. Court DJ, Mann SL, Stone PR, Goldsbury SM, Dixon-McIvor D, Baker JR. Comparison of glucose polymer and glucose for screening and tolerance tests in pregnancy. *Obstet Gynecol* 1985;**66**:491–9.

124. Marquette GP, Klein VR, Repke JT, Niebyl JR. Cost-effective criteria for glucose screening. *Obstet Gynecol* 1985;**66**:181–4.

125. Soltani H, Bruce C, Fraser RB. Observational study of maternal anthropometry and fetal insulin. *Arch Dis Child Fetal Neonatal Ed* 1999;**81**:F122–4.

126. Jimenez-Moleon JJ, Bueno-Cavanillas A, Luna-del-Castillo JD, Lardelli-Claret P, Garcia-Martin M, Galvez-Vargas R. Predictive value of a screen for gestational diabetes mellitus: influence of associated risk factors. *Acta Obstet Gynecol Scand* 2000;**79**:991–8.

127. Davey RX, Hamblin PS. Selective versus universal screening for gestational diabetes mellitus: an evaluation of predictive risk factors. *Med J Aust* 2001;**174**:118–21.

128. Onyeije CI, Divon M. Can early gestational weight gain predict macrosomia in pregnancies complicated by GDM? *Am J Obstet Gynecol* 1999;**180**:S40.

129. Branchtein L, Schmidt MI, Mengue SS, Reichelt AJ, Matos MCG, Duncan BB. Waist circumference and waist-to-hip ratio are related to gestational glucose tolerance. *Diabetes Care* 1997;**20**:509–11.

130. Newman RB. Maternal BMI can help determine which women need screening for GD [abstract]. *Am J Obstet Gynecol* 1999;**180**:S35.

131. Bundred P, Kitchiner D, Buchan I. Prevalence of overweight and obese children between 1989 and 1998: population based series of cross sectional studies. *BMJ* 2001;**322**:326–8.

132. Ford FA, Bruce C, Fraser RB. Fetal macrosomia in potential diabetics with normal oral glucose tolerance: a case control study. *Br J Obstet Gynaecol* 1990;**97**:957–9.

133. Gribble RK, Meier PR, Berg RL. The value of urine screening for glucose at each prenatal visit. *Obstet Gynecol* 1995;**86**:405–10.

134. Hooper DE. Detecting GD and preeclampsia. Effectiveness of routine urine screening for glucose and protein. *J Reprod Med* 1996;**41**:885–8.

135. Watson WJ. Screening for glycosuria during pregnancy. *South Med J* 1990;**83**:156–8.

136. Lind T, Hytten FM. The excretion of glucose during normal pregnancy. *J Obstet Gynaecol Br Commonw* 1972;**79**:961–5.

137. Corcoy R, Gascon N, De Leiva A, Ordonez LJ. Usual delay in sample processing can modify gestational diabetes screening. *Diabetes Care* 2000;**23**:429.

138. Lind T. Antenatal screening using random blood glucose values. *Diabetes* 1985;**34**:17–20.

139. Nielsen IK, Vinther S, Birch K, Lange AP. Random blood glucose sampling as an early antenatal screening test for diabetes mellitus. *Diabetes Res* 1988;**8**:31–3.

140. Jowett NI, Samanta AK, Burden AC. Screening for diabetes in pregnancy: is a random blood glucose enough? *Diabet Med* 1987;**4**:160–3.

141. McElduff A, Goldring J, Gordon P, Wyndham L. A direct comparison of the measurement of a random plasma glucose and a post-50 g glucose load glucose in the detection of gestational diabetes. *Aust N Z J Obstet Gynaecol* 1994;**34**:28–30.

142. Agarwal MM, Hughes PF, Ezimokhai M. Screening for gestational diabetes in a high-risk population using fasting plasma glucose. *Int J Gynecol Obstet* 2000;**68**:147.

143. Reichelt AJ, Spichler ER, Branchtein L, Nucci LB, Franco LJ, Schmidt MI. Fasting plasma glucose is a useful test for the detection of gestational diabetes. Brazilian Study of Gestational Diabetes (EBDG) Working Group. *Diabetes Care* 1998;**21**:1246–9.

144. Jones GC, Walker JD. Poor sensitivity of fasting blood glucose as a screening test for gestational diabetes mellitus and impaired glucose tolerance of pregnancy. *Diabet Med* 1996;**13** Suppl 7:S40.

145. Sacks DA, Greenspoon JS, Fotheringham N. Could the fasting plasma glucose assay be used to screen for gestational diabetes? *J Reprod Med Obstet Gynecol* 1992;**37**:907–9.

146. Yalcin HR, Zorlu CG. Threshold value of glucose screening tests in pregnancy: could it be standardized for every population? *Am J Perinatol* 1996;**13**:317–20.

147. Danilenko-Dixon DR, Van Winter JT, Nelson RL, Ogburn PL. Universal versus selective gestational diabetes screening: application of 1997 American Diabetes Association recommendations. *Am J Obstet Gynecol* 1999;**181**:798–802.

148. Lavin J. Screening of high-risk and general populations for gestational diabetes. Clinical application and cost analysis. *Diabetes* 1985;**34**:24–7.

149. Meriggi E, Trossarelli GF, Carta Q, Menato G, Porta MA, Bordon R, *et al.* Capillary glucose determination in the screening of gestational diabetes. *Diabetes Res Clin Pract* 1988;**5**:55–61.

150. Metzger BE, the Organizing Committee. Summary and recommendations of the Third International Workshop-Conference on Gestational Diabetes Mellitus. *Diabetes* 1991;**40** Suppl 2:197–201.

151. Berkus MD, Stern MP, Mitchell BD, Abashawl A, Langer O. Relationships between glucose levels and insulin secretion during a glucose challenge test. *Am J Obstet Gynecol* 1990;**163**:1818–22.

152. Lewis GF, McNally C, Blackman JD, Polonsky KS, Barron WM. Prior feeding alters the response to the 50-g glucose challenge test in pregnancy: the Staub–Traugott effect revisited. *Diabetes Care* 1993;**16**:1551–6.

153. Cetin M, Cetin A. Time-dependent gestational diabetes screening values. *Int J Gynecol Obstet* 1997;**56**:257–61.

154. Kirkpatrick C, Schwers J, Desir D. Prenatal screening for gestational diabetes throughout office hours. *Eur J Obstet Gynecol Reprod Biol* 1988;**27**:299–306.

155. Watson WJ. Serial changes in the 50-g oral glucose test in pregnancy: implications for screening. *Obstet Gynecol* 1989;**74**:40–3.

156. Jovanovic L, Peterson CM. Screening for gestational diabetes. Optimum timing and criteria for retesting. *Diabetes* 1985;**34**:21–3.

157. Nahum GG, Huffaker BJ. Correlation between first- and early third-trimester glucose screening test results. *Obstet Gynecol* 1990;**76**:709–13.

158. Benjamin F, Wilson SJ, Deutsch S, Seltzer VL, Droesch K, Droesch J. Effect of advancing pregnancy on the glucose tolerance test and on the 50-g oral glucose load screening test for gestational diabetes. *Obstet Gynecol* 1986;**68**:362–5.

159. Super DM, Edelberg SC, Philipson EH, Hertz RH, Kalhan SC. Diagnosis of gestational diabetes in early pregnancy. *Diabetes Care* 1991;**14**:288–94.

160. Meyer WJ, Carbone J, Gauthier DW, Gottmann DA. Early gestational glucose screening and gestational diabetes. *J Reprod Med Obstet Gynecol* 1996;**41**:675–9.

161. Shivvers SA, Lucas MJ. Gestational diabetes: is a 50-g screening result ≤ 200 mg/dL diagnostic? *J Reprod Med Obstet Gynecol* 1999;**44**:685–8.

162. Menon U, Ranjan M, Jasper P, Oommen A. Evaluation of plasma fructosamine as a screening test for gestational diabetes. *Aust N Z J Obstet Gynaecol* 1991;**31**:25–6.

163. Kerbel D, Glazier R, Holzapfel S, Yeung M, Lofsky S. Adverse effects of screening for gestational diabetes: a prospective cohort study in Toronto, Canada. *J Med Screen* 1997;**4**:128–32.

164. Jirapinyo M, Puavilai G, Chanprasertyotin S, Tangtrakul S. Predictive value of 1 hour 50 g oral glucose load screening test for gestational diabetes mellitus compared to 3 hour oral glucose tolerance test in high risk pregnant women. *Asia Oceania J Obstet Gynaecol* 1993;**19**:7–12.

165. Neilson J, Bolton RN, Prins RP, Mark C. Glucose challenge testing in pregnancy. *Am J Obstet Gynecol* 1991;**164**:1673–9.

166. Sacks DA, Abu FS, Greenspoon JS, Fotheringham N. How reliable is the fifty-gram, one-hour glucose screening test? *Am J Obstet Gynecol* 1989;**161**:642–5.

167. Weiner CP, Fraser MM, Burns JM, Schnoor D, Herrig J, Whitaker LA. Cost efficacy of routine screening for diabetes in pregnancy: 1-h versus 2-h specimen. *Diabetes Care* 1986;**9**:255–9.

168. Bergus GR, Murphy NJ. Screening for gestational diabetes mellitus: comparison of a glucose polymer and a glucose monomer test beverage. *J Am Board Fam Pract* 1992;**5**:241–7.

169. Lamar ME, Kuehl TJ, Cooney AT, Gayle LJ, Holleman S, Allen SR. Jelly beans as an alternative to a fifty-gram glucose beverage for gestational diabetes screening. *Am J Obstet Gynecol* 1999;**181**:1154–7.

170. Coustan DR, Widnes JA, Carpenter MW, Rotondo L, Pratt DC. The 'breakfast tolerance test': screening for gestational diabetes with a standardized mixed nutrient meal. *Am J Obstet Gynecol* 1987;**157**:1113–17.

171. Vermes I, Zeyen LJJM, Van Roon E, Brandts H. The role of serum fructosamine as a screening test for gestational diabetes mellitus. *Horm Metab Res* 1989;**21**:73–6.

172. Nasrat HA, Ajabnoor MA, Ardawi MSM. Fructosamine as a screening test for gestational diabetes mellitus: a reappraisal. *Int J Gynecol Obstet* 1991;**34**:27–33.

173. Hughes PF, Agarwal M, Newman P, Morrison J. An evaluation of fructosamine estimation in screening for gestational diabetes mellitus. *Diabet Med* 1995;**12**:708–12.

174. Bor MV, Bor P, Cevik C. Serum fructosamine and fructosamine-albumin ratio as screening tests for gestational diabetes mellitus. *Arch Gynecol Obstet* 1999;**262**:105–11.

175. Grandjean H, Sarramon MF, De Mouzon J, Reme JM, Pontonnier G. Detection of gestational diabetes by means of ultrasonic diagnosis of excessive fetal growth. *Am J Obstet Gynecol* 1980;**138**:790–2.

176. Griffin ME, Coffey M, Johnson H, Scanlon P, Foley M, Stronge J, *et al.* Universal vs. risk factor-based screening for gestational diabetes mellitus: detection rates, gestation at diagnosis and outcome. *Diabet Med* 2000;**17**:26–32.

177. Dietrich ML, Dolnicek TF, Rayburn WF. Gestational diabetes screening in a private, midwestern American population. *Am J Obstet Gynecol* 1987;**156**:1403–8.

178. Rust O, Bofill JA, Carroll SC, Cowan BD, Martin RW, Morrison JC. Two-hour postprandial test versus one-hour, fifty-gram glucola test as screening tools for gestational diabetes: a critical analysis. *J Perinatol* 1998;**18**:49–54.

179. Bevier WC, Fischer R, Jovanovic L. Treatment of women with an abnormal glucose challenge test (but a normal oral glucose tolerance test) decreases the prevalence of macrosomia. *Am J Perinatol* 1999;**16**:269–75.

180. Bebbington MW, Milner R, Wilson RD, Harris S. A randomized controlled trial comparing routine screening vs. selected screening for gestational diabetes in low risk population. *Am J Obstet Gynecol* 1999;**180**:S36.

181. Casey BM, Lucas MJ, McIntire DD, Leveno KJ. Pregnancy outcomes in women with gestational diabetes compared with the general obstetric population [see comments]. *Obstet Gynecol* 1997;**90**:869–73.

182. Uncu G, Ozan H, Cengiz C. The comparison of 50 grams glucose challenge test, HbA_{1c} and fructosamine levels in diagnosis of gestational diabetes mellitus. *Clin Exp Obstet Gynecol* 1995;**22**:230–4.

183. Corcoy R, Cerqueira MJ, Pedreno J, Matas J, Codina M, Pou JM, *et al.* Serum fructosamine is not a useful screening test for gestational diabetes. *Eur J Obstet Gynecol Reprod Biol* 1991;**38**:217–20.

184. Mathai M, Thomas TJ, Kuruvila S, Jairaj P. Random plasma glucose and the glucose challenge test in pregnancy. *Natl Med J India* 1994;**7**:160–2.

185. Fuhrmann K. Targets in oral glucose tolerance. In: Sutherland HW, Stowers JM, Pearson DWM, editors. Carbohydrate metabolism in pregnancy and the newborn. London: Springer-Verlag; 1989. p. 227–37.

186. Tam WH, Rogers MS, Yip SK, Lau TK, Leung TY. Which screening test is the best for gestational impaired glucose tolerance and gestational diabetes mellitus? *Diabetes Care* 2000;**23**:1432.

187. Reece EA, Holford T, Tuck S, Bargar M, O'Connor T, Hobbins JC. Screening for gestational diabetes: one-hour carbohydrate tolerance test performed by a virtually tasteless polymer of glucose. *Am J Obstet Gynecol* 1987;**156**:132–4.

188. Murphy NJ, Meyer BA, O'Kell RT, Hogard ME. Carbohydrate sources for gestational diabetes mellitus screening. A comparison. *J Reprod Med* 1994;**39**:977–81.

189. Zhang Y, Warren PM, Sakura H, Adelman J, Stoffel M, Bell GI, *et al.* No evidence for mutations in a putative beta-cell ATP-sensitive K(+) channel subunit in MODY, NIDDM, or GDM. *Diabetes* 1995;**44**:597–600.

190. Bloomgarden ZT. American Diabetes Association 60th Scientific Sessions, 2000. *Diabetes Care* 2000;**23**:1699–702.

191. Swinker M. Routine screening for gestational diabetes mellitus in a family practice center. *J Fam Pract* 1983;**17**:611–14.

192. Lemen PM, Wigton TR, Miller-McCarthey AJ, Cruikshank DP, Hiss LW, Van Winter J, *et al.* Screening for gestational diabetes mellitus in adolescent pregnancies. *Am J Obstet Gynecol* 1998;**178**:1251–6.

193. Netten A, Curtis L, compilers. Unit costs of health and social care 2000. Canterbury: University of Kent, Personal Social Services Research Unit (PSSRU), 2000.

194. MIMS. Monthly Index of Medical Specialities. Haymarket Medical Ltd. 29-1-0001. URL: http://www.emims.net

195. Berwick DM, Weinstein MC. What do patients value? Willingness to pay for ultrasound in normal pregnancy. *Med Care* 1985;**23**:881–93.

196. Mooney G. The valuation of human life. London: Macmillan; 1977.

197. Comtois R, Desjarlais F, Nguyen M, Beauregard H. Clinical usefulness of estimation of serum fructosamine concentration as screening test for gestational diabetes. *Am J Obstet Gynecol* 1989;**160**:651–4.

198. Office of National Statistics. Health Statistics Quarterly. 8. London: ONS; 2000.

199. Maresh M, Beard RW, Bray CS, Elkeles RS, Wadsworth J. Factors predisposing to and outcome of gestational diabetes. *Obstet Gynecol* 1989;**74**:342–6.

200. Cousins L, Baxi L, Chez R, Coustan DR, Gabbe S, Harris J, *et al.* Screening recommendations for gestational diabetes mellitus. *Am J Obstet Gynecol* 1991;**165**:493–6.

201. Howard ED. Gestational diabetes mellitus screening tests: a review of current recommendations. *J Neonatal Nurs* 1992;**6**:37–42.

202. Damm P, Kuhl C, Bertelsen A, Molsted-Pederson L. Predictive factors for the development of diabetes in women with previous gestational diabetes mellitus. *Am J Obstet Gynecol* 1992;**167**:607–16.

203. Blank A, Metzger BE, Grave GD. Effects of gestational diabetes on perinatal morbidity reassessed: report of the International Workshop on Adverse Perinatal Outcomes of Gestational Diabetes Mellitus. *Diabetes Care* 1995;**18**:127–9.

204. Rossi M, Dornhorst A. Diabetes following gestational diabetes mellitus. In: Dornhorst A, Hadden D, editors. Diabetes and pregnancy: an international approach to diagnosis and management. Chichester: John Wiley & Sons; 1996.

205. Currie CJ, Kraus D, Morgan CL, Gill L, Stott NC, Peters JR. NHS acute sector expenditure for diabetes. *Diabet Med* 1997;**14**:686–92.

206. Grant PT, Oats JN, Beischer NA. The long term follow up of women with gestational diabetes. *Aust N Z J Obstet Gynaecol* 1986;**26**:17–22.

207. Kjos SL, Peters RK, Henry OA, Montoro MN, Buchanan TA. Predicting future diabetes in Latino women with gestational diabetes. *Diabetes* 1995;**44**:586–91.

208. O'Sullivan JB. Diabetes mellitus after GDM. *Diabetes* 1991;**29**:131–5.

209. Coustan DR. Management of gestational diabetes mellitus. *JAMA* 1996;**275**:1199–200.

210. Jovanovic-Peterson L, Peterson CM. Is exercise safe or useful for gestational diabetic women? *Diabetes* 1991;**40**:179–81.

211. Homko CJ, Reece EA. To screen or not to screen for gestational diabetes mellitus. The clinical quagmire. *Clin Perinatol* 2001;**28**:407–17.

212. de Veciana M, Major CA, Morgan MA, Asrat T, Toohey JS, Lien JM, *et al.* Postprandial versus preprandial blood glucose monitoring in women with gestational diabetes mellitus requiring insulin therapy. *N Engl J Med* 1995;**353**:1237–41.

213. Lauszus FF, Paludan J, Kebe JG. Birthweight in women with potential gestational diabetes mellitus – an effect of obesity rather than glucose intolerance? *Acta Obstet Gynecol Scand* 1999;**78**:520–5.

214. Lind T, Anderson J. Does random blood glucose sampling outdate testing for glycosuria in the detection of diabetes during pregnancy? *BMJ* 1984;**289**:1569–71.

215. Benjamin E, Winters D, Mayfield J, Gohdes D. Diabetes in pregnancy in Zuni Indian women. Prevalence and subsequent development of clinical diabetes after gestational diabetes. *Diabetes Care* 1993;**16**:1231–5.

216. American Diabetes Association. Summary and recommendations of the Third International Workshop-Conference on Gestational Diabetes Mellitus. *Diabetes* 1985;**40** Suppl 2:123–6.

217. Begley JP, Stafford KA, Clark JDA. Fructosamine as a screening test for diabetes mellitus. *Pract Diabetes* 1992;**9**:104–5.

218. Huter O, Drexel H, Brezinka C, Soelder E, Koelle D, Patsch JR. Low sensitivity of serum fructosamine as a screening parameter for gestational diabetes mellitus. *Gynecol Obstet Invest* 1992;**34**:20–3.

219. Al-Shawaf T, Moghraby S, Akiel A. Does impaired glucose tolerance imply a risk in pregnancy? *Br J Obstet Gynaecol* 1988;**95**:1036–41.

220. Ammari F, Gregory R. Screening for gestational diabetes in a population at high risk. *Pract Diabetes Int* 1996;**13**:150–2.

221. Ardawi MSM, Nasrat HA, Jamal HS, Al Sagaaf HM, Mustafa BE. Screening for gestational diabetes mellitus in pregnant females. *Saudi Med J* 2000;**21**:155–60.

222. Artal R, Mosley GM, Dorey FJ. Glycohemoglobin as a screening test for gestational diabetes. *Am J Obstet Gynecol* 1984;**148**:412–14.

223. Bartha JL, Martinez DF, Comino DR. Gestational diabetes mellitus diagnosed during early pregnancy. *Am J Obstet Gynecol* 2000;**182**:346–50.

224. Baxi L, Barad D, Reece EA, Farber R. Use of glycosylated hemoglobin as a screen for macrosomia in gestational diabetes. *Obstet Gynecol* 1984;**64**:347–50.

225. Berkus MD, Langer O. Glucose tolerance test: degree of glucose abnormality correlates with neonatal outcome. *Obstet Gynecol* 1993;**81**:344–8.

226. Campbell DM, Sutherland HW, Tuttle S. Standardized test meal in human pregnancy. In: Sutherland HW, Stowers JM, Pearson DWM, eds. Carbohydrate metabolism in pregnancy and the newborn IV. London: Springer-Verlag; 1989. p. 257–65.

227. Cooper A, Granat M, Sharf M. Glucose intolerance during pregnancy. II. A comparative study of diagnostic screening methods. *Obstet Gynecol* 1979;**53**:495–9.

228. Di Cianni G, Benzi L, Casadidio I, Orsini P, Rossi L, Fontana G, *et al.* Screening of gestational diabetes in Tuscany: results in 2000 cases. *Ann 1st Super Sanita* 1997;**33**:389–91.

229. de los Monteros AE, Parra A, Ramirez A. The reproducibility of the 50 g 1-hour glucose screen for diabetes in pregnancy. *Obstet Gynecol* 1993;**82**:515–18.

230. Gillmer MD, Bickerton NJ. Advances in the management of diabetes in pregnancy: success through simplicity. In: Bonnar J, editor. Recent advances in obstetrics and gynaecology. Edinburgh: Churchill Livingstone: 1994.

231. Dornhorst A, Nicholls JSD, Welch A, Ali K, Chan SP, Beard RW. Correcting for ethnicity when defining large for gestational age infants in diabetic pregnancies. *Diabetic Med* 1996;**13**:226–31.

232. Fedele D, Lapolla A. A protocol of screening of gestational diabetes mellitus. *Ann 1st Super Sanita* 1997;**33**:383–7.

233. Forest JC, Masse J, Garrido RM. Glucose tolerance test during pregnancy: the significance of one abnormal value. *Clin Biochem* 1994;**27**:299–304.

234. Greenberg LR, Moore TR, Murphy H. Gestational diabetes mellitus: antenatal variables as predictors of postpartum glucose intolerance. *Obstet Gynecol* 1995;**86**:97–101.

235. Gregory R, Swinn RA, Wareham N, Curling V, Dalton KJ, Edwards OM, *et al.* An audit of a comprehensive screening programme for diabetes in pregnancy. *Pract Diabetes Int* 1998;**15**:45–8.

236. Hadden D. Evidence-based screening for gestational diabetes? *Diabet Med* 2000;**17**:402.

237. Hughes PF, Agarwal M, Newman P, Morrison J. Screening for gestational diabetes in a multi-ethnic population. *Diabetes Res Clin Pract* 1995;**28**:73–8.

238. Knopp RH, Magee MS, Walden CE, Bonet B, Benedetti TJ. Prediction of infant birth-weight by GDM screening tests – importance of plasma triglyceride. *Diabetes Care* 1992;**15**:1605–13.

239. Landon MB, Gabbe SG, Sachs L. Management of diabetes mellitus and pregnancy: a survey of obstetricians and maternal-fetal specialists. *Obstet Gynecol* 1990;**75**:635–40.

240. Lavin JP, Barden TP, Miodovnik M. Clinical experience with a screening program for gestational diabetes. *Am J Obstet Gynecol* 1981;**141**:491–4.

241. Lind T, Phillips PR. Influence of pregnancy on the 75-g OGTT: a prospective multicenter study. *Diabetes* 1991;**40**:8–13.

242. Lindsay MK, Graves W, Klein L. The relationship of one abnormal glucose tolerance test value and pregnancy complications. *Obstet Gynecol* 1989;**73**:103–6.

243. Merkatz IR, Duchon MA, Yamashita TS, Houser HB. A pilot community-based screening program for gestational diabetes. *Diabetes Care* 1980;**3**:453–7.

244. Moses RG, Fulwood S, Griffiths BN, Griffiths R. Gestational diabetes mellitus: resource utilization and costs of diagnosis and treatment. *Aust N Z J Obstet Gynaecol* 1997;**37**:184–6.

245. Narchi H, Kulaylat N. High incidence of Down's syndrome in infants of diabetic mothers. *Arch Dis Child* 1997;**77**:242–4.

246. Rajab KE, Mehdi S. Pregnancy outcome among gestational diabetics with blood glucose levels between 7.7 and 8.3 mmol/l. *Int J Gynaecol Obstet* 1998;**63**:59–61.

247. Rey E. Usefulness of a breakfast test in the management of women with gestational diabetes. *Obstet Gynecol* 1997;**89**:981–8.

248. Rust OA, Bofill JA, Andrew ME, Kincaid TA, Stubbs TM, Miller EH, *et al.* Lowering the threshold for the diagnosis of gestational diabetes. *Am J Obstet Gynecol* 1996;**175**:961–5.

249. Shah BD, Cohen AW, May C, Gabbe SG. Comparison of glycohemoglobin determination and the one-hour oral glucose screen in the identification of gestational diabetes. *Am J Obstet Gynecol* 1982;**144**:774–7.

250. Solomon CG, Willett WC, Rich EJ, Hunter DJ, Stampfer MJ, Colditz GA, *et al.* Variability in diagnostic evaluation and criteria for gestational diabetes. *Diabetes Care* 1996;**19**:12–16.

251. Solomon CG, Willett WC, Carey VJ, Rich EJ, Hunter DJ, Colditz GA, *et al.* A prospective study of pregravid determinants of gestational diabetes mellitus. *JAMA* 1997;**278**:1078–83.

252. Thaisz E, Rappai A, Fovenyi J, Zavodi E. Screening and care of gestational and insulin-dependent diabetic pregnancies: the first four years' experience. *Diabetes Nutr Metab Clin Exp* 1993;**6**:373–5.

253. Van Turnhout HEW, Lotgering FK, Wallenburg HCS. Poor sensitivity of the fifty-gram one-hour glucose screening test for hyperglycemia. *Eur J Obstet Gynecol Reprod Biol* 1994;**53**:7–10.

254. Verma A, Mitchell BF, Demianczuk N, Flowerdew G, Okun NB. Relationship between plasma glucose levels in glucose-intolerant women and newborn macrosomia. *J Matern Fetal Med* 1997;**6**:187–93.

255. Williams CB, Iqbal S, Zawacki CM, Yu D, Brown MB, Herman WH. Effect of selective screening for gestational diabetes. *Diabetes Care* 1999;**22**:418–21.

256. Young C, Kuehl TJ, Sulak PJ, Allen SR. Gestational diabetes screening in subsequent pregnancies of previously healthy patients. *Am J Obstet Gynecol* 2000;**182**:1024.

Appendix 1

Search strategy

Cochrane Library 2000 issue 3

Searched on (gestation* near diabet*) and screen*

MEDLINE 1966–2000/08

#1 explode "Mass-Screening"/all subheadings
#2 screen*
#3 "Diabetes-Gestational"/all subheadings
#4 diabet*
#5 gestation*
#6 diabet* near2 gestation*
#7 gdm
#8 #1 or #2
#9 #3 or #6 or #7
#10 #8 and #9
#11 TG = "HUMAN"
#12 #10 and (TG = "HUMAN")
#13 LA = "ENGLISH"
#14 #12 and (LA = "ENGLISH")
#15 pt=editorial
#16 pt=letter
#17 pt=comment
#18 #15 or #16 or #17
#19 #14 not #18
#20 PT = "CLINICAL-TRIAL"
#21 #19 and (PT = "CLINICAL-TRIAL")

and

#1 explode "Mass-Screening"/all subheadings
#2 screen*
#3 "Diabetes-Gestational"/all subheadings
#4 diabet*
#5 gestation*
#6 diabet* near2 gestation*
#7 gdm
#8 #1 or #2
#9 #3 or #6 or #7
#10 #8 and #9
#11 TG = "HUMAN"
#12 #10 and (TG = "HUMAN")
#13 LA = "ENGLISH"
#14 #12 and (LA = "ENGLISH")
#15 pt=editorial
#16 pt=letter
#17 pt=comment
#18 #15 or #16 or #17
#19 #14 not #18
#20 PT = "CLINICAL-TRIAL"
#21 #19 and (PT = "CLINICAL-TRIAL")
#22 tg=comparative-study
#23 explode "Cohort-Studies"/all subheadings
#24 explode "Case-Control-Studies"/ all subheadings
#25 explode "Research-Design"/all subheadings
#26 #22 or #23 or #24 or #25
#27 #23 or #24 or #25 or #26
#28 #19 and #27
#29 #28 not #21

EMBASE 1981–2000/06

#1	2,685	explode "pregnancy-diabetes-mellitus"/all subheadings
#2	56,070	gestation*
#3	134,146	diabet*
#4	1,692	gestation* near2 diabet*
#5	76,211	explode "screening"/ all subheadings
#6	3,080	#1 or #4
#7	319	#6 and #5
#8	56,070	gestation*
#9	134,146	diabet*
#10	155,023	screen*
#11	418	(gestation* near2 diabet*) with screen*
#12	520	#7 or #11
#13	201	#12 not #7
#14	184,085	explode "clinical-trial"/ all subheadings
#15	31	#11 and #14

and

#1 explode "pregnancy-diabetes-mellitus"/ all subheadings
#2 gestation*
#3 diabet*
#4 gestation* near2 diabet*
#5 explode "screening"/all subheadings
#6 #1 or #4
#7 #6 and #5
#8 gestation*
#9 diabet*
#10 screen*
#11 (gestation* near2 diabet*) with screen*
#12 #7 or #11

#13 #12 not #7
#14 explode "clinical-trial"/all subheadings
#15 #11 and #14
#16 explode "clinical-study"/all subheadings
#17 #11 and #16
#18 #17 not #15
#19 LA = "ENGLISH"
#20 #18 and (LA = "ENGLISH")

PubMed

For the current year on (gestation* AND diabet* AND screen*)

SCI and SSCI

On Web of Science on (gestation* same diabet*) and screen* and random*

NRR 2000 issue 2

Saved completed studies as file NRRcomplet

Appendix 2

Data extraction forms

Studies using FPG

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Agarwal *et al.*, 2000[142]	United Arab Emirates Indian subcontinent (29.1%), Arabs from Jordan, Lebanon and Syria (18.8%), UAE (19.1%), North Africa (14.2%) and East Africa (8.4%) Mean age not given	Observational study, not randomised 430 women *Aim*: to assess the potential value of FBG as a screening test for GDM	Patients who were already referred for a GTT on the basis of a GCT or clinical grounds (GCT or 'clinical grounds' not defined) were available for the study. GDM was diagnosed on the basis of C&C's modified criteria[20] Fasting blood sugar taken from GTT	116 (27%) had GDM (high incidence in this population) Used a 'rule in rule out' strategy for analysing the value of the FBG as a screening test. This involves considering two cut-off values; the higher value, which has an increased specificity, is used to rule in the disease and the lower cut-off with its increased sensitivity to rule out the disease in question
Sensitivity and specificity Using an FBG value ≥ 5.3 mmol/l with a sensitivity of 48.3% and specificity of 97.5% would have eliminated the GTT in 64 of 430 patients Using a lower cut-off of ≥ 4.3 mmol/l with a sensitivity of 93% and specificity of 38.5%, 129 patients would not need GTT Thus fasting glucose alone would have eliminated the need for almost 45% of GTTs, with only 16 (3.7%) being misclassified				
Comments Paper only a brief communication, many aspects of the method and results are lacking Results probably only applicable to local population				

Studies using FPG *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Fuhrmann, 1989[185]	Berlin (GDR) Sample selected from three typical districts in Berlin to ensure representative socio-economic status of sample	Not RCT Prospective cohort *Aim*: to assess: 1. Incidence of GDM 2. The reproducibility of 50-g and 75-g GTTs 3. Screening using traditional RFs 4. Screening using FBG 5. Screening using 1-h and 2-h GCTs 6. Follow-up of women with abnormal glucose tolerance to record later development of DM 2510 women tested with GTT, screened using FBG, 1-h and 2-h glucose load, and RFs Reproducibility of 50-g GTT assessed by repeating GTT after 14 days in 147 women Reproducibility of 75-g GTT assessed by repeating 75-g GTT in 63 women (53 normal and 10 GDM) 8-year follow-up of 50 women with GDM	Capillary whole blood tested Screening test used: FBG (various thresholds); 1-h and 2-h post 50 g glucose load (various thresholds); traditional RFs (see below) Diagnostic test: all women had 50-g OGTT between 26 and 36 weeks GDM diagnosed when 2 or more values ≥ mean ± 2 SD for sample: fasting 5.55; 1-h 8.88; 2-h 7.22 mmol/l Normal when 2 or 3 values ≤ mean ± 1 SD IGT when between normal and GDM Follow-up post-partum using 75-g GTT at 3 months, 6 months, 1 year and 8 years after pregnancy DM diagnosed according to WHO criteria[24]	1% diagnosed as GDM (2 or more values exceeding mean ± 2 SD for sample) 6% diagnosed as IGT Overall reproducibility of 50-g GTT (147 women) = 70% Reproducibility for those testing normal first time = 74%; IGT first time = 19.5%; GDM first time = 0% Insulin response was not significantly different between normal, IGT and GDM groups Follow-up post-partum (50 women with GDM): 3 months: 4% (2/50) had DM; 4% had IGT; 92% were normal 1 year: 18% (9/50) had DM; 20% (10/50) had IGT; 62% were normal 8 years: 46% (23/50) had DM; 16% (8/50) had IGT; 38% were normal Most had NIDDM Follow-up of 29 women with IGT: 8 years: 3% (1/29) had DM; 28% (8/29) had IGT; 69% (20/29) were normal

Sensitivity and specificity

Screening using traditional RFs:

At least one of traditional RFs (see below) present in 85.5%

Single RF detected between 13% (previous infant ≥ 4000 g or proteinuria) and 72% (previous deliveries) of those with GDM

Combination of two RFs detected a maximum of 31% of GDM; three RFs detected maximum of 17% GDM

Screening using FBG (estimated from % reported):

FBG ≥ 4.4 mmol/l: sensitivity 100% (28/28); specificity 74% (1837/2482); PPV 4% (28/673). 27% would need GTT

FBG ≥ 4.72 mmol/l: sensitivity 93% (26/28); specificity 83% (2061/2482); PPV 6% (26/447)

FBG ≥ 5.0 mmol/l: sensitivity 86% (24/28); specificity 89% (2197/2482); PPV 8% (24/309)

FBG ≥ 5.5 mmol/l: sensitivity 75% (21/28); specificity 95% (2362/2482); PPV 15% (21/141). 5.6% would need GTT

Screening using 1-h and 2-h GCT:

1 h post glucose load ≥ 7.22 mmol/l: sensitivity 96% (27/28); specificity 80% (1992/2482); PPV 5% (27/517). 21% (517/2510) would need GTT

1 h post glucose load ≥ 7.77 mmol/l: sensitivity 96% (27/28); specificity 90% (2233/2482); PPV 10% (27/276). 11% (276/2510) would need GTT

2 h post-50 g glucose load ≥ 6.38 mmol/l: sensitivity 75% (21/28); specificity 93% (2296/2482); PPV 10% (21/207)

2 h post-50 g glucose load ≥ 6.94 mmol/l: sensitivity 75% (21/28); specificity 96% (2388/2482); PPV 18% (21/115)

Authors recommend the following screening regime in GDR:

1. Assessment of RFs at 24 weeks (glycosuria before 24 weeks; overweight ≥ 30%; age ≥ 30 years; heredity – especially the mother): identifies ~20% of women

2. FBG. If FBG ≥ 4.4 mmol/l proceed to GTT

3. Diagnostic GTT

Author's conclusions: Reproducibility of a 50-g GTT is good only in pregnant women with normal CHO metabolism; a pathological test is poorly reproducible

The 75-g GTT has no better reproducibility than the 50-g GTT

Women with one pathological GTT are at risk and should be followed up during pregnancy and afterwards

Screening for GDM is possible with BG values in the fasting as well as the post-GCT. Screening with RFs and fasting or GCT is less expensive

Comments

Reports rates of DM after 8 years of 50 women with GDM

All women underwent diagnostic GTT between 26 and 36 weeks

May have underestimated GDM by not detecting women developing GDM later in pregnancy

Very low incidence of GDM reported. GDM diagnosed on basis of population BG values and not according to fetal/maternal outcomes. So sample may not be representative of women with GDM by other criteria. Only 28 women with GDM detected, so evaluation of screening tests based on very small sample size

Small subgroup sizes for assessing reproducibility of 50-g GTT (5 women in GDM group)

No details of characteristics of sample

No cost evaluation of various screening tests

No results reported from an evaluation of a combination of RFs and other screening tests

Studies using FPG *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lauszus *et al.*, 1999[213]	Aarhus, Denmark Inclusion criteria: charts of all women undergoing OGTT for 1992–93 Exclusion criteria: multiple pregnancy; one woman who required insulin Mean age: GDM 31 years vs control 28 years ($p < 0.01$); mean parity: GDM 2.2 vs control 1.8 ($p < 0.05$)	Not RCT Retrospective review of case records *Aim*: to compare the perinatal outcomes of women with GDM, women with a non-diabetic OGTT and the background population 383 case records of GTT reviewed Included 55 women with GDM treated with diet 74 borderline diabetic, and 254 non-diabetic GTT normal women (control) Background population contained 8196 women who seem not to have been tested but are reported to have excluded DM and immunisation (this latter term is not defined)	Screening test used: FBG if PH obesity, FH DM, previous GDM, birth weight > 4500 g, stillbirth, age > 38 years, current glycosuria If two independent FBGs > 4.6 mmol/l referred for GTT Timing of screening test not stated Diagnostic test: 3-h 75-g GTT GDM diagnosed when at least two CPG levels were above mean ± 3 SD: fasting 6.4; 30-min 13.6; 1-h 13.7; 90-min 11.0; 2-h 10.2; 150-min 9.7; 3-h 8.5 mmol/l Advised re diet. Admitted if FBG > 6 mmol/l. Insulin started when mean BG > 7.5 mmol/l Non-GDM women classified as: Borderline diabetic if at least one value > mean ± 2 SD at GTT and max. one value > mean ± 3 SD Normal: all values < mean ± 2 SD > 85% of women without GDM had at least two GTTs to exclude GDM	Denominator for calculation of GDM rates not stated Women with GDM were older and had higher non-pregnant BMI and parity than normal women Infant outcomes: there were more macrosomic infants in the borderline diabetic group than in any of the other groups. Groups who underwent GTT testing had higher rates of macrosomic infants than the background population Birth weight > 4000 g: GDM 27% vs borderline 42% vs normal GTT 27% vs background 17% ($p < 0.01$ for background vs all others) Multiple regression adjusted for age, parity and gestational week showed birth weight was associated with maternal weight at GTT during pregnancy and at delivery 1 week after delivery: 19% (9/47) had borderline diabetic GTT; 19% had diabetic GTT 1 year after delivery: 9% (4/47) had diabetic GTT and 2% (1/47) had borderline GTT

Sensitivity and specificity
N/A
Mean weight gain was greater in women with non-diabetic GTT: GDM 8 kg vs borderline 15 kg vs control 15 kg ($p < 0.01$)
CS was significantly more common in the GDM group compared to the others ($p < 0.01$): GDM 21% (12/55) vs borderline 16% (12/74) vs normal 12% (30/254) vs background 9% (756/8196). Higher but non-significantly increased rates were noted in both acute and elective CS rates (small numbers in these subgroups: acute CS: GDM 9/55 vs borderline 11/74 vs normal 23/254 vs background 475/8196; elective CS: GDM 3/55 vs borderline 1/74 vs normal 7/254 vs background 281/8196)
Authors' comment: a diagnosis of GDM may shift obstetric practice towards CS
No significant difference was noted in rates of hypertensive disorders during pregnancy, trauma due to delivery of macrosomic neonate or neonatal morbidity, although again numbers with adverse outcomes were very small in GDM, borderline and normal groups
Authors' conclusions: even minor hyperglycaemia is associated with increasing birth weight. Birth weight is reduced in GDM when dietary treatment is instituted and effect on weight gain is achieved

Comments
Ethnicity of sample not described
Assumes all cases of GDM detected by screening using two independent FBGs > 4.6 mmol/l in population considered to be at high risk followed by abnormal GTT at whatever time in pregnancy this was done
Criteria for GDM diagnosis determined on basis of centiles of population BG levels rather than fetal/maternal outcomes
Gestation when GDM diagnosed and dietary measure instituted was not reported
No costs reported
Evidence presented supports authors' conclusions. Small numbers of some adverse outcomes may have limited the power to detect significant differences

Studies using FPG *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Reichelt *et al.*, 1998[143]	Six state capitals in Brazil Sample collected 1991–95 Inclusion criteria: women aged ≥ 20 years, with no diagnosis of DM and between 21 and 28 weeks on enrolment Ethnicity determined on basis of skin colour: white 45%; black 14%; mixed 41% Mean age 28 years (SD 5.5) Mean BMI 26 (range 17–50)	Not RCT Cohort of consecutive women *Aim*: to evaluate FPG as a screening test for states of GDM in unselected Brazilian women 5010 (90%) women with completed data out of 5579 enrolled	Screening test used: FPG at 24–28 weeks Several threshold levels evaluated Done at same time as the diagnostic test: 2-h 75-g anhydrous GTT according to WHO criteria[24] between 24 and 28 weeks GDM diagnosed if 2-h PG ≥ 200 mg/dl IGT diagnosed if 2-h PG ≥ 140 mg/dl and < 200 mg/dl	0.3% (16/5010) diagnosed as GDM (WHO criteria: 2-h PG > 200 mg/dl); 7.3% (363/5010) diagnosed as IGT (WHO criteria)

Sensitivity and specificity
Sensitivity and specificity of FPG in diagnosing GDM (WHO criteria): results as reported
FPG ≥ 81 mg/dl (4.5 mmol/l): sensitivity 94%; specificity 51%; PPV 0.6%; NPV 100%; testing positive 49%
FPG ≥ 83 mg/dl (4.6 mmol/l): sensitivity 94%; specificity 58%; PPV 0.7%; NPV 100%; testing positive 42%
FPG ≥ 85 mg/dl (4.7 mmol/l): sensitivity 94%; specificity 66%; PPV 0.9%; NPV 100%; testing positive 35%
FPG ≥ 87 mg/dl (4.8 mmol/l): sensitivity 88%; specificity 72%; PPV 1.0%; NPV 100%; testing positive 28%
FPG ≥ 89 mg/dl (4.9 mmol/l): sensitivity 88%; specificity 78%; PPV 1.3%; NPV 100%; testing positive 22%
ROC curves aimed at jointly maximising sensitivity and specificity indicated threshold of FPG as 89 mg/dl for detection of GDM
Authors' comments: if ADA new criteria for general population were applied in this pregnant population, the performance of FPG would remain similar to that presented (no data supporting this comment); diagnosis of GDM with FPG still awaits validation against obstetric outcomes; basing diagnosis of GDM on FPG requires assurance of state of complete 8-h fast
Authors' conclusion: FPG is a useful screening test for GDM, a threshold of 89 mg/dl maximising sensitivity and specificity

Comments
Clearly written and presented
Unselected population representative of general population of Brazil
All women underwent diagnostic test for GDM
Assumes all cases of GDM were detected by GTT by 28 weeks. May have failed to detect cases developing later in pregnancy
Analysis based on small number of cases of GDM (16 cases out of 5010 women)
Results not related to any fetal/maternal outcomes
No costs reported

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Sacks *et al.*, 1992[145]	USA Unselected women aged over 25 years Only high-risk women included: FH, obesity, age > 25 years No demographic details given	No randomisation, no control group 4561 women Compared FPG with 1-h GCT as a screen for GDM	Tested with 1-h 50-g GCT at first prenatal visit (12 weeks); those tested negative on GTT retested at 29 weeks Those screened ≥ 135 mg/dl given 3-h 100-g GTT Diagnosis on two or more levels at: fasting 105; 1-h 190, 2-h 165; 3-h 145 mg/dl	141 (3.1%) GDM Of these 83 (59%) had FPG < 105 mg/dl on the GTT A significant correlation of fasting glucose level observed between those given GTT twice No significant correlation between first and second glucose screening results

Sensitivity and specificity
Sensitivity reported for FPG levels based upon a ROC curve:
At 88 mg/dl: sensitivity 80%, specificity 40%
At 84 mg/dl: sensitivity 90%, specificity 21%
At 81 mg/dl: sensitivity 95%, specificity 11%
No sensitivity/specificity data given for 50-g glucose screening test
Area under FPG curve significantly greater than under the glucose screening test curve ($p < 0.001$)

Comments
Sensitivities worked out on the assumption that 141 GDM patients total diabetics in sample but no testing after 29 weeks
Use of RF patients only
Authors conclude that FPG may be more discriminating and reproducible screening test for GDM

Studies using RPG

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Neilson *et al.*, 1991[165]	Oregon, USA Average maternal age 27.9 years. CS rate 26.7% (18.6% when excluding breech presentations and repeat CS) No ethnic background given	Study examining whether the results of GTTs were correlated with pregnancy outcome (including macrosomia, birth weight, birth trauma and CS) *n* = 608 Patients with incomplete data were not excluded	Used the 1-h 50-g GCT to non-fasting patients. Screen ≥ 150 mg/dl were given the 3-h diagnostic GTT (NDDG criteria[18]; O'Sullivan and Mahan criteria[19]) after overnight fasting	15% (88/608) of the population had a GCT screen ≥ 150 mg/dl and of these 83 had a GTT. There were 21 (3.5%) abnormal GTTs Of the 21, one was < 24 years, and eight were < 30 years Glucose tolerance parameters were not found to be indicative of macrosomia, birth weight or CS in this population Cost of screen given as US$17.75, GTT as US$59.15. Direct costs only? Cost per case of GDM diagnosed calculated at US$722.31 No adverse effects documented
Sensitivity and specificity Author cites the following sensitivities for age-related screening: 95% for women aged ≥ 24 years; 86% for women aged ≥ 28 years and 62% for women aged ≥ 30 years. However, these assume all women with GDM have been detected. In fact only those with a GCT ≥ 150 mg/dl were, and indeed five of those were not given a GTT				
Comments No data on what costs of screening and diagnostic tests includes 5/88 of the positive screens did not receive a GTT True prevalence uncertain because only those with ≥ 150 mg/dl screen were given the GTT				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
McElduff *et al.*, 1994[141]	New South Wales, Australia No data on characteristics of patients Authors note incidence of 3.9% in the study population attributable to over-representation of immigrants from Indian subcontinent, South-East Asia, Japan and Polynesia. Note figures not representative of Australian population in general	Comparison of RPG and 50-g GCT No randomisation, all patients given both tests *n* = 730, 14 of these were excluded because of incomplete or missing data	Screening was performed at or around 28 weeks and within 2 h of a meal Patients were given an RPG, rated positive if ≥ 6.1 mmol/l. Following this were given the 50 g glucose load and had VPG tested (GCT) after 1 h, rated positive ≥ 7.8 mmol/l Positives underwent diagnostic test: 3-h 100-g GTT (O'Sullivan and Mahan criteria[19]). Two abnormal values of 5.8, 10.6, 9.2 and 8.1 mmol/l at 0, 1, 2 and 3 h, respectively determined GDM	28/714 (3.9%) diagnosed with GDM after a positive screen No adverse effects documented
Sensitivity and specificity RPG: false-positives 13.4% (96/714); sensitivity 46.4% (13/28); PPV 11.9% (13/109) 50-g GCT: false-positives 8.8% (63/714); sensitivity 85.7% (24/28); PPV 27.6% (24/63) False-positives significantly different at 0.001, sensitivity significant at 0.02 *Authors' conclusion*: GCT is superior screening test				
Comments No cost data given Only positive screens given diagnostic test so incidence may be higher if some GDM undetected				

Studies using RPG *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Nasrat *et al.*, 1988[61]	Kuwaiti Arab women (31%); non-Kuwaiti Arab women (50%); 'others' (19%) No further demographic details given Kuwaiti women known high-risk group for GDM	276 unselected women No randomisation All women had RPG followed by GTT Tested predictability of random glucose test in women who had last meal within 2 h and those who had their last meal > 2 h	Screening: RPG Time last eaten noted Two thresholds: Lind and Anderson[214]: > 2-h 6.4; ≤ 2-h 7.0 mmol/l and their 90th percentiles Within 5 days given fasting 75-g GTT (250 women completed both) Bloods at fasting, 1 and 2 h Used WHO (1980[108]) thresholds for DM: fasting ≥ 8.0; 2-h ≥ 11.0 mmol/l and for IGT: 2-h 8–11.0 mmol/l Tested at 28–32 weeks	91% (250) women completed study 3 (1.2%) women GDM; 46 (18.4%) IGT No differences between the three groups of women 7 patients vomited 75 g glucose Screening useful within 2 h of eating but not useful if time since eating > 2 h
Sensitivity and specificity Reported using Lind and Anderson threshold, 7.0 mmol/l if eaten < 2 h, 6.4 if eaten > 2 h: sensitivity 16%; specificity 96%; PPV 47% No data available to check Using 90th percentile of study group: sensitivity 29%; specificity 89%; PPV 38% No data available to check				
Comments No cost data given All patients given GTT – 26 drop-outs High incidence in Kuwait reported – not demonstrated Authors conclude that random glucose of limited predictive value				

GCT and GTT studies

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Benjamin *et al.*, 1993[215]	USA High-risk population (RFs not stated) Patients described as largely being from socially and economically deprived areas Mean age 26.4 (SD 6.0)	Non-randomised, no controls *Aim*: to assess effects of period of gestation on predictability of 50-g glucose screening test 101 women	All patients screened in first trimester with fasting 50-g GCT. Cut-offs 140 mg/dl (Second International Workshop-Conference on GDM[216]) and 150 mg/dl (O'Sullivan and Mahan criteria[19]) Then in second and third trimesters with 3-h 100-g GTT Thresholds for 50-g GCT: O'Sullivan and Mahan 150 mg/dl at 1 h and also modified criteria 140 mg/dl[216] Diagnostic test: 3 days' unrestricted diet with at least 150 mg CHO Fasting 100-g GTT at second trimester (14–26 weeks) and third trimester (27 weeks to term). NDDG thresholds[18]: two or more ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl	Incidence 7.9% women with GDM Screening = ≥ 150 mg/dl 23.8% (24/101) ≥ 140 mg/dl 31.7% (32/101) No additional GDM patients detected using lower threshold No adverse effects documented
Sensitivity and specificity 7/8 patients positive screen, positive diagnostic test *At second trimester* Sensitivity 50-g GCT using 150 mg/dl cut-off = 25%; specificity 83% Sensitivity 50-g GCT using 140 mg/dl cut-off = 25%; specificity 74% *At third trimester* Sensitivity 50-g GCT using 150 mg/dl cut-off = 88%; specificity 82% Sensitivity 50-g GCT using 140 mg/dl cut-off = 88%; specificity 73% No differences between the two screening thresholds In positive screen patients 63% with GDM at third trimester will go undetected at second trimester				
Comments Also compared the response to the 100-g GTT of the non-diabetic women (n = 93) with 131 healthy women matched for age. Demonstrated significantly lower glucose tolerance (higher serum glucose) at 1, 2 and 3 h 'High risk' definition not given Assumes all cases of GDM by third trimester (27 weeks)				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Bergus and Murphy, 1992[168]	Alaska, USA 76 women, 10 did not complete questionnaire Two groups: monomer n = 41, polymer n = 35, baseline characteristics same Excluded if history of DM	Randomised double-blind study No controls Compared glucose polymer with glucose monomer (randomised to either) with sampling method: capillary or venous (all had both)	50-g GCT, regardless of time of last meal 1 h after capillary and venous blood taken Time of testing 24–28 weeks Those ≥ 7.8 mmol/l[216] considered positive All women underwent 3-h 100-g GTT 3 days later Fasting state not stated Questionnaire regarding side-effects and preference for blood sampling technique	No difference in mean BG between glucose groups at 1 h Symptoms (nausea, headache, dizziness, bloated, tired, abdominal discomfort or vomiting) fewer in polymer group (9/33), compared with 17/33 monomer (mean symptom scores significantly different) High correlation CBG and VBG (0.82) Four women ≥ 7.8 mmol/l at screen No cases GDM found No differences in patient preference between capillary or venous sampling
Sensitivity and specificity Reported that capillary blood test had sensitivity by ROC of 75% and specificity of 83% to detect women with a positive GCT (no data to check) No data for venous sampling				
Comments Small sample size No cases GDM identified				

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Cetin and Cetin, 1997[153]	Sivas, Turkey Included if examined < 20 weeks' gestation Exclusion criteria: pre-existing DM, multiple pregnancy, pre-term premature rupture of membranes, pre-eclampsia, delivery ≤ 28 weeks, regular ingestion of any drug Mean age 28 years; mean parity 1; mean BMI 24.5	Not RCT *Aim*: to determine use of different cut-off values with regard to the time of patient's last meal in screening for GDM 274 women analysed 85% (291/344) women were eligible, 274/291 completed the study. Reasons are given for drop-outs	Screening test used: 1-h 50-g GCT between 24 and 28 weeks. Non-fasting Screening cut-offs Women grouped according to timing of last meal before GCT 1. < 2 h. Standard cut-off 140 mg/dl vs suggested 148 mg/dl 2. 2–3 h. Standard cut-off 140 mg/dl vs suggested 142 mg/dl 3. > 3 h. Standard cut-off 140 mg/dl vs suggested 150 mg/dl All women underwent diagnostic test: 100-g GTT after 3 days of a 150 g CHO diet and 8–12-h fast. GDM diagnosed if two or more values ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl Those with one abnormal value were re-tested 1 week later	6.2% (17/274) were diagnosed as GDM (see criteria in adjacent column) No significant differences in the perinatal characteristics between groups Because of the small sample size, statistical comparisons and suggested cut-offs of the 3 groups and all patients could not be performed

Sensitivity and specificity
Values for screening in predicting GDM as reported by time since last meal:
1. < 2 h
Standard cut-off 140 mg/dl: sensitivity 75%; specificity 86%; PPV 27%
Suggested cut-off 148 mg/dl: sensitivity 63%; specificity 91%; PPV 33%
2. 2–3 h
Standard cut-off 140 mg/dl: sensitivity 60%; specificity 89%; PPV 30%
Suggested cut-off 142 mg/dl: sensitivity 60%; specificity 92%; PPV 30%
3. > 3 h
Standard cut-off 140 mg/dl: sensitivity 50%; specificity 89%; PPV 25%
Suggested cut-off 150 mg/dl: sensitivity 50%; specificity 92%; PPV 33%
All patients:
Standard cut-off 140 mg/dl: sensitivity 65%; specificity 88%; PPV 27%
Suggested cut-offs: sensitivity 59%; specificity 92%; PPV 32%
Authors' conclusion: These new values would lead to improved efficiency of the screening test and decreased frequency with which patients required a 3-h GTT

Comments
The basis used for selecting the 'suggested' cut-off values for the GCT was not stated
Those with one abnormal value were re-tested 1 week later. Results from this repeat testing were not reported
Assumes all GDM detected by positive GTT by approx. 30 weeks
Sample too small to detect significant rates in fetal/maternal outcomes using suggested criteria
No costs reported

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Court *et al.*, 1985[123]	Auckland, New Zealand Sample selection: women attending antenatal clinics who did not have overt DM Across groups in RCT: mean age 26 years; mean parity 0.8 and 1.0, weight gain 16 and 17 kg Groups were comparable at baseline	RCT Women randomised to either of 2 screening tests *Aim*: to compare screening for GDM using glucose and glucose polymer and to derive diagnostic criteria for the 100-g glucose polymer test by correlating maternal CHO metabolism with indexes of neonatal metabolic performance 100 women randomised into 2 groups: 1. Group G. 48 women received glucose screening test 2. Group P1. 52 women received glucose polymer screening test Glucose polymer tested on further 178 women: group P2 (giving 230 women on polymer) Neonatal metabolic index used was Kt = rate of glucose use in first hour of infant's life. Assessed in 23 neonates (12 from group G, 11 from group P1, 7 from group P2)	Screening test used: PG 1 h after: 1. 100 g glucose or 2. 100 g glucose polymer Tests taken at about 28 weeks without prior fasting Area under glucose curve measured Diagnostic test: 3-h 100-g GTT as soon as possible after screening test. Prior overnight fast and 3 days of 250 g CHO diet Diagnostic criterion developed from randomised patients then tested on the 230 women who received the polymer	7.8% (18/230) women had glucose polymer curve area > 50 units and were considered to have GDM 5.2% (12/230) diagnosed as GDM by NDDG criteria[18] GDM women (area > 50 units) compared to non-GDM women were significantly older (30.7 vs 25.8 years; $p < 0.001$); had significantly less weight gain during pregnancy (10.9 vs 15.4 kg; $p < 0.01$); had heavier infants (3712 vs 3321 g; $p < 0.001$) Classical RFs present in 67% of those with GDM Glucose polymer test curve correlated significantly with the neonatal Kt ($p < 0.0001$) but the GTT did not Neonatal metabolic results available for 23 infants (12 from group G and 11 from group P2: RCT). Maternal glucose polymer curve area > 50 units detected 4/5 infants with accelerated glucose use (Kt > 1.57%/min)
Sensitivity and specificity Results given below as reported Sensitivity and specificity of glucose polymer screening test using threshold of 8 mmol/l (144 mg/dl): Defining GDM as curve area > 50 units: sensitivity 89%; specificity 81%; PPV 29% Defining GDM as 1–3-h criteria (fasting 5.8; 1-h 10.0; 2-h 8.9; 3-h 7.8 mmol/l): sensitivity 88%; specificity 80%; PPV 25%. The 3-h criteria were elected arbitrarily by modification of NDDG criteria *Authors' conclusion*: The glucose polymer is preferable to glucose for CHO loading in pregnancy because of lower rates of nausea, better reproducibility of test results and better correlation between maternal and neonatal indexes of CHO intolerance				
Comments Women were randomised although method was not reported All patients underwent diagnostic GTT Assumes all cases of GDM diagnosed by GTT at approximately 29 weeks The criterion was developed using 50 women on the glucose polymer and then tested on 230 women (including the original 50 women). Thus results on this sample may be optimistic Diagnostic criteria for GDM (area under curve for polymer vs NDDG criteria not compared on fetal/maternal outcomes) Small number of neonates tested (23/100 from RCT and 7/178 from other group) Not sure how valid the use of chosen neonatal outcome (Kt) is as an indicator of morbidity. Index may be biochemically but not clinically significant Rates of nausea after ingestion of CHO loads and reproducibility of glucose polymer test not reported in this study but references to another study are given				

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Jirapinyo *et al.*, 1993[164]	Bangkok, Thailand Inclusion criteria: women with one or more traditional RFs (FH DM, previous infant ≥ 4000 g, unexplained peri-natal death or fetal anomalies, habitual abortion, repeated glycosuria, poly-hydramnios, obesity (BMI ≥ 27), previously abnormal GTT, age ≥ 35 years) Gestational age ranged from 8 to 38 weeks. 65% screened between 24 and 30 weeks Oriental ethnicity	Not RCT Prospective cohort *Aim*: to evaluate 2 different cut-off values in the 1-h GCT with respect to the traditional 3-h GTT in women at high risk for GDM 396 women involved	Screening test used: RFs plus 1-h 50-g GCT. Threshold VBG ≥ 140 mg/dl and ≥ 150 mg/dl No prior fasting. Most (65%) between 24 and 30 weeks All women underwent diagnostic test: 3-h 100-g GTT performed after 3 days of unrestricted diet (> 150 g CHO) followed by > 8-h overnight fast GDM diagnosed according to Second International Workshop-Conference on GDM criteria:[216] VBG ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl	10.6% (42/396) cases of GDM[216] GDM was more common in those with obesity and those aged ≥ 25 years Time since last meal before GCT was not associated with any marked difference in results for thresholds ≥ 140 mg/dl and ≥ 150 mg/dl
Sensitivity and specificity 1-h GCT ≥ 140 mg/dl: sensitivity 86% (36/42); specificity 65% (230/354); PPV 23% (36/160) 1-h GCT ≥ 150 mg/dl: sensitivity 83% (35/42); specificity 79% (279/354); PPV 32% (35/110) *Authors' conclusion*: The 1-h GCT with cut-off value of ≥ 150 mg/dl is a valuable and practical screening test for GDM				
Comments Oriental population with one or more classical RFs was studied All women underwent 3-h diagnostic GTT Reported sufficient data to permit calculation of validity Assumes all cases of GDM detected by GTT at whatever stage in pregnancy this was performed. May have misclassified any cases of GDM developing in later pregnancy and underestimated rates of GDM Does not relate diagnosis of GDM to fetal/maternal outcomes Costs not reported				

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lewis *et al.*, 1993[152]	Ontario, Canada Two non-hypertensive groups: 1. Control group: non-diabetics with negative 50-g GCT between 24 and 28 weeks without regard to fasting/fed state 2. GDM group: diagnosed on basis of abnormal 50-g GCT and 3-h GTT[150] None received insulin No significant differences between groups for age: 28 vs 30 years; gravidity: 3.5 vs 3.7; BMI at LMP 25.7 vs 26; BMI at study 29.9 vs 29.2 Obese 2/12 vs 3/10	Not RCT Controlled clinical trial *Aim*: to examine the effect of prior meal ingestion on the glucose, insulin and C-peptide response to a 50-g GCT in third trimester of pregnancy GDM group: 10 women Control group: 12 women	Tests performed at 26–32 weeks Screening test used after 12-h overnight fast: 1. 50-g GCT in fasting state 2. 50-g GCT 1 h after test meal 3. 50-g GCT 2 h after test meal Test meal: 500 kcal mixed meal with 50% CHO, 35% fat, 15% protein (full details given) Order of tests determined randomly All women subsequently underwent a 100-g GTT after 12-h overnight fast and 3 days of unrestricted diet (≥ 150 g CHO) GDM diagnosis based on 2 or more values exceeding: fasting 5.8; 1-h 10.6; 2-h 9.2; 3-h 8.1 mmol/l	Rates of GDM in population not reported Control: 58% (7/12) had a false-positive GCT (1-h glucose ≥ 7.8 mmol/l) when test performed in fasting state. 25% (3/12) had false-positive test in 1-h and 2-h postprandial tests. Differences failed to reach statistical significance GDM: no false-negative 1-h glucose results were noted. 1/10 had false-negative results in the 1-h post-prandial study and 1/10 (different woman) had a false-negative in the 2-h postprandial study
Sensitivity and specificity N/A Also reports results of monitoring of serum insulin and plasma C-peptide *Authors' conclusions*: The effect of the prandial state on the PG response to the 50-g GCT used to screen for GDM may be of sufficient magnitude to significantly alter the operating characteristics of this test				
Comments Very small sample size will have limited the power of the study to detect statistically significant differences This is acknowledged by the authors, who suggest a larger study is required Baseline characteristics of groups compared although no details of ethnicity Diagnostic GTT was repeated to confirm/refute diagnosis of GDM No investigation as to whether values related to prandial state were associated with fetal/maternal outcomes Authors' conclusions were supported by the evidence				

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Mathai *et al.*, 1994[184]	India No data on ethnicity of women Mean age in 2 groups 23.9 (SD 3.95) and 15.8 (SD 4.8) Women with previous DM or GDM excluded	Observational study, no randomisation *Aim*: to evaluate sensitivity and specificity of two screening tests across BG levels Consecutive women grouped as high risk for GDM (n = 111) or low risk (n = 121) RFs for high risk: age > 35, FH DM, previous fetal wastage, previous high birth weight baby (> 3800 g), obesity (> 120% ideal body weight), glycosuria, and a symphysio-fundal height measurement > mean ± 2 SD local population	All women had RBG and a 1-h 50-g GCT test for screening at 26–30 weeks Blood taken without regard to time of last meal All women received a 100-g GTT within 2 weeks. GDM diagnosed on C&C criteria,[20] any two: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl	7 (6.3%) high-risk and 4 (3.3%) low-risk women had GDM Difference in mean birth weight not statistically significant
Sensitivity and specificity Using ROC curves various sensitivity and specificity calculated Best results: At a threshold of 90 mg/dl for the RPG the sensitivity was 63% and specificity 66% At a threshold of 115 mg/dl on the GCT sensitivity was 63% and specificity 55% No data available to check calculation				
Comments Study tested all women with GTT Neither random nor GCT screening were able to provide high sensitivity and specificity The population used had a higher incidence of GDM than average and also found a high percentage of IGT, likely to be a result of ethnic group used				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Murphy *et al.*, 1994[188]	USA Mean age 25.9 years No SD or range given No ethnic data available	Randomised trial, 3 groups, no control Comparison CHO sources for screening test 124 women randomly assigned to 1 of 3 sources (not stated how randomised) Tested at 24–28 weeks	Non-fasting screening test: Group 1: 50 g glucose polymer Group 2: standard 50 g glucose solution Group 3: milk chocolate bar 50 g Blood test at 1 h Diagnostic test: 3-h 100-g GTT within 1/52 2 or more O'Sullivan and Mahan[19] (NDDG[18]) criteria for GDM Fasting status not defined	4.6% GDM (5/108): 3/5 standard glucose of which 2 had screening value < 7.5 mmol/l 2/5 glucose polymer Screening test positive in 15.3% (19/124) (glucose ≥ 7.5 mmol/l) Polymer 11% (5) standard glucose 26% (11), chocolate 7.6% (3) 16 patients unable to complete GTT, 5 vomiting, 11 data incomplete No significant differences in mean glucose levels between polymer or standard glucose groups. Chocolate bar significantly lower mean glucose levels than standard glucose Side-effects fewer in polymer test
Sensitivity and specificity (Some numbers of the patients do not appear to add up) Overall: sensitivity 60%; specificity 84%; PPV 16% Standard glucose: sensitivity 33.3%; specificity 73.6%; PPV 9% Polymer: sensitivity 100%; specificity 92.8%; PPV 49% Cannot calculate chocolate				
Comments All patients received GTT Some costs reported: Polymer 71 cents per 50 g dose dry or US$1.06 liquid plus carbonated water US$1.03 per dose Chocolate bars 47 cents per dose No costs for standard glucose				

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
O'Sullivan *et al.*, 1973[100]	Boston, USA Two groups of patients: 1. Positive on initial screening or who met clinical history criteria (previous baby weighing ≥ 9 lbs; history of two or more of fetal death, neonatal death, congenital anomaly, prematurity, excessive weight gain, hypertension, or proteinuria; and FH of DM (1954–59) 2. Prenatal patients (1956–57) ?Overlap between groups No further details in this report – refers to previous report	Cohort Group 1: 18,812 women screened 44% were positive on screening using GCT or history Group 2: 986 women screened with GCT 752 (76%) responded to offer of GTT	Screening test: GCT 1-h venous blood sugar tested after oral 50 g glucose Group 1: only those with blood sugar ≥ 130 mg/100 ml or positive clinical history underwent GTT Group 2 all offered GTT Diagnostic test: 3-h OGTT (quantity of glucose not stated) Positive if 2 or more venous whole blood samples exceeded (mg/100 ml): fasting 90; 1-h 165; 2-h 145; 3-h 125 Not stated whether tested in fasting/non-fasting state	2.5% (19/752) classified as GDM after GTT (O'Sullivan and Mahan criteria[19]) at unknown gestation 44% of Group 1 had one or more RFs No adverse effects documented 1-h screening tests for those who refused GTT reported as no different from those who chose GTT

Sensitivity and specificity

Predictive power of various screening factors in 752 unselected women:
GCT (1-h ≥ 130 mg/100 ml): sensitivity 79%; specificity 87%; PPV 14.8% (15/109)
Clinical history criteria: sensitivity 63%; specificity 56%; 44% (332) required GTT. Both: sensitivity 53%; specificity 93%; 8% (62) required GTT
Age ≥ 25 years (361 women representing 48% of population):
Screening GCT (1-h ≥ 130 mg/100 ml): sensitivity 88% (14/16); specificity 82% (284/345); PPV 19% (14/75); 21% (75/361) required GTT
Clinical history criteria: sensitivity 69%; specificity 35%; 65% (235/362) required GTT
Predictive value of screening (≥ 130 mmol/l) in those with positive clinical history: sensitivity 63% (10/16); specificity reported as 87% (301/345); PPV 19% (10/54). 54 required GTT
Authors' recommendations: screening with 1-h blood sugar level after 50 g of glucose and a GTT if blood sugar ≥ 130 mg/100 ml. Patients aged 25 or over are at greater risk – same screening recommended

Comments

No characteristics of population reported in this study (although reference given)
Timing of testing in relation to stage of pregnancy not stated
No quantity of glucose stated for GTT
76% of Group 2 underwent GTT
Results of immediate GTT only reported – does not appear to have been followed up until end of pregnancy. Timing of diagnostic testing not reported so may be an underestimate of frequency of GDM
No costs
Sample collected between 1956 and 1957 (? relevance to general population in 2000)

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Perucchini *et al.*, 1999[104]	Zurich, Switzerland All pregnant women with singleton pregnancy delivering after 28 weeks eligible Ethnicity: white (European) 63%; Asian 19%; African 6%; others 12% Mean age 28.4 years; age range 17–45 years Mean BMI: 23.8 (SD 0.2) Excluded: pre-existing DM; lack of examination before 24 weeks	Not RCT Prospective, population-based observational study *Aim*: to evaluate whether measuring fasting blood sugar is an easier screening procedure for GDM than the 1-h 50-g GCT 772 eligible; 558 (72.3%) consented 38 (6.8%) subsequently excluded (vomiting, missing data, protocol violation) Those excluded had similar demographic characteristics to those included 520 women remaining in study	Screening test used: standard 1-h 50-g GCT irrespective of last food intake between 24 and 28 weeks of pregnancy Diagnostic test and thresholds: 3-h 100-g OGTT within 1 week of screening. Prior fasting overnight for 12 h and 3 days of 150–200 g (minimum) CHO diet GDM diagnosed according to C&C criteria:[20] two or more values exceeding: fasting 5.3; 1-h 10.0; 2-h 8.6; 3-h 7.8 mmol/l Diagnostic criteria above (75-g GTT without 3-h value or for complete 100-g GTT) adopted by Fourth International Workshop-Conference on GDM[21]	10.2% (53/520) diagnosed as GDM (C&C criteria) GDM by ethnicity: slightly more prevalent in Asian women with 16% (16/99) GDM ($p < 0.05$) or African women with 13% (4/31) GDM ($p = 0.39$) compared to white women with 8% (26/328) GDM GDM by age: significantly more common in those > 30 years (77% vs 46%; $p < 0.001$) GDM by BMI: significantly more common in those with BMI > 25.0 (54% vs 32%; $p = 0.001$) 6.8% (38/558) vomited after 100 g glucose load

Sensitivity and specificity
Results as reported. Raw data not presented
Sensitivity and specificity of fasting glucose:
ROC analysis showed best cut-off value for FPG was 4.8 mmol/l
Universal screening using threshold FPG of 4.8 mmol/l: sensitivity 81%; specificity 76%. 30% (155/520) women ≥ 4.8 mmol/l on screening
Universal screening using threshold FPG of 4.4 mmol/l: sensitivity 100%; specificity 39%. 55% (285/520) women ≥ 4.4 mmol/l on screening
Sensitivity and specificity of 50-g GCT:
ROC analysis showed best cut-off value for 1 h post 50-g GCT was 7.0 mmol/l
Universal screening using threshold for 1 h post 50-g GCT of 7.8 mmol/l: sensitivity 59%; specificity 91%. 14% women ≥ 7.8 mmol/l on screening (from results section of abstract)
Universal screening using threshold for 1 h post 50-g GCT of 7.5 mmol/l: sensitivity 61%; specificity 88%
Universal screening using threshold for 1 h post 50-g GCT of 7.0 mmol/l: sensitivity 68%; specificity 82%
Improved sensitivity by fasting for 2 h prior to test and using cut-off value of 7.0 mmol/l: sensitivity 100%; specificity 71%
Diagnostic 100-g OGTT: 7.6% (4/53) women with GDM would have been missed by omitting the 3-h value. Sensitivity 92.4%

Comments
Sample representative of general population
Characteristics of population given
Influence of time of last meal examined
Assumes that all GDM detected by GTT by 29 weeks. May have underestimated GDM rates because of women developing GDM after 29 weeks not being detected
Costs of different procedures not considered
Authors' conclusions: Measuring FPG using a cut-off value ≥ 4.8 mmol/l between 24 and 28 weeks of pregnancy is an easier screening procedure for GDM than the 50-g GCT and allows 70% of women to avoid the GCT. The 3-h value can be omitted from the GTT with relatively little loss of sensitivity

GCT and GTT studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Schwartz *et al.*, 1994[97]	Texas, USA 81% of the 190 women giving consent were Mexican-Indians (incidence of GDM reported as 10.6% in locality), largely from poor socio-economic background Mean age 27 years (range 13–42 years) Gestation: mean 27 weeks (range 9–40 weeks)	Not RCT *Aim*: to compare deviations in glucose values, insulin values and side-effects between standard 50-g glucose solution and modified solution 132 women completed tests (28 had traditional RFs of previous GDM/GDM in family, 104 tested positive on GCT (defined as ≥ 7.8 mmol/l)	Screening tests: 1. 1-h screening with standard glucose solution (50 g glucose in 150 ml water) 2. 1-h screening using modified glucose solution (50 g glucose in 450 ml water) Pre-test fasting after midnight on eve of each test. At least 3 days between tests. Tests of modified vs standard glucose solution in random order 1-h GCT threshold > 7.8 mmol/l 3-h GTT using standard 100 g glucose (100 g in 300 ml water) Diagnostic GDM test: 2 or more elevated values on 3-h GTT (NDDG criteria[18])	19% (25/132) diagnosed as GDM by NDDG criteria in population of whom 79% had screened positive on GCT Number with nausea/vomiting was significantly less after modified compared to standard solution 11% (15/132) vs 3% (4/132). $p < 0.05$ 90% of patients preferred the taste of the modified solution Side-effects: Nausea/vomiting were significantly more common with the standard solution compared to the modified: 11% (15/132) vs 3% (4/132): $p < 0.05$ Side-effects including sweating and dizziness were more common with the standard solution compared to the modified: 14% (18/132) vs 3% (4/132) Mean patient rankings on taste favoured the modified solution over the standard (1 = good, 5 = bad): 1.4 (0.1) vs 2.9 (1.1): $p < 0.05$. None of the patients complained about the larger volume of the modified solution
Sensitivity and specificity For abnormal GTT using different cut-off values for both types of GCT (1-h values): Modified (50 g glucose in 450 ml water): ≥ 7.22 mmol/l: sensitivity 100%; specificity 36.4% ≥ 7.78 mmol/l: sensitivity 96%; specificity 52.3% ≥ 8.33 mmol/l: sensitivity 96%; specificity 61.5% ≥ 8.89 mmol/l: sensitivity 96%; specificity 78.5% ≥ 9.44 mmol/l: sensitivity 84%; specificity 86.9% Standard (50 g glucose in 150 ml water): ≥ 7.22 mmol/l: sensitivity 92%; specificity 43% (vs modified specificity $p = 0.02$) ≥ 7.78 mmol/l: sensitivity 92%; specificity 52.3% ≥ 8.33 mmol/l: sensitivity 84%; specificity 66.4% ≥ 8.89 mmol/l: sensitivity 80%; specificity 78.5% ≥ 9.44 mmol/l: sensitivity 76%; specificity 85% Non-significant where p not stated Using cut-off ≥ 7.8 mmol/l after GCT: two patients with positive GTT would have been missed by using only the standard solution and one missed by using only the modified solution				
Comments Wide range of gestational age and maternal age. No investigation of influence of gestation/age on outcomes High-risk selected population (79% had screened positive on 50-g GCT with ≥ 7.8 mmol/l) All participants had OGTT No follow-up until end of pregnancy. May be underestimate of GDM No costs *Authors' conclusion*: Modified solution had fewer side-effects than standard solution and was at least as sensitive as the standard solution in screening for GDM				

Fructosamine studies

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Bor *et al.*, 1999[174]	Turkey 54 non-diabetic and non-pregnant women and 96 pregnant women who had positive (over 140) GCTs 24–28 weeks' gestation	Comparison of fructosamine/albumin ratio and OGTT	100-g OGTT – GDM when 2 or more of the following surpassed: fasting 105 (5.9); 1-h 190 (10.6); 2-h 165 (9.2); 3-h 145 mg/dl (8.1 mmol/l) Fructosamine upper limit < 2.8 mmol/l	12 of the 96 women diagnosed GDM after OGTT; only one had fructosamine above normal non-diabetic range No difference in fructosamine/albumin ratio between GDM and non-GDM pregnant groups
This study was done to see if adjustment for the lower protein levels seen in pregnancy would make fructosamine a useful screening test; it did not Also of interest – only 13% of those who had positive GCTs were OGTT positive GDM women were older than non-GDM pregnant women (mean ages 32 and 27; parities 0.8 and 2.1; BMIs 30 and 26)				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Begley *et al.*, 1992[217]	Poole, UK 36 pregnant women having OGTT	Cohort	Fructosamine; reference range 1.5–2.4 mmol/l No correction for protein concentration or gestational age 75-g OGTT and WHO criteria[24]	9 of 36 had abnormal OGTT; all but one had fructosamine in normal range, and one was below normal The pregnant group as a whole had fructosamines in normal range – no difference from non-pregnant women

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Comtois *et al.*, 1989[197]	Montreal, Quebec 100 consecutive women selected on grounds of high risk of GDM, but RFs not given Mean 28 weeks' gestation	Cohort	3-h 100-g OGTT; O'Sullivan and Mahan criteria[19] Fructosamine; reference range 1.9–2.5 mmol/l (reference population not given)	13 had GDM by OGTT, but no difference in fructosamine levels There was a difference in FPG – 4.6 vs 6 mmol/l
Conversion to a fructosamine/protein ratio gave a statistically significant but still not useful result The women had the OGTT as outpatients, but were then admitted and put on supervised diets, with BG levels tested 4 times a day No outcome data given				

Fructosamine studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Hughes *et al.*, 1995[173]	Multi-ethnic group in United Arab Emirates – about 67% Arab and 23% Indo-Pak Two groups studied: 1. Unselected women who had GCT 2. Women selected on basis of clinical risk	Two cohorts	50-g GCT; cut-off 7.8 mmol/l Clinical RFs (previous GDM, FH of DM, past macrosomia, shoulder dystocia, intra-uterine death, gross obesity Fructosamine, reference range not given Corrected fructosamine = measured x 72/total protein; reference range 203–8 for non-diabetics 3-h 100-g OGTT with C&C criteria[20] for normal: fasting < 5.3; 1-h < 10; 2-h < 8.6; 3-h < 7.8 mmol/l	Corrected fructosamine gave 79% sensitivity and 77% specificity, compared to OGTT Fructosamine alone not as good Accuracy, defined as sensitivity + specificity/2, was best for FPG (76%, at cut-off 4.7) and corrected fructosamine (78%, at cut-off 215)
The two groups in this study are not comparable. The first 161 of 682 women had GCTs, but this proved impractical in clinic and the other 521 were selected by RFs Ethnic groups with high DM prevalence, high obesity, high multiparity BG levels after 50-g GCT were higher than in reports from Caucasian populations – at 7.8 mmol/l cut-off, specificity 32%, sensitivity 93%. FPG did better – at cut-off 4.4 mmol/l, sensitivity 67%, specificity 83% Numbers not given in tables May be relevant to ethnic groups in UK. Do thresholds need to vary by ethnicity? No data on outcomes by levels of any test				
Notes The diagnosis of GDM led to an OR of 5.1 for neonatal admission to SCBU, but only 1.34 for birth weight over 4 kg. Diagnosis may influence practice more than required?				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Huter *et al.*, 1992[218]	190 women with singleton pregnancies; authors state that they were unselected 24–28 weeks' gestation; seven over 30 years of age but mean age not given	Comparison of OGTT and fructosamine	100-g OGTT using O'Sullivan and Mahan criteria[19] – GDM diagnosed when 2 or more values exceeded: fasting 90 (5), 165 (9.2), 145 (8.1) and 125 mg/dl (7 mmol/l) Fructosamine upper limit of normal 2.76 mmol/l	10 of the 190 women were diagnosed as having GDM by OGTT; none by fructosamine
There was no adjustment for protein levels, but the range for fructosamine in the GDM group overlapped with that in those with normal OGTT				

Fructosamine studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Menon *et al.*, 1991[162]	India Mean age approx. 27 years No population data described 234 women referred for GTT on basis of RFs	Use of fructosamine as a screening test for GDM in unselected women No randomisation or controls	FBG and fructosamine Threshold used not defined Tested at mean of 30 weeks' gestation All patients had 3-h GTT, glucose load thresholds used not defined Blood for fructosamine taken at fasting sample for GTT Categorised into: 1. Normal (173 women) 2. GDM (28 women) 3. Only FBG normal (15 women) 4. One value other than FBG abnormal (18 women)	Incidence GDM 11.9% FBG significantly higher in GDM than non-GDM patients No differences in fructosamine levels between GDM and non-GDM patients
Sensitivity and specificity N/A				
Comments No cost data available Concern as to quality of report – RFs not defined, thresholds not defined				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Nasrat *et al.*, 1988[61]	Saudi Arabia No ethnic data given	Two parts: 1. Identification of a group of normal pregnant women by exclusion of those with RFs and GCT negative 2. Comparison of fructosamine and OGTT	75-g OGTT – positive if fasting ≥ 5.8 mmol/l and/or 2-h ≥ 10 mmol/l Patients with 2-h > 6.7 mmol/l were classed as IGT Fructosamine – normal range not given in figures but from graph about 2.7 to 0.9 at 24–28 weeks	GDM by OGTT in 6 of 98; 35 had IGT Fructosamine over 90th percentile for gestational age in only 3 of the 6; sensitivity 50%, specificity 90%; for IGT 9% and 94% respectively
Comments Conclusion – fructosamine is not useful as a screening test for GDM Birth weights in GDM group all under 3900 g				

Fructosamine studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Uncu *et al.*, 1995[182]	42 pregnant women in Turkey No data given on selection, but about one-third had GDM by OGTT, so some selection must have operated	Small selected group	100-g OGTT 50-g GCT, 140 mg/dl (7.8 mmol/l) cut-off HbA_{1C} 7.2% cut Fructosamine 2.85 mmol/l cut	Not much difference in test performance amongst fructosamine, GCT and HbA_{1C} in this small study

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Vermes *et al.*, 1989[171]	The Netherlands 106 pregnant women, second trimester (no weeks given); mean age 29, range 16–41 No data on how selected	Cohort	Fructosamine; reference range 1.9–2.6 in non-pregnant healthy subjects OGTT (presumably 75-g) and WHO criteria[24]	12 GDM by OGTT and WHO No difference between fructosamine levels in GDM vs normals Also no difference between FPG (4.5 vs 4.6 mmol/l) or HbA_{1C} (4.7 vs 4.8%)
Comments Correction of fructosamine to allow for the normal drop in protein levels in pregnancy made no difference				

All other studies

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Al-Shawaf *et al.*, 1988[219]	Saudi Arabia No description of women's demographic details given Aged 25.2 (SD 5.1) in control group, 29.4 (SD 5.6) in IGT group and 30.8 (SD 5.3) in GDM group Normal incidence GDM 8% Exclusion and inclusion criteria not stated	Cohort study, with 218 women found to have GDM or IGT and 884 consecutive women with normal screening PG acting as a control Selection of women in GDM, IGT group not defined *Aim*: to assess if IGT has an impact on maternal outcome and if this is related to the degree of impairment	Screening test 75-g GCT, positive if ≥ 7.8 mmol/l at 2 h (fasting status not defined) Diagnosis by fasting 75-g GTT following WHO guidelines[24] Women with fasting glucose < 7.8 mmol/l and 2-h between 7.8 and 11.0 mmol/l diagnosed as IGT Women with fasting glucose > 7.8 mmol/l and 2-h ≥ 11.0 mmol/l diagnosed as GDM Diet controlled unless glucose at 2 h after meal was > 7 mmol/l then insulin prescribed All women repeated GTT 6 weeks post-partum (only 55 (25.5%) did) Newborn infants examined for birth weight, congenital abnormalities and APGAR scores	41 women had GDM, 177 women had IGT, 884 controls Incidence rate cannot be assessed as data not available (control group are a separate group) Significant difference between age of IGT (29.4) and controls (25.2) ($p < 0.001$) but not significant between GDM and IGT Parity significantly different between IGT and controls and between GDM and IGT FH of DM significantly higher in GDM and IGT than controls ($p < 0.001$) No difference between IGT and GDM No difference in gestational age Significantly greater weight (as measured by BMI) between IGT and controls ($p < 0.001$) but not GDM and IGT Mean birth weight significantly greater in IGT than controls ($p < 0.01$) No significant differences between groups in terms of high blood pressure or mode of delivery No difference in APGAR scores Rate of major congenital malformation in GDM and IGT combined = 22.9/1000 and in controls = 11.3/1000 births. No difference in perinatal mortality
Sensitivity and specificity N/A				
Comments At follow-up, persistent abnormal glucose tolerance was found in 10/15 women with GDM and 12/40 with IGT No cost data Authors conclude that IGT does constitute a risk to pregnancy but less than that associated with GDM Authors note that primary care is limited in Saudi Arabia No account made for difference in parity in which GDM and IGT women more likely to have had 5 previous pregnancies				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Ammari and Gregory, 1996[220]	Leicester General Hospital, UK Hospital has 4000 deliveries per year, 35% of which are to women of Indo-Asian origin Quotes reported prevalence of GDM in this population as 11.9% by WHO criteria[24] (Dornhorst *et al*., 1992[115])	Not RCT Retrospective review of case notes looking for evidence of ordering of GTT and results from any maternal BG estimations *Aim*: to investigate reasons for low rates of referral to combined diabetic/antenatal clinic by auditing adherence to the locally agreed policy of screening for GDM using clinical RFs and to evaluate the effect of circulating findings from the audit on screening First audit in 1993: 143 case records (100 Europid and 43 Indo-Asian) Presentation and circulation of results from first audit to clinicians Repeat audit in 1995: 147 case records (95 Europid and 52 Indo-Asian)	Screening test used: clinical RFs (maternal obesity with BMI > 30; first-degree relative with DM; previous GDM; previous large baby > 4000 g at term; previous shoulder dystocia or cephalo-pelvic disproportion; previous unexplained childbirth; subsequent development of polyhydramnios, large for dates fetus, or glycosuria) Official policy: screen with standard GTT those women with above RFs at 24–28 weeks	Compliance with local guidelines was poor Education directed at clinicians was not effective in improving the detection of GDM in high-risk population GDM rates (WHO criteria[24]): 2% (3/147) in first audit; 0% in second audit First audit: 37% (53/143) women eligible for GTT according to local policy. Only 19% (10/53) of those eligible had GTT. Tests performed at median of 31.6 weeks (range 25–38 weeks). 3 Indo-Asian women had GDM (WHO criteria[24]), only one of whom was referred to combined clinic. In addition, 4 women with RFs had fasting or RBG measures performed instead of GTT Second audit: 48% (70/147) women were eligible for GTT of whom 29% (20/70) had GTT Proportion of women tested was similar to first audit (p = 0.30). Tests were performed at a median of 31 weeks (range 28–39 weeks). No GDM detected. No woman with RFs was tested with RBG or FBG

Sensitivity and specificity
N/A
Authors' comments: Arguable whether screening by clinical RFs is the best method for detecting GDM, especially in Indo-Asian women. Definition of macrosomia (> 4000 g) is inappropriate in this population. Maternal obesity is less predictive of GDM in Indo-Asian women than in Europid women (Dornhorst *et al*., 1992[115])
GTTs were generally performed late in pregnancy
GTTs were delayed after one episode of glycosuria until a second was detected
Authors suggest two possible approaches to improving detection and referral of GDM:
1. Improve compliance with local policy by empowering midwives to order GTT at 24 weeks for women with RFs
2. Pilot evaluation of comprehensive screening in all Indo-Asian women (authors favour this approach)

Comments
Highlights lack of compliance with local guidelines in 1996 even after results from first audit on adherence to screening policy had been circulated and presented to clinicians

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Ardawi *et al.*, 2000[221]	Saudi Arabia Women of diverse socio-economic status Mean age reported as two groups, 29.2 (SD 4.6) and 30.7 (SD 4.8) years Known RFs were recorded: FH, previous GDM, 'bad' obstetric history and history of glycosuria Women excluded with hepatic, renal problems or evident DM, previous GDM or other endocrine disorders	Prospective study, women described as being randomly recruited but does not state details of process 1056 women recruited, 818 completed the study (105 delivered in other hospital, 97 did not have GTT and 36 had incomplete data)	All women tested between 24 and 18 weeks with 1-h 50-g GCT. Those ≥ 7.2 mmol/l (≥ 130 mg/dl) (O'Sullivan and Mahan[19]) given 3-h fasting 100-g GTT after 3 days of 200 g CHO per day Diagnosis of GDM by NDDG criteria[18]: any 2 of fasting 5.8; 1-h 10.6; 2-h 9.2; 3-h 8.1 mmol/l Mode of delivery recorded, newborns observed for APGAR score, birth weight, head circumference, fetal length and birth weight centiles, age, and complications of birth	102 GDM cases (12.5% incidence) GDM women significantly older, heavier, higher gravidity, heavier birth weight and greater CS and stillbirths than non-GDM group (all p's < 0.05) The GTT-positive, GDM-negative women were also significantly older than negative screenees Screening PG showed marked correlation with 1- and 2-h GTT values ($p < 0.05$)
Sensitivity and specificity Report maximum sensitivity at 7.8 mmol/l as 88% and specificity as 84% with PPV 82% At 7.2 mmol/l reported sensitivity at 93% but specificity at 64% However, as women not tested if screened lower than 7.2% this is largely redundant Reported PPV at > 10.4 mmol/l as 100% and suggest that this can be used as a cut-off to ensure diagnosis of GDM				
Comments Authors conclude that a cut-off value of 7.8 mmol/l is a reliable screening value, however 22 cases of GDM would be missed at this level				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Artal *et al.*, 1984[222]	USA No description of demographic details No ages of sample given 82 women	Prospective cohort study Women undergoing 3-h GTT for maternal history of DM, age > 35 years, previous unexplained fetal death, congenital fetal abnormality, maternal obesity, previous abortion and hydramnios were randomly selected (no details given)	3-h 100-g fasting GTT following 3-day 150 g CHO diet Blood plasma tested at fasting, 1, 2 and 3 h after ingestion, as was serum for HbA_{1C} GDM diagnosed if 2 or more ≥: fasting 110, 1-h 200; 2-h 150; 3-h 130 mg/dl Correlated with HbA_{1C} rates Tested during third trimester	30 patients had GDM (high incidence 36% as selected RF women) Significant correlation with HbA_{1C} and GTT ($p < 0.006$)
Sensitivity and specificity At HbA_{1C} < 7%, 17 patients had negative GTTs and 8 positive (false-positive rate of 41%, false-negative 26%) At HbA_{1C} > 7%, 33 had negative glucose tests and 22 positive Overall, the agreement for the two tests at best will only be 69% of the patients				
Comments Findings confirm previous published reports that HbA_{1C} is not sufficiently sensitive and specific for screening procedures				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Bartha *et al.*, 2000[223]	Spain No demographic details given Two groups of women aged 33.6 (SD 5.4) and 32.6 (5.3) years	Compared women diagnosed with GDM in early pregnancy with those in late pregnancy on pregnancy and neonatal outcomes No randomisation, no control 3986 women screened 235 women with GDM, some numbers lost at follow-up	Screened at first antenatal visit (mean 18.1 weeks) with 50-g GCT (fasting status not defined) If glucose ≥ 140 mg/dl given 100-g fasting GTT Thresholds if 2 or more >: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl If negative screen or positive screen followed by negative GTT given repeat test at 24–28 weeks Defined two groups: those diagnosed early pregnancy and those late pregnancy All women diagnosed GDM admitted for 1 day and tested before and after meals and defined as A1 diet control, A2 insulin depending on glucose levels	Overall incidence 5.9% (235 women) 65 women early and 170 late Analysis on 183 women who delivered at the same hospital Differences noted between two groups: Pre-gestational BMI, hypertension and pre-eclampsia and need for insulin control all significantly greater in early group. However, no differences in age, parity, rate of previous abortion or previous CS delivery. Women in early group gained significantly less weight during pregnancy than late group Higher rate of neonatal hypoglycaemia and antenatal deaths in early onset group. However, mothers of two stillbirths had other medical complications
Sensitivity and specificity N/A				
Comments Authors comment that the increased fetal and neonatal risks are proportional to the degree and duration of maternal hyperglycaemia. Also note that testing at early pregnancy may find previously unrecognised type 2 DM No comparison made between those not diagnosed in terms of outcomes, etc. No cost data available				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Baxi *et al.*, 1984[224]	USA No demographic details given Age of women not reported	Cohort study with control group Selected women with high risk of GDM ('history of or demonstration of prediabetes') Those positive GDM compared to those negative (control) 180 women (33 diabetics, 147 controls, 10 excluded)	Screened with: HbA_{1C}, FPG and PG 2 h after a 100 g glucose load tested Fasting plasma ≥ 130 mg/dl or a 2-h post-load ≥ 140 mg/dl given a 3-h 100-g GTT Time of test not defined HbA_{1C} taken as positive if 1 SD above control group mean (6.78%)	Mean HbA_{1C} in control group 6.17 (SD 0.61), and in diabetics 6.87 (SD 0.73) 21 of the 33 diabetic patients had an elevated HbA_{1C} at screening 27 of the control patients had an elevated screening HbA_{1C} Mean HbA_{1C} values significantly different between controls and diabetics 10 of the diabetic patients gave birth to LGA babies. HbA_{1C} was elevated in all of these patients
Sensitivity and specificity Sensitivity of HbA_{1C} 63.6%, specificity 81.6%, false-negative rate 36.4% and false-positive rate 18.4%				
Comments *Authors' conclusion*: HbA_{1C} is less sensitive than the conventional GCT. However, an elevated HbA_{1C} value in the third trimester may predict the possibility of a macrosomic infant				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Berkowitz *et al*., 1992[99]	New York, USA Inclusion criteria: women without overt DM who delivered singleton infant between 1986 and 1991 who were screened both < 24 and > 24 weeks Exclusion criteria: women who had not been screened; women who had been screened but not retested later in pregnancy; those who had only been screened after 24 weeks; and those who had an abnormal screen but a normal GTT	Not RCT Prospective cohort *Aim*: to evaluate the yield of early routine screening for GDM and to determine whether maternal characteristics and neonatal outcome differ according to time of diagnosis 2776 women included Those eligible represented 36% (2776/7762) of the total population	Screening test used: 1-h 50-g glucola GCT at first visit (< 24 weeks), non-fasting state Abnormal if PG ≥ 135 mg/dl. If normal, re-screened ≥ 24 weeks Diagnostic tests: 3-h 100-g GTT. GDM diagnosed if 2 or more values ≥: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl (O'Sullivan and Mahan[19] as modified by C&C[20]) Those with abnormal screen but normal GTT were re-tested with second GTT ≥ 24 weeks Patients divided into 3 groups and compared: Early diagnosis group (102 women) Late diagnosis group (252 women) Controls with normal screen < and > 24 weeks (2422 women) GDM treated with ADA diet and home monitoring Insulin was started if FBG > 95 mg/dl or postprandial BG > 120 mg/dl on 2 or more occasions during a 2-week interval	GDM rates in local population not reported 354 diagnosed as GDM 28.5% (102/354) diagnosed < 24 weeks and 71.2% (252/354) ≥ 24 weeks After adjusting for BMI ≥ 27.3, PH GDM, FH DM and PH stillbirth, the following were significantly more common in early compared to late diagnosis: insulin use, maternal age ≥ 30, low weekly weight gain, and PET/chronic hypertension Early vs late diagnosis: insulin use (56% vs 30%), maternal age ≥ 30 (64% vs 45%), low weekly weight gain (46% vs 31%), and PET (15% vs 8%); chronic hypertension (12% vs 4%) See later for perinatal outcomes

Sensitivity and specificity
N/A
Perinatal outcomes: Early diagnosis compared to late diagnosis tended to have higher frequency of adverse neonatal outcomes (RDS, IVH, congenital malformations, neonatal infection and hypoglycaemia) but none reached statistical significance
Early diagnosis compared to controls had significantly higher rates of pre-term delivery (19.6% vs 7.6%), RDS (3% vs 0.5%), IVH (2% vs 0.2%), neonatal infection (4% vs 1%) and hypoglycaemia (4% vs 1%)
Late diagnoses compared to controls had significantly higher rates of pre-term delivery (17.9% vs 7.6%) and LGA infants (17.1% vs 10.9%)
No significant difference was noted between groups for low APGAR scores, birth injuries, meconium aspiration or transient tachypnoea of the newborn
Neonatal outcome and insulin use: Macrosomic infants tended to be more common in the early diagnosis group who were not treated with insulin compared to those who were (N/S, 17.8% vs 10.5%; $p = 0.29$). Similar rates were observed in late diagnosis group treated with and without insulin (18.7% vs 16.4%)
Compared to the general population, the study group were more likely to be Hispanic, nulliparous and have one or more RFs
Author's conclusion: A sizeable proportion of GDM patients can be diagnosed early in pregnancy. The difference in maternal characteristics and insulin requirements between early and late diagnosis groups suggests heterogeneity of GDM or the possibility of pre-existing IGT in the early diagnosis group

Comments
Compared study group to the general population. Findings suggested study population self-selected. Only 36% of the population met the inclusion criteria
All those included were screened on two occasions
Assumes all cases of GDM detected by GCT ≥ 135 mg/dl and abnormal GTT around 24 weeks

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Berkus and Langer, 1993[225]	San Antonio, Texas, USA Non-hypertensive gravidas including non-diabetic pregnant women found to have 1-h 50-g GCT ≥ 140 mg/dl and scheduled for GTT	Not RCT Cohort study *Aim*: to investigate whether the extent of glucose abnormality as reflected by the number of abnormal GTT values correlates with levels of CHO intolerance in pregnancy, and whether increasing degrees of abnormality signify greater adverse outcome 1400 subjects (764 with GDM and 636 with abnormal screening but normal GTT values) Classified according to number of abnormal values of 3-h 100-g GTT: 0 abnormal: 636 women 1 abnormal: 86 women 2 abnormal: 347 women 3 abnormal: 237 women 4 abnormal: 94 women	Screening test: 1-h 50-g GCT. Threshold ≥ 140 mg/dl Testing during morning clinic. Timing and fasting status not stated Diagnostic test: 3-h 100-g GTT. GDM diagnosed using modified NDDG values[18] with 2 or more values ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl Those with GDM scheduled for glucose control: regular diabetic diet (ACOG guidelines[4]); insulin when fasting glucose > 105 mg/dl or 2-h postprandial values consistently > 140 mg/dl Control adequate if fasting glucose 60–90 mg/dl; preprandial 60–105 mg/dl; 2-h postprandial < 120 mg/dl; and 2–6 am value 60–90 mg/dl	53% (746 out of 1400 women with abnormal GCT screening) GDM (NDDG criteria) Significant differences in demographics between 0 abnormal and other groups: less likely to be obese; lower mean glucose screening value LGA infants (≥ 90th percentile for gestational age based on institution) incidence significantly less for 0 abnormal group than other groups ($p < 0.01$): 0 abnormal: 13% 1 abnormal: 24% 2 abnormal: 27% 3 abnormal: 23% 4 abnormal: 23% Rates of treatment with insulin differed significantly between groups 2, 3 and 4 (14% vs 25% vs 48%)

Sensitivity and specificity

Results as reported. Raw data not available

Sensitivity and specificity of predicted value of 3.5 h taken for BG to return to fasting levels after 100-g GTT in diagnosing GDM (NDDG): sensitivity 86%; specificity 80%

Sensitivity and specificity of predicted value of 3.5 h taken for BG to return to fasting levels after 100-g GTT in identifying LGA infants: sensitivity 60%; specificity 46%

No significant increase in LGA with good control except for those with 2 abnormal GTT values (authors suggest that reduced use of insulin in the group may influence results)

Rates of LGA according to number of abnormal GTT values and level of glucose control (poor if fasting glucose > 90 mg/dl):

0 abnormal: 13% (untreated)

1 abnormal: 23% (untreated)

2 abnormal: good control 18% vs poor control 29%

3 abnormal: good control 16% vs poor control 24%

4 abnormal: good control 10% vs poor control 22%

Rates of LGA according to number of abnormal GTT values and level of glucose control (poor if 2-h postprandial > 120 mg/dl):

2 abnormal: good control 23% vs poor control 31%

3 abnormal: good control 18% vs poor control 31%

4 abnormal: good control 10% vs poor control 27%

Authors' conclusions: One or more abnormal GTT values were associated with comparably elevated incidence of LGA infants in patients with poor glycaemia control. Achievement of recommended glucose control decreased adverse outcomes to near normal levels

Authors' comment: Control of glucose during pregnancy becomes a confounding factor that influences the incidence of LGA

Comments

Population all had abnormal screening GCT

Outcomes defined. Confounding factors considered

LGA rates compared using two different criteria for diagnosing poor control

No costs

Gestational state of subjects not reported

Because gestational stage at which GDM was diagnosed was not reported may have underestimated GDM rates due to later development of GDM

? Did any of those with abnormal screening and initially normal GTT go on to develop GDM or those with 1 or more abnormal levels later develop a greater degree of glucose abnormality

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Bobrowski *et al.*, 1996[101]	USA 66.8% African-American, 26.4% Caucasian, 6.8% other Mean age not stated RF patients also included, i.e. maternal age > 30, obesity, previous macrosomia, PH GDM	Selected women with abnormal GCT No randomisation or control 422 women *Aim*: to investigate the 50-g GCT as a diagnostic test	All women had had 50-g GCT at 24 (± 7) weeks Those abnormal (≥ 135 mg/dl) given 3-h fasting 100-g GTT following 3-day CHO loading diet Diagnosis made on 2 or more glucose that meet NDDG criteria[18] (fasting ≥ 105; 1-h ≥ 190; 2-h ≥ 165; 3-h ≥ 140 mg/dl) Also compared with C&C criteria[20] (fasting ≥ 95; 1-h ≥ 180; 2-h ≥ 155; 3-h ≥ 140 mg/dl) Those with fasting glucose ≥ 140 mg/dl not given GTT as considered GDM	Incidence 29% with NDDG criteria; 38% with C&C 27 patients had GCT between 200 and 219 mg/dl (67% had GDM with NDDG and 70% C&C) and 27 had GCT > 220 mg/dl (100% had GDM with both criteria, 13 not given GTT as fasting glucose ≥ 140 mg/dl) Highest screen value without evidence of GDM was 216 mg/dl
Sensitivity and specificity N/A, as only those with positive screening were followed up PPV using NDDG criteria = 29% PPV using C&C criteria = 38%				
Comments High incidence in this study – not discussed in paper Time of GTT not defined No cost data Authors conclude that GCT result ≥ 220 mg/dl can obviate the need for a 3-h GTT. If fasting glucose ≥ 140 mg/dl no need for 3-h GTT				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Bochner *et al.*, 1987[71]	Los Angeles, California Study conducted 1984–86 Inclusion criteria: women with GDM diagnosed before 34 weeks' gestation (O'Sullivan and Mahan criteria[19]) Managed with diabetic pregnancy diet, weekly fasting and postprandial PG maintained < 100 mg/dl and 120 mg/dl respectively on diet alone Accurate early pregnancy dating Delivered after 36 weeks' gestation Mean maternal age 31 years Nulliparous 40% Mean pre-pregnancy weight 125 lbs Mean gestational age at which GDM was diagnosed 30.5 weeks (SD 1.06)	Not RCT Observational study *Aim*: to determine whether an early third trimester fetal AC measurement can be used in patients with GDM to predict the presence/absence of macrosomia and labour dystocia at term 201 women participated	Screening test used: ultrasonographic fetal measurement between 30 and 33 completed weeks of pregnancy AC percentile ranks based on local population Diagnostic test: 3-h 100-g GTT GDM diagnosed according to O'Sullivan and Mahan criteria (two or more values ≥: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl) Outcomes compared for those with AC > 90th percentile and those with AC ≤ 90th percentile according to values for local population	No GDM rates reported 15% (31/201) had macrosomia (defined as birth weight > 90th percentile for population) Group with AC > 90th percentile had significantly greater overall rates of CS, shoulder dystocia, and birth trauma ($p < 0.001$) than those with AC < 90th percentile Overall CS: 41.7% vs 14.3% Shoulder dystocia 16.7% vs 0.7% Birth trauma 3.6% vs 20.8%

Sensitivity and specificity

Of fetal AC measurement at 30–33 weeks for predicting macrosomia at term: sensitivity 87.8% (36/41); specificity 82.5% (132/160); PPV 56.3% (36/64)

NPV of AC ≤ 90th percentile for predicting the absence of macrosomia = 96.4% (132/137)

No significant differences between groups (AC > 90th percentile vs AC ≤ 90th percentile) in maternal age, % nulliparous, mean pre-pregnancy weight, mean gestational age at which GDM was diagnosed, or timing of AC measurement

Outcomes of 31 infants with birth weight > 4000 g: 8 had CS for failure to progress; 4 had shoulder dystocia; 3 had birth trauma

Group with AC ≤ 90th percentile were not at increased risk compared to general local population, for CS due to failure to progress (15.7% vs 16.6% or shoulder dystocia (0.8% vs 0.8%)

Authors' conclusions: Patients with non-insulin-dependent GDM with fetal AC measurements ≤ 90th percentile for population at 30–33 weeks are not at increased risk for macrosomia, CS or birth trauma at term as long as weekly glucose testing remains within normal limits

Comments

Well presented with defined outcomes

Conclusions supported by the evidence, although small numbers with shoulder dystocia ($n = 5$) and birth trauma ($n = 10$)

Ethnic mix of population not described

% of GDM population adequately controlled on diet not reported

Diagnosis of GDM seemed to be made rather late (mean 30.5 weeks)

90th percentile birth weight not reported for this population

Costs of measuring AC not reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Branchtein *et al*., 1997[129]	Porto Alegre, Brazil Consecutive women in general prenatal care units Inclusion criteria: age ≥ 20 years, 21–28 weeks pregnant, no PH DM outwith pregnancy, attended 2 antenatal clinics between 1991 and 1993 Mean age 28 years; BMI before pregnancy 24; gestation on enrolment 23 weeks FH DM in 15% Ethnicity: white 66%; black 16%; other 17% (including mixture of black, white, American Indian)	Not RCT Cross-sectional study *Aim*: to evaluate the relationship of central fat distribution with GDM 1113 women eligible 1025 (92%) reported as completing study but 959 (86%) included in analysis Multiple linear regression model of 2-h glycaemia vs WHR or waist included the following variables: age, sum of skinfold thicknesses, height, ambient temperature, parity, FH DM, uterine height, skin colour, obstetric antecedents, years of education, and prenatal clinic. Only variables with $p < 0.05$ included in final model	Diagnostic test: standardised 2-h 75-g GTT GDM diagnosed using WHO criteria[24] (≥ 11.1 mmol/l) IGT if 7.8–11.0 mmol/l	0.4% (4/959) women were diagnosed as GDM (WHO criteria[24]) 6.4% (71/959) women were diagnosed as IGT Positive association between the two measures of central fat deposition (WHR and waist circumference) and 2-h post-GTT glycaemia independent of other factors (see types of study column) Glycaemia increased 0.11 for each increase in SD of WHR (SD 0.06) and 0.13 mmol/l for each increase in SD of waist circumference (SD 0.06) $p < 0.02$ Limiting analysis to 870 women with uterine height ≤ 26 cm (where there was little change in fat measures with height): Glycaemia increased 0.13 for each increase in SD of WHR (SD 0.06) and 0.19 mmol/l for each increase in SD of waist circumference (SD 0.06) $p < 0.02$ 1 SD change in skinfold thickness (SD 24.7 mm) was associated with glycaemia change of 0.22, and 1 SD in age (SD 5.5) with 0.19 mmol/l

Sensitivity and specificity
N/A
Uterine height ≤ 26 cm: the WHR changed little (WHR 0.0015 larger with each additional cm of uterine height)
Uterine height ≥ 27 cm: marked increase with height (0.007/cm of height)
Authors' conclusion: Central fat distribution in pregnancy is an independent RF for gestational IGT

Comments
Sample considered to be representative of general local population of Brazil
Analysis controlled for potential confounding factors
Number completing study differs from number analysed (1025 completed study, 959 analysed on one occasion and 955 on another occasion). Discrepancies in these values were not explained
No demonstration that relationship between either WHR or waist circumference and glycaemia was linear. Interactions among variables in model not reported. No report of amount of variability explained by the model or comment on any outliers
Sample contained only 4 women diagnosed as GDM (WHO criteria[24]), thus limiting power of results
Relationship between measure of central fat distribution and either fetal/maternal outcomes not examined
No assessment of any change in value of screening for GDM using RFs with the addition of either of these measures of central fat distribution
No assessment of inter-examiner variability in estimating variables such as uterine height, skinfold thickness, WHR or waist circumference

All other studies *contd*

Reference	Study population and selection	Types of study	Outcome
Campbell *et al.*, 1989[226]	Ethnic groups not described Country of study not specified ? Aberdeen (recipe for 'Aberdeen mixed nutrient meal' in appendix)	Not RCT. Predictive power of screening tests not evaluated. Diagnostic test not used *Aim*: to assess the A: reproducibility and variability of BG response to standardised pre-packaged test meal (2 test meals, 7 days apart in 20 healthy women with no RFs for GDM at 30–32 weeks' gestation. BG measured at 4, 10, 15, 20, 30, 40, 50, 60, 90, 120 min after meal) B: reproducibility of 75-g OGTT (2 GTTs, 2 weeks apart in 10 healthy women at 30–34 weeks' gestation. BG measured at 4, 10, 20, 30, 40, 60, 90, 120 min after meal) C: comparability of BG responses to glucose test meal at breakfast vs lunchtime, the latter with no fasting (10 healthy women at 30–32 weeks' gestation. BG measured as in A) D: comparability of BG responses to test meal at breakfast and lunchtime according to activity levels (5 women ate meal at breakfast time on 2 occasions, 5 ate meal at lunchtime. On one occasion women rested, on the other travelled from home to clinic. BG as in A) E: BG response to test meal according to stage of pregnancy (18 primigravida ate test meal at 16, 26 and 36 weeks. BG measured at 30, 60, 90, 120 min after meal)	No GDM rates reported A: reproducibility and variability of BG response to standardised pre-packaged test meal: paired *t*-tests. No difference in summed PG or in maximum glucose reached B: reproducibility of 75-g OGTT: differences between tests N/S C: comparability of BG responses to glucose test meal at breakfast vs lunchtime: only the 90-min BG value differed. It was significantly lower at lunch than at breakfast D: comparability of BG responses to test meal at breakfast and lunchtime according to activity levels: differences between tests associated with activity levels N/S E: BG response to test meal according to stage of pregnancy: differences between tests according to stage of pregnancy N/S

Sensitivity and specificity
N/A
Authors' conclusion: PG response following a standardised pre-packaged test meal is highly reproducible, does not vary significantly with time of day or level of activity, is acceptable to pregnant women though a few found difficulty in eating it all in early pregnancy, and it does not change with the stage of pregnancy
Provided it can be shown to be of clinical value in detecting pathological pregnancies, it could be used in place of GTT for the diagnosis of GDM

Comments
Very small sample size would not have had the statistical power to detect differences in BG responses
Test meals evaluated in women with no GDM – responses in women with GDM were not evaluated. Results may differ in women with abnormal glucose tolerance
Evaluation of test meal did not include the ability to diagnose GDM
Reports that 'a few found difficulty in eating' all the test meal in early pregnancy, but numbers experiencing problems not reported
Analysis involved multiple statistical tests with no adjustment of significance levels
Costs of test meal compared with standard glucose load not reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Carpenter and Coustan, 1982[20]	USA Private patients aged > 25 years (range 25–44) 93.8% white 49.1% nulliparous Previous diabetics, multiple gestation excluded	381 unselected women. No randomisation, no controls *Aim*: to investigate screening test thresholds with aim of increasing sensitivity	Screening test: 50-g oral GCT CHO loading 3 days before and fasted after midnight Testing mainly between 24 and 33 weeks (3 outliers) Threshold: O'Sullivan and Mahan criteria[19] 143 mg/dl Modified criteria cut-offs: 130 and 135 mg/dl Diagnostic test if GCT ≥ 130 mg/dl a 3-h fasting 100-g GTT using modified O'Sullivan and Mahan criteria due to measuring glucose in plasma not in whole blood 2 or more values ≥: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl	Incidence GDM = 6.29% No adverse effects of test noted At O'Sullivan and Mahan threshold (143 mg/dl), 4 GDM patients would have been missed Below 135 mg/dl, no patients diagnosed with GDM, although no patients under 130 mg/dl given GTT
Sensitivity and specificity Using O'Sullivan and Mahan threshold, 143 mg/dl: Sensitivity 83% (20/24); specificity 87% (312/357); PPV (20/65) 30% Using 135 mg/dl: Sensitivity 100% (24/24); specificity 79.8% (285/357); PPV (24/96) 25%				
Comments Those screened below 130 mg/dl not given GTT (272 patients) No cost data available Authors propose, in women > 25 years, lowering GCT threshold to 135 mg/dl, will reduce to almost zero those missed who have GDM (based on this data and O'Sullivan and Mahan's)				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Casey *et al.*, 1997[181]	Dallas, Texas, USA Sample collected 1991–95 from community clinics Inclusion criteria: Study group: singleton, cephalic-presenting pregnancies in women with class A1 GDM Controls: states 2 controls matched for each case on ethnicity (black, white, Hispanic, Asian and other), maternal age, maternal weight and parity. But actually mainly uses general obstetric population as control Exclusion criteria: Non-cephalic presentation or multiple gestation Ethnicity: White: A1 GDM 11%; control 12% Black: A1 GDM 16% vs control 27% ($p < 0.001$) Hispanic: A1 GDM 69%; control 58% Mean age: A1 GDM 28 years vs control 24 years ($p < 0.001$) Maternal weight: A1 GDM 186 lbs vs control 167 lbs ($p < 0.001$) Parous: A1 GDM 72% vs control 59% ($p < 0.001$)	Not RCT Retrospective cohort *Aim*: to compare pregnancy outcomes in women with and without glucose intolerance Study group: 874 women with A1 GDM General obstetric population: 61,209	Screening test used: all women had random serum BG (RBG) on initial visit (threshold ≥ 130 mg/dl) Women with RFs (glycosuria, RBG ≥ 130 mg/dl, PH GDM, macrosomia/hydramnios, symptoms of DM) screened with 1-h BG after 50-g GCT at any time during pregnancy Women with other RFs (FH DM, previous stillborn, malformed or macrosomic infant) screened 24–28 weeks Threshold for GCT 140 mg/dl Fasting/fed status not reported Diagnostic test: 3-h 100-g GTT after overnight fast. Abnormal if 2 or more abnormal values according to NDDG criteria[18] Class A1 GDM: at least 2 abnormal levels of GTT excluding fasting level. Received dietary counselling Class A2: elevated FBG. Received insulin if elevated FBG persisted	1.2% (874/72,572) diagnosed as class A1 GDM (NDDG criteria without elevated FBG). This presumably excludes those diagnosed as GDM class A2 Pregnancy outcomes: The following outcomes were significantly more common in class A1 GDM compared to the general population: pregnancy-induced hypertension (17% vs 12%: $p = 0.001$); CS (30% vs 17%: $p < 0.001$); repeat CS (16% vs 9%: $p < 0.001$); CS for dystocia (8% vs 4%: $p < 0.001$); shoulder dystocia (3% vs 0.9%; $p < 0.001$) Infant outcomes: The following outcomes were significantly more common in class A1 GDM compared to the general population: LGA (35% vs 14%: $p < 0.001$); fractured clavicle (2% vs 0.9%: $p = 0.017$)
Sensitivity and specificity N/A Women with GDM were significantly older, heavier and more likely to be parous or Hispanic (see under study population) Infant outcomes: there was no statistically significant difference in the following outcomes between women with A1 GDM and those without GDM: admission to SCBU (47/874 vs 3214/61209); 5-min APGAR score ≤ 3; Erb's palsy; stillbirth; neonatal death; and umbilical artery pH ≤ 7 Multiple logistic regression used to examine relationship between LGA and class A1 GDM after adjusting for maternal weight: RR = 1.54 (95% CI, 1.35 to 1.74). LGA rate among women without GDM matched for age, parity, weight and ethnicity was 23% vs 35% in class A1 GDM women. 12% risk attributable to class A1 GDM *Authors' comments*: Vaginal delivery after CS not practised in this unit for women with GDM. This could account for higher repeat CS rates *Authors' conclusion*: The main consequence of class A1 GDM is excessive fetal size leading to increased risk of difficult labour and delivery. This study estimates that one in eight women with class A1 GDM will deliver an LGA infant			
Comments Good sample size though small numbers with neonatal morbidity in class A1 group (5-min APGAR score ≤ 3, $n = 0$; Erb's palsy $n = 3$; stillbirth $n = 4$; neonatal death $n = 1$; and umbilical artery pH ≤ 7 $n = 4$) Predominantly Hispanic population (69% of study group) Confused re methodology: stated 2 matched controls for each case of GDM but then goes on to use general obstetric population as the comparison group (61,209 women) for much of the report Assumes all cases of GDM were detected by BG ≥ 140 mg/dl on screening with GCT followed by abnormal GTT at whatever time in pregnancy tests were performed Class A1 women in study were diagnosed and received dietary counselling but this does not seem to have been effective in reducing risks of LGA infants to that of normal population. No comment on this was made in article No mention of rates of diagnosis of class A2 GDM or outcomes in those classified as A2 GDM			

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Cooper *et al.*, 1979[227]	Haifa, Israel Setting: antenatal clinic of University hospital Sample selected from women with one or more classical RFs for GDM (previous infant ≥ 4500 g or neonatal death, low fertility, FH DM, previous GDM, suspected large fetus, glycosuria, toxaemia, hypertension, hydramnios, marked obesity ≥ 90 kg) Women with persistent glucose intolerance were considered established GDM and were subsequently excluded from the study Ethnicity of sample not reported Mean age 27 to 34 years across groups	Not RCT *Aim*: to compare the relative validity of the 2-h post-prandial glucose test with the rapid intravenous GTT as predictors of outcome in pregnancy 77 women were included	All women underwent both the following combined screening/ diagnostic tests within a week at start of third trimester 1. rapid IVGTT: intravenous administration of 25 g glucose in 50% solution after overnight fast. BG tested at 10 and 55 mins. Abnormal if K < 1.15% K = ((ln (C1/C2)) / (t2–t1)) x 100 where t = time and C1 and C2 = BG at t1 and t2. 2. 2-h post-prandial VBG 2 h after standard provocative breakfast (100 g CHO). Abnormal if BG ≥ 145 mg/100 ml Women with abnormal levels of either test treated by low-calorie, low-CHO diet. Women with persistently abnormal tests given more intensive treatment including insulin and early termination of pregnancy	44% (34/77) women had abnormal IVGTT tests 35% (27/77) women had abnormal 2-h pp test Women classified into 4 groups using results: 1. 2-h pp abnormal, IVGTT normal (11 women) 2. IVGTT abnormal, 2-h pp normal (18 women) Both groups 1 and 2 received advice re diet 3. both tests normal (32 women) 4. both tests abnormal (16 women). This group received intensive treatment No significant differences were noted between groups 1, 2 and 3 on maternal age or length of pregnancy The 'only IVGTT abnormal' group had significantly heavier babies than the 'only 2-h pp abnormal' (3935 vs 3160 g: $p = 0.007$) and the normal group (3935 vs 3428: $p < 0.02$) Babies of women with 'only abnormal 2-h pp' had significantly lower rates of perinatal morbidity than groups 2 and 3. 'Only 2-h pp abnormal' 18.2% (2/11) vs 'only IVGTT abnormal' 27.8% (5/18) vs both tests normal 28.1% (9/32): $p < 0.05$)

Sensitivity and specificity
N/A
Linear regression was used to investigate the correlation between birth weight and results of screening. Found that IVGTT was significantly correlated with birth weight ($p < 0.05$) but that 2-h pp was not
However, IVGTT accounted for only 11.1% of the variance in the regression
% agreement on IVGTT and 2-h pp = 62% (48/77)
Authors' comment: The IVGTT has not been compared to the standard full GTT
Authors' conclusion: The rapid IVGTT is a more dependable test than a single postprandial BG estimation in screening pregnant women for GDM
NB: Perinatal morbidity defined as prematurity, respiratory distress syndrome, congenital anomalies, hypoglycaemia, neonatal asphyxia, plethoric habitus)

Comments
Sample was not selected from general population but from those with classical RFs for GDM
Ethnicity not reported
Assumes all cases of abnormal GT were diagnosed by tests in early third trimester. Would have missed out the chance of treating women who developed GDM at an earlier stage of pregnancy
Birth weight does not seem to have been adjusted for gestational age. There was no adjustment for other potential confounding factors
Small numbers in subgroups
No costs were reported
Presents profile of all those pregnancies with perinatal morbidity. No comment made about the greater perinatal morbidity in the group testing normal on both tests compared to the group with only 2-h pp abnormal (28% vs 18.2%). Perinatal morbidity rates appear similar in the normal group and the group with only IVGTT abnormal (28.1% vs 27.8%). This would not seem to support use of the IVGTT. Small numbers in groups 1 and 2 (11 and 18 women)
Without further investigation of the effect of potential confounders on these results, the authors' conclusions cannot be considered supported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Corcoy *et al.*, 2000[137]	Spain No details of sample 158 consecutive women used No inclusion or exclusion criteria given	Not RCT *Aim*: to assess if the delay in processing of blood samples for GDM screening influences the results	5 ml blood taken consecutively Sample 1 immediately centrifuged and PG was analysed Sample 2 kept at room temperature and analysed after the usual time delay Collection tubes were standard 2.5 g/l NaF vacutainers	Mean glucose concentrations were 6.75 in sample 1 and 6.5 in sample 2 ($p < 0.001$) PG decrease was significant in 21 samples measured with a processing delay of < 30 minutes ($p < 0.01$) whereas further decreases were not significant Of 35 women with a positive screening test in sample 1 samples, 6 were misclassified as negative in sample 2 samples
Comments The overall rate of GDM screening misclassification (3.8%) is similar to that described for OGTT May have been influenced by the concentration of NaF used in the collection tubes Conclude that 2.5 g/l NaF is not sufficient to prevent PG decrease in the first 30 minutes				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Corcoy *et al.*, 1991[183]	Spain No description of population or age range Taken from centre for high-risk pregnancies	Unselected pregnancies No randomisation Used normal screens as a control 104 samples Known incidence of 16.2%	Screening test: 50-g non-fasting GCT tested at 1-h ≥ 7.8 mmol/l serum fructosamine: thresholds taken from values of 464 (59% high-risk women) = mean plus 2 SD Diagnostic test. 100-g fasting GTT. Thresholds not defined Screened at first ante-natal visit, 24–28 weeks and 32–35 weeks Compared GCT predictability for GDM with SF by testing differences between SF in normal and SF in positive screen or diabetic GTT	46% positive screening tests had GDM Screening with fructosamine predicted only 4/48 GDM Screening with 50-g GCT predicted 48/48
Sensitivity and specificity Those screened positive on GCT PPV 46% Sensitivity of SF = 8.33% Specificity 100%, PPV 100% (authors' data). However, taken from those who screened positive on GCT				
Comments No description as to population used No description of test results in respect to the three different times of testing Those screened negative with 50-g GCT not included in diagnostic test No cost data given				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Corrado *et al.*, 1999[102]	Sicily, no demographic details of women given, mean age 29.0 (SD 4.8)	Cohort study, no randomisation or control. Unselected, consecutive women screened. 1236 women in cohort, 1028 underwent screening. *Aim*: to establish the prevalence of GDM with both the NDDG[18] and C&C[20] criteria	Screened at mean 25 weeks, with a 1-h 50-g GCT. Those over threshold > 135 mg/dl given 3-h fasting 100-g GTT. NDDG criteria, 2 or more: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl. C&C criteria: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl	Of 1028 women 12 patients excluded due to multiple pregnancies, 16 did not have follow-up GTT Incidence GDM by NDDG criteria 34 (3.4%), by C&C 46 (4.6%) No differences between those screened and those not in terms of age, gestational age, parity or BMI Age of women significantly higher ($p < 0.0001$) in GDM than non-GDM women Non-GDM women significantly less likely to have RFs, such as FH, previous GDM, previous poor obstetric outcome or previous macrosomic infant (significance at least $p < 0.01$ for all)
Sensitivity and specificity No data available to calculate sensitivity or specificity and no women with negative screening test given GTT				
Comments RFs not defined before study commenced Some variability in the reporting of the numbers of women with positive screening values				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Coustan *et al.*, 1987[170]	Rhode Island, USA 70 women, 20 with known GDM and 50 being screened. Recruited from private practices	Comparison of breakfast vs 50-g GCT; crossover with randomised order	50-g GCT "Standard" breakfast: 600 cals, 47% fat, 35% CHO Gold standard was 3-h 100-g OGTT	Breakfast tolerance test performed better in ROC curve than standard GCT
Sensitivity of meal was 75% and specificity 94% using 120 mg/dl as cut-off; 96% and 74% using 100 mg/dl. Of 50 patients presumed normal, 45 had post-breakfast PG under 120 mg/dl, of whom 2 had GDM; 5 were over 120 mg/dl of whom 3 had no GDM on OGTT Meal would not be considered satisfactory by today's nutritional standards				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Coustan *et al.*, 1989[4]	USA Unselected women Aged < 20 to > 40 years (no mean given) 87.9% white, 8.1% black, 1.2% Asian, 0.5% Indian, 0.1% Chinese, 2.0% unknown	Cohort study No randomisation No control 6034 women with 6214 pregnancies over 2.5 years Assessed costs and sensitivity of screening test	Testing at 24–28 weeks 1-h 50-g GCT without regard to last meal Threshold of ≥ 130 mg/dl had 3-h 100-g GTT NDDG criteria[18] for diagnosis Women asked for information regarding age and RFs	Incidence GDM 2.0% As maternal age increased there was an increased likelihood of GDM ($p < 0.001$) However, 56% of GDM below 30 years. Of these, 58% had one or more RFs Also observed GDM cases if screening cut 140 mg/dl. 13 (10%) of women would have been undetected
Sensitivity and specificity Reported sensitivities (no data to check): 1. ACOG[4] (≥ 30 years or younger if RF) 65% 2. Universal screen if ≥ 25, or younger if RF with threshold 130 mg/dl – 85% 3. Universal screen if ≥ 25, or younger if RF with threshold 140 mg/dl – 95% 4. Universal screening threshold 140 mg/dl – 90% 5. Universal screening threshold 130 mg/dl – 100%				
Comments Reported costs (1989) based on: 1-h 50-g test = US$2.45; 3-h 100-g GTT = US$11.00 per 1000 cases: ACOG criteria[4] (current practice) (1); US$2470 (7 cases missed) Screen ≥ 25 years (2); US$3264 (3 missed) Screen ≥ 25 years (3); US$4085 (1 missed) Universal screen (4); US$4980 (to diagnose all) Universal screen (5); US$3990 (2 cases missed) Only women with positive screen given diagnostic test				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Damm *et al.*, 1992[202]	Denmark Women who had previously had GDM diagnosed from routine screening in clinic (based on RFs and FBG levels) and confirmed by 3-h 50-g GTT Cohort from 1978–85 Median ages at pregnancy: GDM women 30.1, controls 26.7 and GDM women who were not followed up 31.3 years	Observational study, no randomisation *Aim*: to determine the incidence of DM in women who had previously had GDM 355 women with previous dietary controlled GDM, 241 participated Control group of 74 (57 participated) women who did not have GDM Subgroup of 91 women who had GDM and in whom plasma insulin was measured in late pregnancy	All GDM women had a 50-g GTT at 8 weeks post-delivery and invited for yearly testing Controls tested for 2–11 years post-delivery Glucose tolerance was tested in all women except insulin-controlled GDM women by fasting 75-g GTT following 3 days 150 g CHO diet Thresholds by WHO criteria[24] Insulin-controlled women had an intravenous glucagon test for β cell function following overnight fast	42 (17.4%) women developed DM (3.7% insulin and 13.7% diet-controlled DM) No controls had DM Insulin group: median interval between pregnancy and diagnosis: 14 months (range 3–52). Significantly younger and leaner than those dietary controlled (p = 0.05) Diet group: median interval for diagnosis 48 months (range 2–117)
Predictive factors By univariate analysis (groups combined): parity ≥ 2, early diagnosis GDM, fasting glucose levels at diagnosis or at 2-month follow-up > 5 mmol/l, non-white race, pre-term delivery (< 3 weeks), and LGA (> 90th percentile) baby Maternal overweight (pre-pregnancy BMI ≥ 25) or a pre-pregnancy weight ≥ 120% of ideal weight, maternal age or FH of DM not predictors. Nor were pregnancy complications such as pre-eclampsia When assessed as independent variables in a logistic regression analysis: fasting glucose at diagnosis (OR 3.34; 95% CI, 1.84 to 6.12), pre-term delivery (OR 3.64; 95% CI, 1.14 to 11.65) and an OGTT that showed DM at the 2-month post-partum examination (OR 4.88; 95% CI, 2.14 to 11.15) When glucose values and glucose areas under the curve from GTT at diagnosis and post-partum were assessed as predictors, the fasting glucose at diagnosis and pre-term delivery remained significant predictors. Also the 2 h value on the diagnostic GTT and the 0–2 h glucose area under the curve at the post-partum GTT were significant predictors Subgroup analysis: fasting glucose at diagnosis and 0–2 h glucose area under the curve and the 0–60 min insulin area under the curve significant predictors of DM			
Comments Women with insulin-controlled GDM not tested Different diagnostic tests and criteria used for GDM Unable to evaluate predictive values for insulin and non-insulin dependent diabetics separately due to small numbers of insulin-dependent diabetics. Therefore, analysis on 'diabetics'. Risk for DM increased significantly if 2-month post-partum test positive. Authors suggest that important to test post-partum				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Danilenko-Dixon *et al.*, 1999[147]	USA 92.6% non-Hispanic white, 4.9% Asian, 1.7% Hispanic, 0.7% African-American and 0.1% other Ages in two groups 28.8 (SD 5.3) and 31.3 (SD 4.9) Possible sample 18,834 but 330 excluded as data not available as to presence of any RFs	Retrospective study *Aim*: to analyse the impact of selective screening (ADA, 1997[2]) with universal screening No randomisation, or controls 18,504 women	Between 24 and 30 weeks' gestation all women screened with a 1-h 50-g GCT following an overnight fast. Those with glucose ≥ 140 mg/dl given diagnostic 3-h 100-g GTT (fasting status not defined) Thresholds for diagnosis based on NDDG criteria[18] – 2 or more glucose at fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl	20% (3683 women) had a positive screening test 564 women (3%) diagnosed GDM If used selective screening as defined by ADA 3% (17 cases GDM) would have been exempt from screening Women with GDM were significantly older, more obese, more likely to smoke, to have had previous GDM and to have had a macrosomic or still birth No significant differences as a result of ethnicity (although only small numbers not non-Hispanic white)
Sensitivity and specificity Assume sensitivity 100%, specificity 83%, PPV 15%				
Comments ADA selective screening guidelines exclude women < 25 years, with normal body weight, BMI of < 27, not a member of a high prevalence ethnic group for DM with no first-degree relative with DM The use of this selective screening would only exempt 10% of this study population from screening (however only 17.8% of this population were younger than 25 years) Presumes that women screened negative not GDM No cost data available				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Deerocha-nawong *et al.*, 1996[110]	Bangkok, Thailand Inclusion criteria: women attending antenatal clinic were randomly selected Exclusion criteria: women with pre-existing DM All women were Thai Mean age 27 years Mean BMI 22.4	Not RCT *Aim*: to compare the criteria for diagnosis of GDM by NDDG[18] and WHO[24] criteria and to study the outcome of GDM when diagnosed by these criteria 709 women were included	Screening tests used: 50-g GCT (threshold 7.8 mmol/l at 1 h) between 24 and 28 weeks. Plus, in same week, WHO diagnostic 75-g GTT. Tests in random order If positive 50-g GCT underwent 100-g GTT Diagnostic tests NDDG criteria: 3-h 100-g GTT within 7 days of 75-g GTT. GDM diagnosed if BG ≥ fasting 5.8; 1-h 10.6; 2-h 9.2; 3-h 8.1 mmol/l WHO criteria: 2-h 75-g GTT with threshold 7.8 mmol/l at 2 h	1.4% (10/709) diagnosed as GDM by NDDG criteria 15.6% (111/709) diagnosed as GDM by WHO criteria All those diagnosed as GDM by NDDG criteria were GDM by WHO criteria Reported that GDM by WHO criteria (111 women): 35 received diet, 6 received insulin GDM by NDDG criteria (10 women): 2 treated with insulin, 3 with diet

Sensitivity and specificity

Sensitivities reported, specificities and PPV calculated

Sensitivity and specificity of GDM criteria in detecting macrosomia (> 4000 g):

NDDG criteria: sensitivity 21% (3/14); specificity 99% (688/695); PPV 30% (3/10)

WHO criteria: sensitivity 43% (6/14); specificity 85% (590/695); PPV 5% (6/111)

Sensitivity and specificity of GDM criteria in detecting LGA (> 90th percentile):

NDDG criteria: sensitivity 10% (4/41); specificity 99% (662/668); PPV 40% (4/10)

WHO criteria: sensitivity 41% (17/41); specificity 86% (574/668); PPV 15% (17/111)

The following were more common in women with GDM (by both NDDG and WHO criteria) than controls: pregnancy-induced hypertension, CS, macrosomia, LGA, hypoglycaemia

NDDG criteria (10 women):

Pregnancy-induced hypertension OR 20.48 (95% CI, 6.24 to 67.2)

CS OR 8.55 (95% CI, 2.24 to 32.6); macrosomia OR 21.28 (95% CI, 6.13 to 73.8); LGA OR 10.86 (95% CI, 3.19 to 36.9); hypoglycaemia OR 46.60 (95% CI, 15.50 to 140)

WHO criteria (111 women):

Pregnancy-induced hypertension OR 2.72 (95% CI, 1.71 to 4.31); CS OR 1.89 (95% CI, 1.18 to 2.41); macrosomia OR 2.84 (95% CI, 1.51 to 5.33); LGA OR 2.95 (95% CI, 1.96 to 4.44); hypoglycaemia OR 3.99 (95% CI, 2.34 to 6.83)

Authors' conclusion: Findings suggest that WHO criteria patients had significantly worse outcome of pregnancy and fewer perinatal complications missed than the more cumbersome NDDG criteria

Comments

Sample randomly selected from Thai population. Method of random selection was not reported

No adjustment for potential confounding factors in estimating OR of outcomes of pregnancy

Cost-effectiveness of alternative strategies was not compared

Small number of adverse outcomes led to wide 95% CI, especially for NDDG criteria

Criteria for insulin use in GDM women (however defined) was not reported. Some women who were diagnosed as GDM appear to have received no treatment – reasons for this are not reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Deitrich *et al.*, 1987[177]	Omaha, Nebraska, USA Setting: large private medical practice Subjects predominantly Caucasian (95%) with high school or higher level of education and full financial support for medical costs Mean maternal age 24–25 years; obesity 29%; strong FH DM 14–15%	Not RCT Controlled trial *Aim*: to compare the value of routine vs selective screening for GDM Allocated according to attending physician 2000 women overall, including 1. 1000 women allocated to routine GDM screening 2. 1000 women allocated to selective screening with screening only in the 453 with standard RFs: age > 30 years, obesity, first-degree FH of DM, poor obstetric history, previous macrosomic infant, or persistent glycosuria	Women received either routine/selective screening 1. Routine screening: all women started between 24 and 28 weeks. All had GCT 2. Selective screening: only those with RFs underwent GCT, started as early as conveniently possible Screening test: 1-h post 50-g GCT with commercially prepared solution. Abnormal if BG ≥ 150 mg/dl on photometer machine. Fasting not required Diagnostic test: 3-h 100-g GTT within 1 week of abnormal GCT. Prior 3-day CHO load advised GDM diagnosed if two or more BG values elevated: fasting > 105; 1-h > 195; 2-h > 145; 3-h > 130 mg/dl	Overall 2% (40/2000) diagnosed as GDM (see previous GDM criteria) Routine screening: 10.6% (106/1000) had abnormal GCT (≥ 150 mg/dl); 2.1% (21/1000) diagnosed as GDM Selective screening: 453/1000 screened, of which 22% (84/453) had abnormal GCT (≥ 150 mg/dl); 4.2% (19/453) diagnosed as GDM Incidence of infant birth weights ≥ 4000 g, 5-min APGAR scores, 7, and perinatal deaths (0 vs 0.1%) were not significantly different between the routine and select screening groups. No figures given

Sensitivity and specificity

Results below assume detected all cases of GDM

Sensitivity and specificity of routine screening in all women using GCT ≥ 150 mg/dl: sensitivity 100% (21/21); specificity 91% (894/979); PPV 20% (21/106)

Sensitivity and specificity of selective screening in high-risk women using GCT ≥ 150 mg/dl: sensitivity 100% (19/19); specificity 85% (369/434); PPV 23% (19/84)

Costs: estimates that in this practice with 1700 deliveries, using selective screening would allow 850 fewer 1-h GCT each year with annual savings of US$8500 (1987)

Authors' comment: Others should be encouraged to undertake a similar analysis of their own experience before altering any previously proposed standards

Authors' conclusion: This assessment of clinical practice has allowed the authors to safely eliminate the need for GDM screening in more than half of their private patients. This will reduce office times, patients inconvenience and expense

Comments

Selected population of middle/upper class patients. Results may not be applicable to other populations

Patients not randomly allocated to screening group. Groups comparable on age, race, nulliparity, gestation at delivery, obesity, FH DM

Groups comparable on age, race, nulliparity, gestation at delivery, obesity, FH DM

Timing of screening differed between groups with routine screening at 24–28 weeks and screening in the selective population starting as soon as possible. Actual timing of screening not reported

Assumes all cases of GDM detected by GCT screening test ≥ 150 mg/dl in first group and assumes all cases of GDM in selective screening group were detected by the presence of RFs followed by a GCT ≥ 150 mg/dl. Thus may have failed to detect some cases of GDM

Gestational stage of actual screening not reported

Comments from other authors about this study include (5 clinicians commented):

1. Questions whether this study identified all women with GDM by screening using a cut-off level ≥ 150 mg/dl and if not, then the authors cannot assess the importance of RFs. Cost analysis needs to take the cost of missed diagnosis into account.
2. The conclusion reached is at variance with other data, which suggest that 40% of women with known GDM have no historical RFs. Notes that age was not a significant predictor in this study in contrast to other studies.
3. One clinician comments that rates of emesis, nausea and vomiting after glucola testing were 11% to 14% in their population and they changed to full 100 g CHO meal.
4. Asks if conclusions are applicable only to their type of private practice and says that conclusions do not mean that others should not screen all their patients.
5. In current litigious society would be difficult to defend that rare missed individual who has DM.

Dietrich responded:

1. Population is primarily middle and upper middle class. Patients seen very early (average 8.3 weeks).
2. Study was not designed as a recommendation for the general obstetric population.
3. They had no significant problems with nausea and vomiting after glucola.

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Di Cianni *et al*., 1997[228]	Italy No demographic details of women given, mean age 29.5 (SD 4.7)	Retrospective observational study of women undergoing standard screening practices in clinic 2000 cases reported from unselected women	At 24–28 weeks all women had GCT (glucose load not defined). RF women given screen at 14–18 weeks. Those with BG > 140 mg/dl given 3-h 100-g GTT and diagnosis made on C&C criteria[20] Women with GDM given dietary control unless fasting glucose and/or 1 h glucose higher than 95 mg/dl and 130 mg/dl respectively Obstetric and metabolic parameters, maternal and neonatal outcomes noted RFs: FH DM, age > 35 years, previous macrosomia and previous GDM	102 women with RFs had GDM and 23 without had GDM (overall incidence 6.25%) Took GDM women and 110 women with one abnormal value on GTT together and compared with non-GDM women for effects of RFs and outcomes. Higher mean age ($p < 0.05$), pre-pregnancy BMI ($p < 0.01$), more women > 35 years ($p < 0.05$), previous macrosomia ($p < 0.01$) and previous GDM ($p < 0.01$) demonstrated in GDM/one abnormal GTT result women. Prevalence of preterm deliveries ($p < 0.02$), CS ($p < 0.01$) also higher in GDM/one abnormal GTT result women Prevalence of macrosomia higher in GDM than one abnormal GTT and non-GDM women ($p < 0.05$). Macrosomia more prevalent in women who were obese pre-pregnancy than those of normal weight ($p < 0.01$) A higher prevalence of hyperbilirubinaemia, hypoglycaemia and hypertrophic cardiomyopathy in new-borns of GDM women ($p < 0.01$)
Sensitivity and specificity Reported sensitivity (expressed as the term sensibility ? translation error) 75.1%, specificity 44% and PPV 46.4% and NPV 74% however no data available to check, and not based on patients who had a negative screen				
Comments No clear details of GCT or GTT used reported *Authors' conclusion*: GCT has a good effectiveness for GDM diagnosis, however, cannot conclude that any negative screened women would not have had GDM				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
de los Monteros *et al.*, 1993[229]	160 Mexican women Mean reported ages between 27 and 32 No further descriptions of population characteristics Only those without history of DM included	Unselected women for two groups based on gestational age (early and late) randomly assigned to subgroups *Aim*: to see how reproducible the 50-g screening test was over 2 consecutive days Early women (80) tested at 12–13 weeks Late women (80) tested at 24–28 weeks Four subgroups, day 1 and day 2: Fasting–fasting Fed–fed Fasting–fed Fed–fasting	50-g oral GCT used Fasting category, overnight fast for 10–12 h Fed category, eaten normal breakfast within 2 h of test (no detail as to types of breakfast) Three thresholds considered: 130, 135 and 140 mg/dl No diagnostic test given to any women	No diagnostic test used No statistically significant glucose levels in any individual women. Mean glucose levels significantly different between day 1 and 2 in: early fasting–fasting, fasting–fed and fed–fast and late fasting–fasting, fasting–fed and fed–fast. In both fasting–fasting groups this was higher day 1 than 2, in other groups was higher day 2 than day 1 No significant differences in either fed–fed groups Serum glucose levels varied < 10% from each other in 32.5% early and 40% of the late patients Reproducibility in early groups, day 1 and 2 having levels over: 130 mg/dl = 61%, 135 mg/dl = 47%, 140 mg/dl = 43% In late groups 83% overall, no individual levels reported In both groups possibility of levels under 130 mg/dl = 90%
Sensitivity and specificity No data available as no diagnostic test used Significance values, day 1 vs day 2: Early group: fasting–fasting, $p < 0.046$, fasting–fed 0.05, fed–fasting 0.01 Late group: fasting–fasting, $p < 0.49$, fasting–fed, 0.02, fed–fasting, 0.001				
Comments No cost data available Mean values lower on the fed day in fasting–fed and fed–fasting groups, in accordance with previous reports *Authors' conclusion*: The screening test is moderately reproducible Reproducibility better under fed conditions, and in later testing				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Dornhorst *et al.*, 1992[115]	London, UK Sample: All women not known to be diabetic who attended clinic in teaching hospital and delivered in the hospital between 1984 and 1988 Multiracial population: 44.5% from ethnic minority groups; Indian sub-continent 10.5%; Afro-Caribbean 17.7%; SE Asia 5%; others including Middle-east 11.4%	Not RCT Retrospective analysis *Aim*: to determine the frequency of GDM according to age, BMI, parity and ethnic origin in women without known pre-existing DM and to analyse the influence of RFs separately for each ethnic group 11,205 women White (Northern European Caucasian) 6135 women Indian (Indian subcontinent) 1218 women Black (Afro-Caribbean) 1977 women SE Asian (ethnic Chinese) 572 Miscellaneous (including Arab, Mediterranean, Middle East) 1303	Screening test used: all screened at first antenatal visit with 1-h 50-g GCT (Lucozade®) in non-fasting state. Women with RFs (obesity, glycosuria, FH DM, previous macrosomic infant, unexplained stillbirth or GDM) re-screened at 28 weeks Abnormal if 1-h PG ≥ 7.8 mmol/l Diagnostic test: 3-h 50-g GTT. Diagnosed as GDM if area under glucose curve ≥ 43 units (Gillmer and Bickerton, 1994[230])	Overall 1.5% (170/112,056) diagnosed as GDM (Gillmer and Bickerton criteria) Women with GDM were significantly older (32.3 vs 28.3 years; $p < 0.001$), had higher BMI (27.7 vs 23.8; $p < 0.001$), more likely to be parous ($p < 0.001$) and from ethnic minorities (55.4% vs 15.3%; $p < 0.0001$) GDM rates by ethnicity: white 0.4% (26/6135); black 1.5% (29/1977); miscellaneous 3.1% (41/1303); SE Asian 3.5% (20/572); Indian 4.4% (54/1218) After adjusting for age, BMI and parity RR as below with white as reference category Black RR = 3.1 (95% CI, 1.8 to 5.5); miscellaneous RR = 5.9 (95% CI, 3.5 to 9.9); SE Asian RR = 7.6 (95% CI, 4.1 to 14.1); Indian RR = 11.3 (95% CI, 6.8 to 18.8)

Sensitivity and specificity
N/A
Age and GDM
The proportion with GDM increases with age from 0.1% in 16–19-year-olds to 5.2% in mothers > 40 years
BMI: The proportion with GDM increases with booking BMI from 0.5% in women with BMI < 20 to 10.5% in mothers with BMI ≥ 40
Parity: The proportion with GDM increases with increasing parity from 0.8% in primigravidas to 4.5% in mothers with parity ≥ 4
After adjusting for ethnicity, age and BMI, RR become non-significant
Influence of age, BMI and parity within each ethnic group
Summary results below were supported by adjusted RR given for each outcome:
Age ≥ 35 years was independently associated with increased risk of GDM in white, black, Indian, miscellaneous but not SE Asian women
BMI ≥ 27 was independently associated with increased risk of GDM in white, black and miscellaneous but not SE Asian or Indian women
Parity ≥ 3 was independently associated with increased risk of GDM in white, black and SE Asian women but not in miscellaneous or Indian women
Authors' conclusions: Ethnic origin has a major influence on the prevalence of GDM and the importance of other RFs varies between ethnic groups. These findings have important implications for the screening of women in pregnancy

Comments
Large sample size
Sample representative of local population
All women screened for GDM
Assumes all cases of GDM detected by positive screen at first clinic visit in women without RFs or by GCT at 28 weeks in women with RFs followed by abnormal GTT at whatever stage of pregnancy this was performed
Analysis adjusted for potential confounding factors
The author's conclusions were supported by the evidence presented

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Dornhorst *et al*., 1996[231]	London, UK Inclusion criteria: Asian and Europid with GDM (WHO criteria[24]) or with positive screening (> 7.8 mmol/l on GCT) and normal GTT Across groups: Mean age: 30–35 years BMI: 24–26	Not RCT Observational *Aim*: to assess Asian and white /Europid LGA infants of mother with GDM and those with lesser degrees of glucose intolerance 47 women with GDM (21 Asian and 26 white) 155 women with a positive screening test for GDM but normal GTT (34 Asian and 121 white)	Screening test used: universal screening with 1-h post 50-g GCT at 20 weeks' gestation. Abnormal if ≥ 7.8 mmol/l Fasting/fed status not reported Diagnostic test: 2-h 75-g GTT between 24 and 28 weeks using WHO criteria (2-h value ≥ 7.8 mmol/l) Women (? All/only those with GDM) received dietary advice. Those with GDM home tested glucose 2–4 times/day. Insulin started if FBG > 6 mmol/l or 2-h postprandial BG > 8 mmol/l 2 types of percentile charts used to classify infants: 1. Asian and Europid percentile charts calculated for each sex and both ethnic groups 2. Standard MRC 'Birth weight for Dates'	GDM rates N/A Overall the MRC charts compared with the ethnically derived charts tended to underestimate LGA Asian infants and overestimate Europid LGA infants (see results below) LGA regarded as > 90th or > 95th percentiles

Sensitivity and specificity

N/A

LGA defined as > 90th percentile

Overall (GDM + abnormal screening but normal GTT) the standard MRC charts identified fewer Asian LGA infants than the ethnically derived charts: 7/56 (13%) vs 15/56 (27%): $p = 0.06$)

Overall the standard MRC charts identified more Europid LGA infants than the ethnically derived charts: 33/147 (22%) vs 21/147 (14%) $p = 0.07$

9 Asian mothers (8 GDM and 1 with positive GCT but normal GGT) required insulin compared to 3 Europid mothers (1 GDM and 2 with positive screen but normal GTT)

Birth weights of Asian infants and Europid GDM mothers were similar (3539 g vs 3436 g). Asian and Europid mothers with GDM were of similar age, booking weight and BMI

Asian mothers with positive screen but normal GTT had lighter infants than Europid mothers (4141 g vs 3485 g: $p < 0.005$). Asian mothers with positive GCT and normal GTT were shorter and weighed less at booking

Multivariate analysis (no details of type of analysis – seems to have been done separately for each ethnic group) of birth weight with gestational age, maternal height, weight, and 2-h GTT glucose showed that the 2-h GTT was a significant predictor of birth weight in Asian mothers ($p < 0.0005$) and maternal height and gestational age were significant predictors of birth weight in Europid women ($p < 0.0001$)

Authors' conclusions: Ethnic influences are important when defining LGA infants. Mild disturbances of maternal glycaemia have a greater influence on the birth weight of Asian than white/Europid infants

Comments

Small sample size limited statistical power (21 Asian and 26 white/Europid women with GDM). The findings of this small study of the influence of maternal glycaemia on infant weight differing according to ethnicity would need confirming by larger study

Assumes all cases of GDM were diagnosed by positive screen at 20 weeks followed by abnormal GTT (WHO criteria) at 24–28 weeks

Seems to have been delay in diagnosing women with abnormal screening

Some women with abnormal screening and non-diabetic GTT required insulin. How this need was detected was not stated

No comment was made on adequacy of control of GDM

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Fedele and Lapolla, 1997[232]	Italy Low middle class women, no further demographic details given Mean age of women not reported Known diabetic women excluded	Observational study, 490 women, two groups: RF (n = 234) and non-RF (n = 256). Defined RFs: FH, maternal macrosomia, history of diabetic tendency, age > 35 years, habitual abortions, history of toxaemia, macrosomia, unexplained stillbirths, congenital abnormalities, obesity, glycosuria, polyhydramnios, pyelonephritis or recurrent urinary tract infections	50-g non-fasting GCT (O'Sullivan and Mahan[19]) PG ≥ 140 mg/dl given 3-h 100-g fasting GTT, interpreted by C&C criteria[20] (not given). At-risk women tested at 10–14 gestational weeks, if negative repeated 24–28 weeks, and if negative again at 32–34 weeks. If positive GCT but negative GTT then just the GTT was repeated the next time. Non-RF women tested at 24–28 weeks and if negative repeated at 32–34 weeks	Overall incidence GDM 10.8%. Incidence in RF women (22.3%) significantly greater than non-RF women (9.3%), $p < 0.01$ Women at high risk had a high prevalence of GDM in the last testing period (11.3% first testing, 15% second and 32.6% last testing in RF. 12.5% first test and 7.3% second test in non-RF women) No difference between RFs in GDM and non-GDM women in the RF group (reported percentages for FH, BMI, age > 35, maternal macrosomia, fetal macrosomia or fetal malformations only) CS were significantly greater in GDM women than non-GDM women ($p < 0.05$)
Sensitivity and specificity No data given to calculate sensitivity or specificity Women with negative screen were given repeat screen but percentages reported and cannot assume that these are percentages of total or percentages of those repeated testing from data given				
Comments No note made whether maternal outcomes noted were from women diagnosed at which testing period No demonstration that early screening alters outcome Authors suggest that useful to retest women with negative GCT or GTT later in pregnancy, although this may not alter outcome				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Forest *et al.*, 1994[233]	Canada Population predominantly white French Canadian women from a low to middle socio-economic environment with free access to healthcare Mean ages in different groups: 29.8 (SD 4.7), 30 (SD 5.1) and 28.2 (SD 4.3)	Retrospective study of pregnancy outcome in women diagnosed GDM, women with only one abnormal value on GTT, and normal women 6380 women delivered, of these 4314 women were screened and classified into 4 groups, I normal screening or normal GTT (4183), II GDM (237), III one abnormal GTT treated as GDM (85) and IV one abnormal GTT untreated (69)	Between 24 and 32 weeks women screened with 1-h 50-g GCT. Those with glucose ≥ 7.8 mmol/l given diagnostic GTT. Also 183 women just given GTT (no explanation given) Diagnostic 3-h 100-g GTT given following 3-day high CHO diet and fasting overnight Abnormal values defined as fasting 5.0, 1-h 11.3, 2-h 10.6 and 3-h 8.9 mmol/l (95th percentile from authors' previous experience) 32 cases of GDM also diagnosed on basis of repeated fasting and postprandial BG (not stated why)	Incidence GDM 3.7% No difference in group characteristics such as initial weight, weight gain, pregnancy duration, or time when GTT performed Area under glycaemic curve was similar between groups III and IV and statistically less than group II (GDM), $p < 0.001$ GTT periodicity (referred to another reference for definition) was greatest in GDM group Maternal outcomes: GDM women increased incidence of delivery before 37 weeks, chronic hypertension, and CS than normal mothers ($p < 0.01$) No significant differences between one abnormal values and normals Neonatal complications: incidence macrosomia not significantly different, metabolic and respiratory disorders and malformations were significantly different between GDM and all other groups ($p < 0.01$)
Sensitivity and specificity Unable to calculate sensitivity or specificity as different screening methods used, i.e. some women given GCT, others just GTT and others not screened but diagnosed GDM on the basis of repeated fasting and postprandial BG measurements				
Comments Used a higher criteria for diagnosis of GDM than NDDG[18]. If NDDG used, some more women from groups III and IV would have been classed as GDM No differences in maternal or neonatal outcomes between women with one abnormal value who were treated as GDM and those untreated *Authors' conclusions*: One abnormal value on GTT has no measurable negative impact on perinatal outcomes (although small sample in these two groups)				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Grandjean *et al.*, 1980[175]	Toulouse, France Sample collected 1978–79 Sample was a mixture of pregnant women with traditional RFs for GDM (previous infant ≥ 4500 g or unexplained stillbirth, FH DM, fasting glycosuria, hydramnios, obesity) and those with FATD > 95th percentile on charts for local population as measured by ultrasound Exclusion criteria: women with known DM Characteristics of subjects not described	Not RCT Retrospective review *Aim*: to examine the relationship between GDM and FATD 113 women including 92 women with traditional RFs for GDM and 21 women with FATD > 95th percentile FATD measured by real-time scanner at right angles to the longitudinal axis level of umbilical vein	Screening test used: classical RFs for some women or excessive FATD for others Diagnostic test: OGTT with 1 g glucose/kg body weight after 32 weeks GDM diagnosed if VBG values ≥ 167 mg/100 ml (9.3 mmol/l) at 120 min and ≥ 8.0 mmol/l at 150 min GTT performed after 72 h of normal food intake followed by 10-h overnight fast GDM treated with restricted diet (1600 to 2000 calories) with 40% CHO	Selected population Overall 19% (21/113) had GDM GDM was significantly more common in those with FATD > 95th percentile than in those with FATD ≤ 95th percentile 36% (12/33) vs 11% (9/80): $p < 0.01$ Only 29% (6/21) of those with GDM had biparietal diameter > 95th percentile Birth weights of GDM infants were similar to that of general population (3460 g vs 3250 g at 38 weeks)
Sensitivity and specificity FATD > 95th percentile: sensitivity 57% (12/21); specificity 77% (71/92); PPV 36% (12/33) Only one infant from a mother with normal GTT died in first 6 days. Rest were healthy *Authors' conclusion*: Excessive FATD seems to be a useful addition to the classic criteria for GDM				
Comments Selected population. Characteristics of this population were not described No comment made on inter-observer variability in measuring FATD Screening using only FATD as reported would have missed about 40% of the cases of GDM Does not report on the validity of incorporating FATD into RFs for GDM Study does not justify using FATD as the sole screening test for GDM and does not report on the validity of adding FATD to the classical RFs for GDM				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Grant *et al.*, 1986[206]	Australia No demographic details of women given except median age range at diagnosis (30–34 with normal tests and 35–39 with impaired tolerance)	Retrospective study over 14 years (1971–84) 1144 women who had previously had GDM contacted in 1985 and asked to attend for a follow-up, with 39.1% response rate (447 women)	All women given a 75-g GTT (WHO criteria[24]) Other details collected on FH, subsequent obstetric or surgical history, symptoms of hyperglycaemia, and BMI	49 (11%) patients diabetic 35 (7.8%) patients IGT 14 of the diabetic patients were known to be diabetic at the time of the recall and already on appropriate therapy FH not a predictive factor Of 210 women who had had subsequent pregnancies, only 120 had received a GTT despite previous GDM Time of diagnosis: 51% diabetics identified within 4 years and 56% not identified until this review 73.5% of the diabetics were overweight, compared to 62.9% IGT and 44.7% normal In 63 women a postnatal GTT was performed, this was a predictor of later DM (no report of statistical testing of this)
Comments Different diagnostic thresholds were used over the period studied. When analysed separately the incidence of DM at follow-up was greater in those diagnosed in the early period when the threshold was higher No data on the remaining 697 women with GDM who did not respond				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Greenberg *et al.*, 1995[234]	San Diego, California, USA Inclusion criteria: women with GDM detected in pregnancy who were followed up for 6 weeks after pregnancy Sample ethnicity: Hispanic 68%; white 12%; African-American 9%; Asian 7%; Pacific Islander 3%; other 1% Mean age 30 years; multiparous 79% BMI 29	Not RCT Retrospective review of case notes *Aim*: to determine whether antepartum variables can predict post-partum glucose intolerance (51 antepartum variables analysed) 238 mother–infant pairs reviewed 94 (39%) completed post-partum GTT	GDM detected by screening using 1-h 50-g glucose load at 24–28 weeks' gestation/at first antenatal visit if had RFs for GDM. Threshold ≥ 140 mg/dl GDM diagnosed if GCT ≥ 200 mg/dl or using 3-h 100-g GTT with two or more values ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl (Second International Workshop-Conference on GDM[216]) Post-partum Diagnostic test: 75-g GTT performed 6 weeks post-partum for all those with GDM. NDDG definitions[18] used to classify women as normal, IGT or diabetic	At 6 weeks post-partum, using NDDG criteria[18]: 34% (32/94) with GDM had abnormal GT 18% (17/94) had IGT 16% (15/94) had DM The factors most individually predicative of glucose intolerance and included in regression model were: PH GDM; 1-h GCT ≥ 200 mg/dl; requirement for insulin in pregnancy; total insulin dose at delivery ≥ 100 U/day; and any 2-h post-prandial BG ≥ 150 mg/dl RR for persistent post-partum glucose intolerance at 6 weeks with all three factors (1-h GTT > 200 mg/dl, any postprandial BG ≥ 150 mg/dl, insulin required): 19.68 (95% CI, 2.88 to 134.42). All 3 factors present in 18 women
Sensitivity and specificity Sensitivity and specificity of factors in predicting glucose intolerance 6 weeks post-partum in women with GDM in pregnancy: All 3 factors (1-h GTT > 200 mg/dl, any postprandial BG ≥ 150 mg/dl, insulin required): sensitivity 52%; specificity 96%; PPV 89% 1-h GTT > 200 mg/dl: sensitivity 68%; specificity 71%; PPV 56% Any postprandial BG ≥ 150 mg/dl: sensitivity 64%; specificity 68%; PPV 56% Insulin required: sensitivity 97%; specificity 52%; PPV 51% Insulin ≥ 100 U/day required: sensitivity 19%; specificity 100%; PPV 100% Above as reported. Data not presented to allow checking *Authors' comments*: Sizeable proportion of patients had 'pregnancy only' medical insurance and follow-up GTT could not be scheduled; differences in racial profile may account for difference in rates of post-partum DM in this study and other studies; significant number of patients have medical insurance cover only for emergencies or pregnancy and could have had undiagnosed pre-pregnancy DM *Authors' conclusion*: Post-partum screening is not warranted for women at low risk who do not require insulin in pregnancy. Women with RFs should receive post-partum screening for DM. Patients receiving ≥ 100 U/day insulin have a 100% incidence of post-partum glucose intolerance			
Comments Population predominantly Hispanic. Results may only be applicable to this specific population High dropout rate from follow-up GTT (61%) Compared characteristics between those in study and non-returners Discussed potential reasons for high rates of non-return for follow-up GTT Large 95% CI intervals for RR indicate small numbers with factor present				

All other studies *contd*

Reference	Research question	Outcome
Gregory *et al.*, 1993[84] Clinical commentary	What are the estimated potential savings in healthcare costs that would result from a primary prevention programme for women with GDM? Economic model: Used data from the National Center for Health Statistics (USA) in which there were 4,179,000 births in 1990 If 3% had GDM then 125,370 cases would be expected Previously reported that 50% of women with GDM will develop non-insulin DM within 10 years. In this case, 62,685 women Model assumes that the rate of progression to DM is constant over the 10-year period (6.7%). The annual healthcare cost per patient was US$2265 in 1986, adjusted to US$2834 in 1990 Based on this model 8400 women will develop DM in first year, costing healthcare system approximately US$23.8 million. After 2 years, the cost translates into US$44 million during year 2. The cumulative net cost over 10 years is US$818 million Costs for education: Model assumes that education and counselling will occur at routine semi-annual contraceptive or yearly gynaecologic evaluations. Thus no new charges for physician services. Estimates for DM testing at this time, glucose determination US$2 and dietary consultation US$29. The ongoing cost for the entire cohort over 10 years is US$39.8 million	If 10% of the cases could be prevented or delayed for 10 years the savings would be US$71,757 million If the incidence was reduced by 25% approximately US$179 million would be saved over 10 years If the incidence was halved, to 3.35% a year, US$371 million would be saved 10-year savings minus cost of preventive therapy: 10% reduction, US$31,946; 25% US$139,369 and 50% US$331,369
Authors' comments: Savings may be overestimated as no cost for initial screening Rate of progression to DM may not be a linear 6.7% per year Cost-per-case estimate based on that of a prospective study Assumed incidence of GDM was 3% (does not reflect higher incidence in certain ethnic groups) Calculation based on a 5% discount rate for future dollars, which attempts to adjust for both inflation and rate of return had money been invested elsewhere A theoretical benefit may arise from the family of those treated as long-term changes may impact others		

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Gregory *et al.*, 1998[235]	UK Mainly Caucasian, no other details given Mean age 29.1 (SD 5.2) years	Audit over 1 year of screening programme 4016 women booked for delivery at centre, 3316 completed the screening protocol	At first antenatal visit (mean 13.5 weeks, SD 3.7) all women had a random VBG taken. Any value > 7.0 mmol/l given 75-g fasting GTT interpreted by WHO (1980) criteria[108]: GDM if fasting ≥ 7.8 mmol/l or 2-h ≥ 11.1 mmol/l Those with normal screening or GTT are screened again at 28 weeks (range 19–37 weeks) with a 1-h 50-g GCT with Lucozade Glucose ≥ 7.8 mmol/l given 100-g GTT as above	195 women did not have an RBG at first visit, of these 2 were referred for GTT due to individual clinicians choice 23 random glucose were > 7.0 mmol/l and 18 of these had GTT A further 42 GTT were ordered by clinician despite glucose levels < 7.0 mmol/l. 4 GDM cases from these 62 found 682 had no subsequent BG measurements (no reasons recorded in hospital data) Remaining 3278 had 50-g GCT and 223 positive. 42 had GDM 70 negative screens also sent for GTT and 8 GDM cases found Further 49 proceeded directly to GTT without GCT (clinician choice), 12 had GDM Overall incidence GDM = 1.6%
Sensitivity and specificity Women with GDM were significantly older (31.3, SD 5.3) than those without GDM (29.0, SD 5.2), $p < 0.001$ Unable to calculate sensitivity or specificity				
Comments The yield of GDM from RPG was low, however the time of testing was early *Authors' comment*: Obstetricians were unhappy to rely on 'normal' ranges				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Gribble *et al.*, 1995[133]	USA Mean age approximately 27 years No further descriptions of population characteristics Women with pre-existing DM, multiple gestation excluded Only those who had at least 2 urinalysis tests during first 2 trimesters included	Retrospective observational study No randomisation *Aim*: to assess ability of urinalysis for glucose to predict GDM or gestational outcomes, including persistent hypertension, uterine bleeding, fetal heart rate abnormality, CS delivery, shoulder dystocia, or birth weight percentile	Categorised into 2 groups, negative or positive glycosuria groups All women screened with 50-g GCT at 24–28 weeks (fasting status not stated) Threshold: 140 mg/dl Positive screenees started 3-day CHO load, and fasting 100-g GTT Threshold 2 or more ≥ fasting 105; 1-h 190; 2-h 165 and 3-h 145 mg/dl Negative screenees comparison of the 2 glycosuria groups in terms of outcomes	Incidence GDM 3.1% 47/2745 women had positive glycosuria – 6 of these had GDM (12.8%) 79 women with GDM in negative glycosuria group (2.9%). These two percentages significantly different from one another (p = 0.002 Women in positive glycosuria group with GDM more likely to require insulin for control (67% vs 19%, p = 0.02) In terms of outcomes, the only difference between positive glycosuria and negative was that the former had lower rate of hypertension and higher mean birth weight ratio
Sensitivity and specificity Sensitivity of glycosuria as a predictor of GDM 7% (6/85); specificity 98% (2619/2660); PPV 12.7% (6/47)				
Comments 3217 potential candidates, 252 excluded due to inclusion/exclusion criteria. Analysis on 2745 patients, 220 patients lost without being accounted for No cost data Authors suggest early detection glycosuria can allow diagnosis by blood screening earlier than 24–28 weeks, but only small number of women with glycosuria had GDM				

Reference	Type of study
Hadden, 2000[236]	Not an original study Appraisal of paper by Perucchini *et al.*, 1999[104] (compares ease of screening for GDM using an FPG and 50-g GCT) Hadden comments on this paper and points out that: 1. The paper did not study the outcome of the pregnancy 2. Not fully population-based as over 25% of the eligible population did not take part Considered that the question that should be asked is whether either measurement has an effect on the outcome of the pregnancy either in the short term for the baby or in the longer term for the mother Advises that although the paper shows that a true fasting glucose is more sensitive than a random glucose in predicting glucose intolerance and that it may be that the authors' threshold value of 4.8 mmol/l between the 24th and 28th weeks would be a more useful value in predicting the fetal results that we must await better studies which include at least the short-term outcome for the baby before making further decisions The Hyperglycaemia and Adverse Pregnancy Outcomes study is currently funded largely by the US NIH, the ADA and the BDA. Study has enrolled 25,000 pregnant women in 16 different centres around the world and will undertake standard 75-g GTT at 28 weeks in all women. Outcomes include: maternal fasting and 2-h post-glucose values, maternal insulin response to glucose load, and the following infant outcomes: umbilical venous and early neonatal CBG values and anthropomorphic measure of skin thickness. DNA analysis of mother and baby. Many neonatal and maternal markers Hadden considers there is a need to know how to manage what is likely to be a continuously rising risk at levels of glucose well below those presently defined as DM or even IGT. The question is when to start insulin

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Hanson and Kallner, 1984[111]	Stockholm, Sweden Inclusion criteria: Women with traditional RFs/RFs developing in current pregnancy Characteristics not described	Not RCT Observational *Aim*: to compare the discriminatory power of a 2-h GTT with that of a 3-h GTT using the Gillmer and Bickerton method[230] 182 GTT in 168 patients	Screening test used: screened using traditional RFs. GTT performed in early pregnancy if the following RFs present: FH DM, previous baby > 4500 g, extreme obesity, previous unexplained fetal death, repeated abortions/ premature labour If initial GTT negative, GTT repeated at 30 weeks GTT performed as soon as possible if glycosuria 2+, polyhydramnios, accelerated fetal growth, other indications Diagnostic test: 50-g GTT after fasting 3 days of CHO-rich diet. VPG tested fasting and after 30, 60, 90, 120, 150 and 180 min GDM diagnosed if area under glucose curve > 44 units for 3-h test (Gillmer and Bickerton, 1994[230])	13% (22/168) high-risk patients were diagnosed as GDM (area > 44 units on the 3-h test) Linear correlation between 2-h and 3-h area: $r = 0.97$, $p < 0.0005$ Cut-off point in the 2-h test was chosen to have a similar diagnostic specificity and sensitivity as the 3-h test: cut-off = 34 units for 2-h 2 patients judged normal by 2-h test who would be GDM by 3-h test (one had normal infant, the other had biphasic glucose curve and was treated from 33 weeks, delivered normal infant). 1 patient judged normal by 3-h test was diagnosed GDM by 2-h test (no treatment, healthy child)

Sensitivity and specificity

Data not available to calculate sensitivity and specificity for 2-h test

Agreement of classification = 98% (165/168)

Relationship between CBG and PG ($n = 160$): PG = 1.033 capill + 0.17 ($r = 0.99$). Plasma cut-off = 34 units is equivalent to CBG of 32 units

Authors' comment: In a number of patients a second peak in the PG is obtained; this may be missed by a shorter test

Authors' conclusions: The discriminatory power of the 2-h test is comparable to the 3-h test with VPG values

Comments

Sample had either historical/current RFs. Other characteristics of study sample were not described

Diagnostic GTT performed on two occasions to exclude GDM, although assumes all cases of GDM detected by abnormal GTT at 30 weeks

Linear correlations were given but no evidence was presented illustrating that the relationship between variables tested was linear

Small number of patients diagnosed as GDM (22 cases). With only 3 patients differing in classification by 2-h and 3-h tests

Costs of 2-h and 3-h tests not reported

The authors' conclusions are supported by the evidence

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Harlass *et al.*, 1991[105]	Army Medical Centre, Tacoma, Washington, USA Included: those scheduled to undergo diagnostic GTT after abnormal screening test Excluded: those unable to tolerate glucose solution Mean age: 27 years Nulliparous: 34% Mainly dependants of Army-enlisted personnel. Mean income US$11,000 to US$16,000. Tertiary centre that takes referrals from other military installations Ethnicity: white 72%; black 16%; Hispanic 5%; Oriental 5%; other 2% Gestation: first trimester 19%; second 66%; third 16%	Not RCT *Aim*: to evaluate the reproducibility of the 100-g GTT 64 women enrolled	Screening test: all undergo 1-h 50-g oral GCT at 24–28 weeks after fasting with diet > 150 g CHO for previous 3 days. Those with RFs (age > 25 years, obesity, FH of DM, previous stillbirth or macrosomia) also undergo screening at initial visit (Second International Workshop-Conference on GDM[216]) Abnormal GCT if glucose ≥ 135 mg/dl Diagnostic test: 100-g OGTT. GDM diagnosed when 2 or more values ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl Second OGTT performed under identical conditions in approximately 1–2 weeks	GDM rate not reported Overall 50 of the 64 (78%) GTTs were reproducible Women classified as: Group 1: normal and normal (48 women = 75%) Group 2: normal and abnormal (11 women = 17%) Group 3: abnormal and normal (3 women = 5%) Group 4: abnormal and abnormal (2 women = 3%) No statistically significant differences in mean GTT values over whole population between 2 testing periods or reproducibility between trimesters (1: 100%; 2: 74%; 3: 70%) Significant differences between group 1 (normal, normal) and group 2 (normal, abnormal) at fasting (p = 0.02) and 1-h (p = 0.003) and 2-h (p = 0.02)
Sensitivity and specificity N/A *Authors' conclusion*: Overall the reproducibility of the 100-g GTT was 78%. We recommend the GTT be repeated when the 1-h value is abnormal or when the fasting blood sugar, 1-h or 2-h values are near the upper end of the normal range				
Comments Small sample size (64 women) limited power to detect statistically significant differences No evaluation of costs of recommended strategy				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Hatem *et al.*, 1988[62]	UK study No demographic details given Age of women 26.0 (19–34) years Excluded if diabetic, FH DM, previous baby > 4000 g, stillbirth or infant with congenital malformations, neonatal death, weight > 85 kg, those over 35 years, or with a multiple pregnancy	No randomisation No description of selection process 215 women recruited, 3 excluded due to vomiting *Aim*: to determine reference values for the 75-g GTT during pregnancy	All women given 75-g GTT after an overnight fast Fasting, 1- and 2-h BG levels taken Subjects divided into 3 groups depending on time of gestation at testing: Group 1 (n = 43) 14–20 weeks; group 2 (n = 125) 28–32 weeks; group 3 (n = 44) 36–37 weeks	PG at 1 and 2 h not normally distributed thus transformed logarithmically and expressed as geometric means and 2.5 and 97.5 centile limits The fasting glucose did not differ significantly between the three groups Groups 2 and 3 did not differ significantly on the 1 or 2 h glucose and the data were combined. Significant differences between 1 and 2 h glucose were demonstrated between this combined group and group 1 (both p's < 0.001) No congenital malformations observed, 3 babies less than 2.5 centile and one greater than the 97.5 centile. One women shown to be diabetic (fasting 5.9, 1-h 12.8, 2-h 13.0 mmol/l) and her results excluded
Sensitivity and specificity No screening test used				
Comments No cost data available Noted the difference in mean PG between first and second/third trimester Authors suggest that in similar populations, the upper limits of normal for PG 2 h after a 75-g GTT are 7.5 and 9.6 mmol/l for the first and second trimester respectively				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Helton *et al.*, 1997[117]	North Carolina, USA Setting: university community with diverse socio-economic population attending a family practice centre of large southern academic medical centre Inclusion criteria: all women who enrolled before 26 weeks' gestation, carried pregnancy ≥ 28 weeks, gave birth between 1/1988 and 12/1993 and had no personal history of DM (considered to be low risk of GDM) Mean age 27 years (range 15–46 years) Medicaid 40%; private insurance/managed care 60% Ethnicity: non-Hispanic white 60%; black 35%; other 5%	Not RCT Retrospective review of case records *Aim*: to determine the rate of abnormal GCT, number of true-positive GCT, and whether women with GDM could have been selected for screening on historical RFs 595 women were eligible 526 (88%) were used in analysis (dropouts failed to complete protocol)	Screening test used: universal screening with 1-h 50-g GCT at 24–28 weeks. Non-fasting. Women sedentary during test Abnormal GCT if 1-h PG > 140 mg/dl Charts reviewed for RFs including: obesity, FH DM, previous poor pregnancy outcome, previous macrosomic infant Diagnostic test: 3-h 100-g GTT. Prior fasting, remained sedentary. GDM diagnosed according to NDDG criteria[18] as two or more values >: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl Women with GDM received dietary advice. None received insulin	2.5% (13/526) diagnosed as GDM (NDDG criteria) 31% (4/13) of women with GDM had no RFs 1.1% (4/353) women with no RFs were diagnosed as GDM 5.2% (9/173) women with one or more RFs were diagnosed as GDM 14% (76/544) had positive GCT Rates of infants > 4000 g were as follows: 15% (2/13) in women with GDM 17% (3/18) in women with positive GCT who did not undergo diagnostic GTT 4% (2/51) in women who were not screened 15% (69/468) in women with negative GCT. 2 mothers in this group had unexplained intra-uterine death

Sensitivity and specificity

1-h GCT > 140 mg/dl in predicting GDM (NDDG criteria): sensitivity 100% (13/13); specificity 91% (468/513); PPV 22% (13/58)
Screening using traditional RFs in predicting GDM as defined: sensitivity 69% (9/13); specificity 68% (349/513); PPV 5% (9/173)
Assumes GCT > 140 mg/dl followed by abnormal GTT test to be 100% sensitive in detecting GDM
Authors' comments: The lack of screening of 69 women may have falsely lowered the rate of GDM
Authors' conclusions: Screening with a 1-h 50-g GCT only those women who have identifiable RFs for GDM is a reasonable approach to identifying the disease in a low-risk population

Comments

Assumes all cases of GDM were detected by GCT value > 140 mg/dl followed by abnormal GTT at whatever time in pregnancy this was performed
Population reported to be at low risk for GDM
Small number of women diagnosed as GDM (13 women) limiting power to detect differences in rates of infants > 4000 g in various groups
24% (18/76) of women with positive GCT did not undergo diagnostic GTT though birth weight was reported for this subgroup
Full details of management of GDM women were not described
Screening only those with traditional RFs as recommended by the authors would have missed 31% (4/13) cases of GDM in this study
No costs were reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Hong *et al.*, 1989[103]	USA No mean age given, women subdivided into four groups: < 24 years, ≥ 24 years, ≥ 25 years and ≥ 30 years to allow comparisons to be made 61% black, 23% Hispanic, 10% Indian, 5% Caucasian, 1% Oriental	Prospective observational study, no controls, no randomisation 999 women tested at first visit regardless of gestational age and then classified into groups: first 13 weeks (n = 228), 14–23 weeks (n = 354), 24–28 weeks (n = 122) and > 28 weeks (n = 195)	Screened with 1-h 50-g GCT. Fasting status not defined. Those ≥ 130 mg/dl given 3-h fasting 100-g GTT Diagnosis compared by NDDG[18] and C&C[20] criteria Values not given	Incidence GDM with NDDG 3.9% Incidence with C&C 4.4% With a screening value of 130–139 mg/dl 2 GDM cases with NDDG and 5 with C&C All patients (16) with screening value > 200 mg/dl had GDM Fewer women under 24 years had GDM compared to those over 24 years (5 vs 32, but report does not state if tested for significance) Number of GDM in each gestation age shows that 2/3 were diagnosed after 23 weeks, and most diagnosed > 28 weeks, again no statistics

Sensitivity and specificity
No data to calculate sensitivity or specificity
Mean screening value was significantly lower in the younger group < 24 years, than means of the older groups (p < 0.05)

Comments
No patients who were screened negative were given GTT and of those screened early on, there is no report that they were re-screened later
Cost data per case of GDM given: (based on total cost of GCT (US$3.50) and total cost of GTT (US$15.00))
If universal screening, and cut-off 130 mg/dl, US$184; cut-off 140 mg/dl, US$158
If selective screening of women ≥ 25 years, and cut-off 130 mg/dl, US$122; cut-off 140 mg/dl, US$106
If selective screening of women ≥ 30 years, and cut-off 130 mg/dl, US$120, cut-off 140 mg/dl, US$102

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Hooper, 1996[134]	Atlanta, Georgia, USA Sample of records from women attending antenatal clinic. None had known pre-existing renal disease Exclusion criteria: pre-existing DM	Not RCT Retrospective review of 610 case records *Aim*: to assess the sensitivity of routine urine analysis in detecting pre-eclampsia and DM	Screening test used: urine testing for glucose and protein (≥ 30 mg/dl) using Ames Bili-Labstix. Plus a 1-h GCT with 50 g glucose load between 24 and 28 weeks If GCT > 135 mg/dl, undergo GTT Diagnostic tests: 3-h 100-g GTT GDM and pre-eclampsia (PET) diagnosed according to ACOG criteria[4]	1.8% (11/610) diagnosed as GDM (ACOG criteria) 2.5% (15/601) had glycosuria on at least one occasion. Only 36% (4/11) of those with GDM had glycosuria 18% (109/610) had proteinuria (≥ 30 mg/dl) on at least one occasion. Only 18% (3/18) of those with proteinuria had proteinuria preceding hypertension

Sensitivity and specificity
Table 1 evaluated the use of glycosuria in predicting GCT > 135 mg/dl. Results for GDM prediction calculated by present authors
Glycosuria in predicting GDM: sensitivity 36% (4/11); specificity 98% (588/599); PPV 27% (4/15)
Glycosuria in predicting GCT > 135 mg/dl: sensitivity 46% (6/13); specificity 98% (588/597); PPV 40% (6/15)
Proteinuria in predicting PET: sensitivity 70.6% (12/17); specificity 83.6% (496/593); PPV 11% (12/109)
Authors' conclusion: The results support the use of routine 50-g GCT and careful monitoring of blood pressure in detection of GDM and PET rather than routine urine analysis

Comments
Characteristics of population were not described
Assumes all cases of GDM were detected with 1-h GCT > 135 mg/dl by 28 weeks followed by diagnostic GTT. The authors do mention this assumption
No fetal/maternal outcomes assessed according to GDM screening methods
Mentions that the cost of urine testing is greater that the relatively low cost of using dipsticks but no actual costs were reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Hughes *et al.*, 1995[237]	Al Ain, United Arab Emirates (UAE) Sample from routine antenatal clinic in highly pro-natal society where early serial parity is the norm Multi-ethnic population: 33% Indian subcontinent; Chami Arabs 19%; UAE 19%; North African Arabs 10%; Asian Arabs 9%; East African Arabs 7%; others 4% Excluded women delivering infants < 1000 g in weight	Not RCT Cohort *Aim*: to demonstrate and quantify the influence of levels of CHO intolerance on fetal and maternal morbidity 1722 women eligible. Results from 1401 (81%) women analysed. Reasons given for drop-outs	Screening test used: universal screening with venous glucose 1-h post-50 g glucose load. Non-fasting. Performed between 24 and 28 weeks. GCT abnormal if BG ≥ 140 mg/dl Diagnostic test for those with abnormal GCT: 3-h 100-g GTT GDM diagnosed according to C&C criteria[20]: two or more levels ≥: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl GDM managed by diet and/or insulin Screen-positive women given dietary advice	Not possible to calculate GDM rates. Only 35% (97/277) with positive screen underwent GTT Overall 19.8% (277/1401) were screen-positive. Rates varied with ethnicity Admission to SCBU was more common among screen-positive women than screen-negative women (15.2% vs 8.7%; $p = 0.002$) Birth weight > 3999 g was significantly more common in screen-positive women than screen-negative women (30.5% vs 11.7%; $p < 0.001$) 19.6% (19/97) of those with positive screens were diagnosed with GDM (C&C) Admission to SCBU was significantly more common among GTT-positive women than GTT-negative women (36.8% vs 10.3%; $p = 0.01$) Birth weight > 3999 g was not significantly different in GTT-positive women vs GTT-negative women (31.6% vs 25.6%)
Sensitivity and specificity Screen-positive rates (GCT ≥ 140 mg/dl) according to ethnicity: 'high rates' East African Arabs 23%; Indian subcontinent: 22%; UAE national 20%; Chami Arabs 20% 'Low rates' North African Arabs 16%; Asian Arabs 13% Screen-positive women were significantly older than screen-negative women (29.4 years vs 27.2 years) Babies of screen-positive women were more likely to be admitted to SCBU: OR 1.87 (95% CI, 1.25 to 2.81) The number of babies with birth weight > 3999 g was significantly greater in screen-positive women: OR 1.99 (95% CI, 1.39 to 2.84) For subgroup of 97 screen-positive women having GTT Small numbers of women with GDM prevented ethnic analysis on GDM rates Babies of GDM women were more likely to be admitted to SCBU: OR 5.10 (95% CI, 1.35 to 19.61) There was no difference in number of babies > 3999 g between GDM and non-GDM women. OR 1.34 (95% CI, 0.39 to 4.48) *Authors' comments*: Adjusting screening threshold to select the same percentage of diagnostic cases per ethnic group is in danger of ignoring the primary indication for screening, which is to detect the subset of patients with an increased risk of fetal/maternal morbidity and mortality				
Comments Authors acknowledge the limitations of the study and discuss the aims of screening for GDM Authors do not consider the rate of GDM reported to be a valid estimate of the sample population due to the poor compliance rate with diagnostic testing following a positive screening result. Only 35% (97/277) who were screen-positive underwent diagnostic GTT. Characteristics of screen-positive women who underwent GTT were not compared with those not undergoing GTT. Sample analysed may not be representative of local population Assumes all GDM detected by abnormal GCT at 24–28 weeks. May have underestimated overall positive screen and GDM rates by failing to detect women who develop CHO intolerance later in pregnancy No investigation of the influence of potential confounding factors on OR *Authors' conclusions*: Clinical attention should be directed to outcome assessment in order to properly evaluate the nature and place of screening for GDM in specific populations				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Jovanovic and Peterson, 1985[156]	NYC, USA 80% white, 13% black, 5–6% Hispanic Ages ranged from 20 to 40 years; 60% of patients were aged 25–29 years and 25% between 30 and 34 years	*Aim*: to determine optimal timing and criteria for re-testing. Also ability of screening test to predict birth weight of > 4000 g The first group of patients (n = 300) were screened at 3 intervals: 9–20 weeks, 27–31 weeks and 33–36 weeks A further group (n = 300) were screened at 2 intervals: 27–31 weeks and 33–36 weeks	Patients instructed to fast for 4 h before appointment Screening was performed with the 1-h 50-g GCT with an abnormal rating classified as ≥ 150 mg/dl Patients screening positive were given a 100-g GTT (O'Sullivan and Mahan[19]) within 2 days Patients were either confirmed with GDM or completed the testing schedule	Prevalence (defined by screening criteria) was 3.2% (19/600) 1 patient out of 300 (0.03%) had a positive GTT between 9 and 20 weeks. 12 of the remaining 599 (2%) had a positive GTT when tested between 27 and 31 weeks. 6 of the remaining 587 (1%) had a positive GTT when tested between 33 and 36 weeks Had RFs been used instead of GCT, only 1 of the 19 women with GDM in this sample would have been identified A retrospective study of the women who delivered babies > 4000 g found that 55 such babies were born. 48 of those had an abnormal GCT (no information is given to match with positive GTTs) Authors recommend screening at 27–31 weeks to detect the majority of GDM supplemented by screening at 33–36 weeks for: those with a positive GCT at 27–31 weeks, age ≥ 33 years, and obesity (defined as > 120% ideal body weight). Also found this to be correct timing in terms of maximising PPV of large (> 4000 g) babies

Sensitivity and specificity
50-g GCT, positive ≥ 150 mg/dl
For a positive diagnosis of GDM
9–20 weeks: PPV 11.1% (1/9); sensitivity 5% (1/19) assuming that GCT is 100% sensitive and that all women with GDM are detected by later screens
27–31 weeks: qualified PPV* 6.9% (12/173); sensitivity 67% (12/18) assuming that GCT is 100% sensitive and that all women with GDM are detected by later screen
33–36 weeks: qualified PPV* 3.7% (6/161); assumed sensitivity 100% (6/6)
* Not true PPV; these results are both qualified by fact that some women with GDM were identified earlier, which may have skewed results. Could be viewed as "PPV of what's left"
For predicting those who will give birth to > 4000 g infants
9–20 weeks: PPV 44.4% (4/9); sensitivity 12.5% (4/32); specificity 98% (263/268)
27–31 weeks: PPV 27.7% (48/173); sensitivity 87.3% (48/55); specificity 77% (419/544)
33–36 weeks: PPV 22.3% (36/161); sensitivity 65% (36/55); specificity 76.5% (407/532)

Comments
The authors' calculation of specificities of the GCT (in terms of predicting > 4000 g babies) are seriously flawed (see above for correct figures)
Only positive screens given diagnostic test so prevalence may be higher if some GDM went undetected. Assumes GCT 100% sensitive in detection of GDM
Sensitivity and PPV figures (in terms of prediction of GDM) are skewed (see opposite)
Only women not found to have GDM screened at each point. Latest screening point may have had highest detection rate if GDM women were not previously diagnosed
No costs given

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Jowett *et al.*, 1987[140]	UK Pre-selected women referred due to heavy glycosuria, polyhydramnios, poor obstetric history, large baby or FH DM Mean age 25 (SD 4.1) range 17–38	Not randomised No controls *Aim*: to assess sensitivity of RBG screening for detecting GDM 110 consecutive women	Tested between 27 and 31 weeks Patients admitted to hospital for five venous glucose levels: 0800, 1200, 1500, 1700 and 2200 hours. Fed at normal meal times and snacks in between All then had 75-g GTT (fasting status not stated but states 'as detailed in WHO protocol') Diagnosed normal (55%) if 2-h < 8.0 mmol/l, impaired (34%) if 2-h > 8 and < 11.0 mmol/l and diabetic (11%) if 2-h > 11.0 mmol/l Assessed sensitivity if BG > 6.0 mmol/l within 2 h of eating or > 5.6 mmol/l if more than 2 h	Incidence GDM 11% Incidence IGT 34% No adverse effects noted
Sensitivity and specificity Different sensitivity depending on time of test: (NB: authors' calculation based upon incorrect formula and based upon cases of IGT together with GDM cases) Sensitivities using correct formula: If threshold > 5.6 mmol/l: 0800 sensitivity 66%; specificity 94%; 1200: sensitivity 50%; specificity 86%; 1500: sensitivity 58%; specificity 67%; 1700 sensitivity 58%; specificity 84%; 2200 sensitivity 50%; specificity 72% If threshold > 6.0 mmol/l: 0800: sensitivity 58%; specificity 96%; 1200: sensitivity 41%; specificity 94%; 1500: sensitivity 50%; specificity 74%; 1700: sensitivity 41%; specificity 86%; 2200: sensitivity 41%; specificity 87%				
Comments Women pre-selected due to RFs *Authors' conclusion*: RBG not suitable screening test No cost data available				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Kerbel *et al.*, 1997[163]	Toronto, Canada Mean age 30, many high education and income level. Majority had English as first language and born in North America but some Chinese origin 897 women tested at 2 intervals of whom 809 tested at 3 intervals History of GDM or DM and multiple pregnancies excluded	Unselected, non-randomised women undergoing universal screening *Aim*: to assess the psychological effects of screening. Categorised into 5 groups based on the woman's self-report following screening: false-positives; true-positives; perceived test negatives; uncertain; not tested	Universal screening with 50-g GCT between 24 and 28 weeks followed by 100-g GTT if ≥ 7.8 mmol/l Fasting status not defined Given state-trait anxiety inventory, centre of epidemiological studies depression scale, mothers perception of health and perception of health of newborn question at 12–24, 32 and 36 weeks	1.6% incidence (true-positives not analysed further) 9.8% false-positives; 55.1% perceived test negative; 25.8% not tested; 7.8% uncertain (not in analysis) No differences in anxiety, depression or concern about the newborn scores between false-positive and perceived negative groups At 32 weeks, half as many false-positives rated health as excellent compared to test negative or not tested group (p = 0.0001) The proportion of women perceiving their health to be excellent decreased over time in false-positive group No differences in those rating health as very good, good, fair or poor between groups or over time
Sensitivity and specificity N/A				
Comments False-positives outnumbered true-positives approximately 6:1. These false-positive women have a decreased perception of their own health Used patient's perception of test outcome, not always confirmed on checking with medical records No validation of questions regarding maternal or newborn's health. Anxiety and depression scales are validated scales and no differences noted here Authors suggest taking these adverse effects into account when deciding about a policy of screening of all (low-risk) pregnant women				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Khine *et al.*, 1999[95]	New Haven, Connecticut, USA Inclusion criteria: gravidas < 19 years of age at time of pre-natal registration Ethnicity of locale: white 70%; black 18%; Hispanic 2%; other 2%	Not RCT Retrospective review of 632 case records *Aim*: to determine the prevalence of GDM in adolescents and identify RFs 352 adolescent gravidas without GDM used as controls. Authors considered this group comparable but no baseline comparison was reported	Screening tests used: mostly universal screening with 1-h glucose post 50-g GCT between 24 and 28 weeks. No prior fasting. Positive if VPG ≥ 130 mg/dl (C&C[20]) Diagnostic test: 3-h 100-g GTT GDM diagnosed if two or more values ≥: fasting 95; 1-h 180; 2-h; 3-h 140 mg/dl (C&C criteria) GDM managed with 30–35 kcal/kg/day diet Insulin started when FBG persistently > 105 mg/dl or 2-h postprandial > 120 mg/dl	1.7% (11/632) of those aged < 19 years diagnosed as GDM Compared with GDM rates in other age groups: 19–24 years 3.4%; 25–35 years 4.9%; > 35 years 7.4%; overall 4.8% No difference was detected between GDM and non-GDM adolescents in race, FH DM, or presence of medical disorders. Adolescents with GDM more commonly had BMI > 27 than non-GDM (82% vs 16%: p < 0.001) Reports outcome of 11 GDM pregnancies (4 > 4000 kg; 2 mild shoulder dystocia)
Sensitivity and specificity N/A *Authors' recommendation*: That selective screening of adolescents be performed on the basis of BMI				
Comments Small number of cases of GDM in adolescent (11 cases) limited power of the analysis Assumes all cases of GDM detected by GCT ≥ 130 mg/dl followed by GTT by 28 weeks Sensitivity and specificity of using BM > 27 as a screening criteria for GDM in adolescents was not reported				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Kirkpatrick *et al.*, 1988[154]	Brussels, Belgium Inclusion criteria: consecutive women without historical/clinical RFs attending prenatal clinic Excluded: pre-existing DM	Not RCT Prospective cohort *Aim*: to test the feasibility of a screening programme for GDM and evaluate the influence of timing of glucose ingestion on BG variations Total of 1666 women screened (155 with historical/clinical RFs and 1511 without RFs) 1511 included in the analysis	Screening tests used: Those without RFs had 50-g GCT at first visit without prior fasting Test done between 8 am and 4 pm Timing of tests varied from 5 to 42 weeks Abnormal if VBG ≥ 135 mg/dl. If normal was repeated in third trimester Those with clinical/historical RFs had GTT Diagnostic test: 2-h 50-g GTT after overnight fast in first trimester Abnormal if one or more BG values ≥: fasting 105; 1-h 165; 2-h 150 mg/dl. If normal was repeated in third trimester	Overall 4.5% (73/1666) diagnosed as GDM (see criteria in adjacent column) 15% (25/155) with RFs had GDM 14% (226/1511) without RFs had GDM Mean PG in GCT was significantly higher in: third compared to first trimester (113 vs 96 mg/dl: $p < 0.001$) after 11 am compared to before 11 am (107 vs 99 mg/dl: $p < 0.001$) both in first and third trimesters women with body weight ≥ 150% of ideal body weight at start of pregnancy than those with less/no weight excess (124 vs 104 mg/dl: $p < 0.001$) More women over 35 years had BG > 135 mg/dl compared to younger (44% vs 19%; $p < 0.005$)
Sensitivity and specificity N/A *Authors' comments*: Some patients were misclassified because obstetricians had omitted consideration of one of the RFs Used expensive flavoured test solution to improve acceptability *Authors' conclusions*: The variations in glucose tolerance related to time of day, gestational stage and body weight are of limited amplitude and should not be considered in the determination of the cut-off point of the screening test				
Comments Excluded women with clinical/historical RFs Characteristics of sample were not described Screening performed twice (first and third trimester) Assumes all cases of GDM in those without RFs were detected by GCT ≥ 135 mg/dl followed by abnormal GTT It was not reported whether the time of day/gestational stage influenced the classification of any patients as GCT normal/abnormal and whether this classification was associated with fetal/maternal outcomes Costs (direct costs included PG determination, test solutions and phlebotomist's salary) Total of 2493 GCT and 716 GTT were performed Total cost for 1666 patients US$18,913 Cost per patient screened US$11.24 Cost per case GDM detected US$256.3				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Kjos *et al.*, 1995[207]	USA Observation of incidence of DM in Latino women only Ages at diagnosis of GDM not given	3300 women with GDM by NDDG criteria[18] between 1987 and 1993 684 women not diagnosed DM at post-partum testing returned for testing, 671 from Mexico or Central America were used Variables: parity, previous GDM, previous infant with anomalies, stillbirth or macrosomia, FH DM, gestational age at screen and test, BMI, hypertension tested as predictors for DM. As were glycaemia, fasting and serum lipid concentrations	Those with fasting glycaemia > 7.8 mmol/l prior to discharge were advised to return for a 75-g GTT 4–16 weeks later Those attending who were not diagnosed DM at that time were advised to return annually for 75-g GTT following overnight fast DM was defined by NDDG criteria	146 women developed DM during the follow-up. The cumulative incidence rates were 47% after 5 years and 55% after 6 years The only variables with significantly predictive values were: the area under the glucose curve of the GTT at 4–16 weeks, the gestational age at time of diagnosis of GDM, the area under the antepartum glucose tolerance curve, and the highest fasting glucose concentration during pregnancy
Comments No data available on those not returned for re-testing No ages of women given				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Knopp *et al.*, 1992[238] Same subjects as Magee *et al.*, 1993[94]	Seattle, Washington, USA Inclusion criteria: pre-natal registrants at 2 hospitals of a Health Maintenance Cooperative between 1985 and 1986 Excluded: women with GDM who required insulin 53% of those eligible consented	Not RCT Cohort study *Aim*: to evaluate the potential value of GHb, GPro, IRI and TG in addition to glucose for GDM screening and infant macrosomia (GHb, glycosylated Hb; Gpro, glycosylated protein; IRI, immuno-reactive insulin; TG, triglyceride) 2019 were screened The following were included in the analyses: 521 negative screenees selected randomly and 365 women with positive GCT	Screening tests used: 1-h PG after 50 g glucose load between 24 and 32 weeks after overnight fast. GCT positive if PG ≥ 7.77 mmol/l Diagnostic tests: 3-h 100-g GTT GDM diagnosed according to C&C criteria[20]: fasting 5.3; 1-h 10.1; 2-h 8.7; 3-h 4.8 mmol/l Patients classified as negative screenees; positive screenees with negative GTT; and GDM	4.6% (96/2019) were diagnosed as GDM (C&C) TG was the only value consistently significantly associated with birth weight ratio
Sensitivity and specificity N/A Relationship between birth weight ratio and TG: evaluated using univariate correlations and multivariate linear regression: TG was statistically significant (p = 0.019) after adjusting for maternal weight gain and BMI *Authors' comments*: Relationship between birth weight ratio and TG in the GDM subjects is variable and not statistically different from horizontal NB: Relationship is non-linear for TG > 2.4 mm and there is considerable scatter *Authors' conclusion*: Further evaluation of plasma TG in GDM screening is justified				
Comments Sensitivity and specificity of glucose plus TG in predicting GDM was not reported Relationship between TG and birth weight ratio was not linear, though some analyses were apparently based on this assumption Assumes all cases of GDM detected by GCT ≥ 7.77 mm followed by abnormal GTT by 32 weeks Cost-effectiveness of including TG with glucose in GDM screening was not evaluated There does not appear to be convincing evidence supporting the authors' conclusion				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Kvetny *et al.*, 1999[109]	Denmark Selected women from all pregnancies with RFs: previous GDM, fetal macrosomia (> 4500 g), glycosuria, BMI > 29, FH, prior stillbirth No demographic detail given	Prospective descriptive study No randomisation, no control 6158 women of whom 1220 were included due to RFs 20 of these refused Mean age 29 years (21–39)	Tested at 24–28 weeks with 2-h 75-g fasting GTT. No screening test Those with 2-h value ≥ 6.7 mmol/l diagnosed GDM Also compared numbers diagnosed with WHO criteria[24] ≥ 6.9 mmol/l Repeated GTT 3 months and 1 year after delivery	220 pregnancies (3.6% of 6158 women) had GDM with their criteria (2.8% with WHO criteria) GDM management with monthly follow-up from dietician. Diet controlled, unless BG > 7.2 mmol/l then insulin given Increased frequency malformations and pre-eclampsia (both significant) and tendency for increased CS in GDM when compared to those not tested. No difference macrosomia GTT still abnormal in 54% 3 months after delivery and in 33% after 1 year
Sensitivity and specificity No screening test used				
Comments Assumes that all GDM cases detected on grounds of RFs Authors conclude that women with GDM had a higher frequency of malformations and pre-eclampsia than women with normal glucose tolerance				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lamar *et al.*, 1999[169]	USA 72% white and 27% Hispanic or African-American Mean age 26 (SD 5.3) years Women with DM were excluded	Prospective study *Aim*: to assess the use of jelly beans as an alternative to a standard 50-g GCT All women randomised to study group (jelly beans vs standard glucose), and on second visit given the alternative source 160 women Women asked for preference for both glucose sources	Between 24 and 28 weeks women given either 50-g GCT with standard glucose or with 50 g glucose equivalent in jelly beans Within 1 week given the alternative source A BG ≥ 140 mg/dl considered abnormal All women went on to have 3-h 100-g fasting GTT, thresholds: any two, fasting ≥ 105, 1-h ≥ 190, 2-h ≥ 165 and 3-h ≥ 145 mg/dl	136 women completed all phases of the study, no details of why 24 dropped out/excluded Mean serum glucose levels not significantly different after ingestion of standard glucose or jelly beans The mean difference in individuals also non-significant 5 cases of GDM were identified (3.7%): 3 by standard glucose alone, 1 by jelly alone and 1 by both Reports of nausea, dizziness or headache for standard glucose 38% and for jelly beans 20% ($p < 0.001$) Women preferred jelly bean source ($p < 0.001$)
Sensitivity and specificity Reported sensitivity for standard glucose 80%; specificity 82%, PPV 15% and NPV 99% Reported sensitivity for jelly beans 40%, specificity 85%, PPV 9% and NPV 97% (sensitivity checked and correct but no data on numbers screened positive therefore can't check specificity) Reports that there is no significant difference in screening performance for jelly beans and the standard glucose				
Comments Authors comment that the lack of statistical difference between the 2 sources may be a type two error as only a small proportion of women with GDM Authors suggest that due to different rate of absorption over a time period greater than an hour, probably cannot be recommended for the 100 g glucose load for a GTT High proportion of reported side-effects				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Landon *et al.*, 1990[239]	Philadelphia, Pennsylvania, USA Inclusion criteria: 'high-risk' patients (not defined)	Not RCT *Aim*: to compare 1-h post-GCT CBG measurements with VBG levels measured in a laboratory 125 'high-risk' women	Screening test used: standard 50-g GCT at 26–28 weeks in fasting/fed state 1-h capillary blood obtained by finger stick, incubated with test strip for 1 min and read using AccuChek reflectance meter. All tests performed by a single trained nurse who calibrated meter Values compared with 1-h venous blood sample, taken at same time and sent to lab, assay performed within 1 h (threshold 135 mg/dl) Diagnostic test: 3-h GTT GDM diagnosed according to NDDG criteria[18]	6.4% (8/125) classified as GDM (NDDG criteria) in 'high-risk' population Costs: Laboratory testing of VBG US$5.90 per assay vs Capillary reflectance meter US$0.50 per assay plus US$150.00 for meter Cost savings per 1000 patients (included cost of additional 32 GTT per 1000 patients when using meter) = US$4866.00 (1986 prices)
Sensitivity and specificity ROC analysis based on 1-h CBG with screening considered positive if corresponding lab PG ≥ 135 mg/dl gave optimal level for performing 3-h GTT as 160 mg/dl Sensitivity and specificity for all values of meter CBG with test results being classified as positive if corresponding lab PG ≥ 135 mg/dl: sensitivity 93% (28/30); specificity 96% (91/95); PPV 88% (28/32) 32 patients ≥ 160 mg/dl on CBG vs 30 patients ≥ 135 mg/dl for VBG *Authors' comments*: There is a need for larger studies to confirm these findings. Training and skill are required for those performing capillary estimations *Authors' conclusion*: Potential advantages for the use of a reflectance meter at a clinic include both cost and efficiency				
Comments 'High-risk' patients were studied but this category was not described Influence of fasting/fed status of women undergoing GCT not examined Only those with CBG ≥ 160 mg/dl or with VBG ≥ 135 mg/dl before 28 weeks underwent GTT. May have underestimated GDM rates Compares sensitivity and specificity according to VBG levels and not GDM diagnosis No fetal/maternal outcomes Costs of two methods of measuring BG compared				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Landy *et al.*, 1996[112]	Miami, USA 345 white (67%); 169 black (33%) Women with a BG level ≥ 140 mg/dl Mean age 30.7 years (SD 6.3); 30% had a FH of DM Mean pre-pregnancy BMI was 27.2 (SD 6.3). 376/514 of the women had had prior pregnancies – 8.5% were complicated by GDM, congenital abnormalities 2.7%, stillborns 3.5% and macrosomia 42.4% (defined as > 4 kg)	Retrospective study seeking to determine whether an elevated level of BG detected by a screening test was sufficient to diagnose GDM Uses ROC curve analysis (sensitivity plotted against 1 – specificity). [area under curve of 0.5 = no better than chance, 1.0 = perfect diagnostic ability] *n* = 514	Women who had previously shown a BG level of ≥ 140 mg/dl on screening (at 24–28 weeks; 1-h GCT; no info. on fasting) were given a 3-h diagnostic GTT (O'Sullivan and Mahan criteria[19])	Prevalence was assumed to be 2.7% (based on previous 2 years data) for purposes of ROC curve calculations 202/514 women were diagnosed with GDM whilst 312 were normal (non-GDM group) 333 vaginal deliveries (29 assisted), 181 CS. Mean birth weight 3468 g (SD 653 g). 30.5% (157) were LGA (defined as at or above the 90th percentile for gestational age) 20% of women with GDM had had a history of GDM compared to 6.3% in the non-GDM group. 45.5% had a history of macrosomia compared to 18.8% in the non-GDM group. The CS rate was 37.7% (vs 33.3% in the non-GDM group), average birth weight was 3565 g (vs 3333 g in non-GDM group), LGA rate was 44.3% (vs 28.6% in non-GDM group), hypoglycaemia rate was 23.4% (vs 9.5% in non-GDM group) The authors identified a BG level of 186 mg/dl with the GCT to be sufficient for a "probable diagnosis" of GDM indicating a GTT unnecessary (based on the diagnostic stats below) 21% of the sample in this study had GCT levels above this point No adverse effects documented
Sensitivity and specificity 50-g GCT at ≥ 186 mg/dl: sensitivity 38.1% (77/202); specificity 95.9% (492/513); false-positive rate = 4.1% (21/513); PPV 78.6% Data exist to calculate these statistics at other levels of screening (at 5 mg/dl intervals)				
Comments Study indicates that high glucose levels on screening precludes the need for a GTT and would result in some cost savings (over giving GTTs to all positive screens) The ROC curve was estimated assuming that none of the women who had previously had negative screens (i.e. < 140 mg/dl; *n* = 201; so grand total 715) had GDM (i.e. assuming 100% sensitivity of GCT) ROC curve calculations based on prevalence of 2.7%, uncertain of incidence in study group Sensitivity of GCT in diagnosing GDM calculated as 38.1% not 36.1% as shown in the paper No other scenarios presented but can calculate from raw data 100% of women with screens ≥ 226 mg/dl (16) were diagnosed with GDM No costs given				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lavin, 1985[148]	USA No data on ethnic background, etc. given here but referred to previous paper (Lavin *et al.*, 1981)[240]	Comparison of RF screening with GCT *n* = 2077	Patients were divided into 2 groups, those with established RFs (group 1; *n* = 959) and those without (group 2; *n* = 1118) Established RFs used were: previous GDM pregnancy; family member with DM; previous macrosomic infant (> 4000 g); previous unexplained stillborn; previous infant with congenital malformations; history of recurrent abortion; obesity; monilial vaginitis; glucose uria; polyhydramnios; suspected LGA infant Group 1 patients were given the 1-h GCT at their initial visit and again at 28 weeks. Group 2 patients were given a 1-h GCT screening test between 28 and 32 weeks Criteria for an abnormal screening test was > 150 mg/dl (Second International Workshop-Conference on GDM[216]) Diagnostic test: 3-h GTT (O'Sullivan and Mahan[19]) Two or more of the following values had to be obtained to confirm the diagnosis: initial ≥ 105; 1-h ≥ 190; 2-h ≥ 165; 3-h 145 mg/dl No information was provided on whether there was fasting prior to testing	Incidence of GDM was determined by the combination of those detected in either group, i.e. 30 (14 in group 1 and 16 in group 2). Incidence rates (detected) given as 1.5% and 1.4%, respectively Group 1 method detected 46.7% of those gestational diabetics identified in the study whilst group 2 detected 53.3% of those detected in the study Total cost of screening programme in this study was US$9869; cost per patient screened US$4.75 and cost per case of GDM detected US$328.96 (all 1985 prices) No adverse effects documented

Sensitivity and specificity
50-g GCT: [≥ 150 mg/dl] PPV 24% (16/68); not enough data to calculate sensitivity or specificity

Comments
Incidence of GDM in both groups could be expected to be higher since in both groups, only the positive screens were advanced for diagnosis. Detection rate of GCT-based screening could be expected to be higher in a normal group of pregnant women since it would include some of those with RFs
Only positive screens given diagnostic test so incidence may be higher if some GDM undetected. Assumes combination of both tests = 100% sensitivity
Costs significantly outdated and in US dollars
Lab costs included phlebotomy supplies, reagents, equipment and depreciation and technician time. Pharmacy costs included the glucose solutions. Phlebotomist time was included but physician interpretation time was not
Few of the group 1 women attended for a second GCT at 28 weeks (211/959); had they done so this may have increased incidence

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lemen *et al.*, 1998[192]	Wisconsin, USA Adolescents classified as < 20 years. Mean age was 16.86 years (SD 1.49). 75.2% were black, 16.4% were white and 7.6% were Hispanic and 0.8% were from other racial groups. 78.8% were nulliparous, 17.9% were primiparous and 3.3% had delivered two or more children previously	*Aim*: to determine incidence and cost of screening for GDM in an adolescent population *n* = 509 Original pregnant population was 638 but 129 were excluded as they were not screened for GDM	Screening was performed using the 1-h 50-g GCT at 24–28 weeks, or earlier in the presence of RFs Those women with a screen ≥ 140 mg/dl were given the 3-h 100-g GTT (O'Sullivan and Mahan[19] values) after a 2-day CHO load and overnight fasting	There were 23 (4.5%) positive screens. This includes 2 women who had a negative result when screened earlier than 24–28 weeks. 14 of the positive screens had at least one RF. No significant differences were detected between the positive and negative screen groups Incidence was 1.18% (6/509) Of the 6 women diagnosed with GDM, 4 (1.06%) were black and 2 were Hispanic (5.26%). All 6 had a FH of DM and 5 had glycosuria 30 (5.9%) women had macrosomic babies (> 4000 g), none of these were to the women with GDM. 1 of the mothers of a macrosomic baby had a positive screen but negative diagnosis. 2 others had screens between 130 and 139 mg/dl Laboratory costs of a GCT and GTT were given as US$25 and US$50, respectively. The cost of testing the study population was calculated at US$16,400 and cost per case of GDM diagnosed was US$2733 If screening were done on the basis of RFs, the cost per case of GDM detected would have fallen to US$1258 (all the women had a FH of DM) Using their incidence results, the authors further calculated that universal screening for all adolescents in the USA would cost US$13.9m or US$2292 per case diagnosed No adverse effects documented
Sensitivity and specificity Insufficient data to calculate accurately as only positive screens given diagnostic test Assuming all women with GDM detected, GCT [≥ 140 mg/dl] sensitivity 100% (6/6); specificity 95% (480/503); PPV 21% (6/29)				
Comments Only laboratory costs included, no nursing or administration costs Only positive screens given diagnostic test so incidence may be higher if some GDM undetected				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Li *et al.*, 1987[64]	Hong Kong High-risk population Inclusion criteria: PH of GDM, previous macrosomic infant (> 4 kg); positive FH; accelerated fetal growth in present pregnancy; recurrent glycosuria; maternal obesity (> 120% ideal weight); previous unexplained stillbirth; congenital abnormality Mean age 28.8 years (SD 4.3); mean gestational age of diagnosis of glucose intolerance 29.2 weeks (SD 6.3)	Comparison of 100-g GTT (NDDG criteria[18]) and 75-g GTT (WHO criteria[24]) 3084 screened with 100-g GTT 401 (13%) were abnormal (NDDG criteria) 347 (87%) of these women underwent subsequent 75-g GTT (WHO criteria) 13% dropped out	High risk on clinical factors tested using 100-g OGTT. Those positive on NDDG criteria underwent 75-g OGTT Positive 100-g GTT on NDDG: fasting 5.8; 1-h 10.6; 2-h 9.2; 3-h 8.1 mmol/l Classified after 75-g GTT on WHO criteria as normal glucose tolerance; IGT and GDM 2-h PG: WHO criteria: normal < 8.0 mmol/l; IGT 8–10.9 mmol/l; GDM > 11.0 mmol/l NDDG criteria: normal < 9.2 mmol/l; GDM ≥ 9.2 mmol/l	13% (401/3084) abnormal response to 100-g GTT using NDDG criteria Area under glucose response curve estimated in units and compared after 100-g and 75-g GTT No adverse effects documented
Sensitivity and specificity 347 had 75-g GTT after abnormal 100-g GTT (NDDG criteria). These 347 classified by WHO criteria after 75-g GTT as: 157 (45%) normal, 178 (51%) IGT, 12 (3%) GDM Comparing classification after 100-g GTT at 2 h using WHO (8.0 mmol/l) and NDDG (9.2 mmol/l) criteria: reports agreed on 60% of cases (IGT only). Comparing classification after 100-g GTT vs after 75-g GTT using WHO criteria at 2 h: agreed on 47% of cases Comparing mean BG values in 3 groups (normal, IGT, GDM) using classification by WHO after 75-g GTT: no significant difference between groups on fasting or 1-h values; significant difference between all groups on 2-h values; no significant difference between normal and IGT groups in 3-h values *Authors' conclusion*: The maternal glucose response is dependent upon the size of the glucose challenge				
Comments [NB: no NDDG negative/WHO positive classifications possible, thus figures below flawed] Selected high-risk population Gestational age not considered No follow-up till end of pregnancy so may underestimate incidence of GDM Only those testing positive by NDDG criteria on 100-g GTT underwent 75-g GTT using WHO criteria, thus omitting possibility of detecting those negative on NDDG and positive on WHO tests and negating analysis presented Conditions under which GTT was performed not stated (i.e. fasting/previous diet)				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lind, 1985[138]	UK study No demographic details given Women aged between 17 and 44 in study one. No details given for study two Those with pre-existing DM excluded	Unselected women in two studies. Study one 186 women, study two 2403 women *Aim*: to evaluate the usefulness of RBG as a screen for GDM	Study one: RBG between 24 and 32 weeks Study two: Two RBG measurements, one at first antenatal visit, second at 28–32 weeks BG grouped by 30-min intervals from time of last meal and then regrouped by hourly intervals and the mean glucose values calculated. The average SD was used to calculate the 95 and 99% upper limits. The 99% limits were chosen for screening thresholds For study one these were: 6.4 mmol/l within 2 h of eating and 5.8 mmol/l > 2 h For study two: 6.1 within 2 h and 5.6 > 2 h Those above the 99% values were given 75-g GTT (fasting status or thresholds not defined)	Study one incidence – nil Study two incidence – 0.16% (4 women; 2 had such high random glucose levels that diagnosis made on basis of this not the GTT) Women were also referred for GTT on the basis of glycosuria (2/7 referred had IGT), previously large baby (6, none of which had abnormal GTT), and previous GDM (1/2 diagnosed IGT) None of the patients with GDM had any RFs Three were treated with insulin, the fourth with diet
Sensitivity and specificity No data available to calculate sensitivity or specificity In study two, PPV (including 2 patients whose GDM was diagnosed on the basis of fasting glucose alone) 11.7%				
Comments In study one there were no patients who had not eaten within 30 minutes of the test so data from a previous sample used As blood obtained at stage in pregnancy where other bloods being taken, suggests that the costs shared with other determinations and less disruptive to clinic routine and to patients No discussion of variance in 99% cut-offs between the two studies				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lind and Phillips, 1991[241]	10 centres throughout Europe Reports that a general prenatal population was selected. No details of methods used to select sample in different centres Exclusion criteria: All centres excluded those with insulin dependent DM. One centre also excluded women with FH DM in a first-degree relative Mean (range across centres): BMI 26 (range 23–29); age 27 years (range 25–31) 14% tested in early pregnancy (< 117 days); 43% tested in mid-pregnancy (117 to 196 days); 43% tested in late pregnancy (> 196 days)	Not RCT Multicentre observational study *Aim*: to report responses to a 75-g GTT 1009 women provided data from 1187 separate tests	Test used: 75-g OGTT following WHO criteria[24] Different centres tested plasma, VBG or CBG levels 1 h after GCT Fasting/fed status not reported GDM diagnosed according to WHO criteria as 2-h VBG > 11 mmol/; IGT as 2-h VBG 8–11 mmol/l; non-diabetic 2-h VBG < 8 mmo/l	7.8% (79/1009) had either IGT or GDM (WHO criteria 2-h VBG or CBG > 8 mmol/l) ??? 1.3% (13/1009) diagnosed as GDM (WHO criteria) Those with IGT/GDM were comparable to normal women on parity, maternal height, gestation at delivery, and weight of infants A total of 7 stillbirths, 1 neonatal death and 1 abortion were recorded. None were associated with high 2-h BG

Sensitivity and specificity

N/A

Mean (range across centres) was reported for the following factors:

Gestation at delivery 278 days (range 272–281); birth weight 3367 g (range 3301–3377); vaginal delivery 81% (range 67–94%); forceps 9% (range 0–21%); CS 10% (range 6–25%)

Authors' comments: Missing data from some patients

12/13 women with fasting VPG ≥ 6.9 mmol/l came from one centre

Mean BG values were significantly higher in the third trimester compared with earlier pregnancy

95th percentile for VBG:

Early pregnancy: fasting 4.9 mm; 2-h 6.8 mm

Mid-pregnancy: fasting 6.9 mm; 2-h 9.0 mm

Late pregnancy: fasting 5.2 mm; 2-h 9.0 mm

95th percentile for CBG:

Early pregnancy: fasting 4.8 mm; 2-h 7.3 mm

Mid-pregnancy: fasting 4.7 mm; 2-h 7.2 mm

Late pregnancy: fasting 4.5 mm; 2-h 8.2 mm

Only 9/79 women with 2-h BG ≥ 8 mm had FBG ≥ 6.9 (7 came from 1 centre)

By considering fasting and 1-h BG as well as 2-h BG, the number of women classified as abnormal would be reduced from 79 to 15

By increasing the 2-h cut-off value to ≥ 9 mm, the number of women classified as abnormal would be reduced from 79 to 32

Comments

Mixed sample from 10 different centres with different rates of interventions and population characteristics. It is not clear whether the samples were representative of the general population within each participating country

Assumes all GDM detected by tests used in each centre

There was no systematic investigation of variability in results from different centres though one centre had clustering of high BG values. There was no discussion of possible reasons for this clustering

Assumes sample came from homogeneous population

Small number of neonatal morbidity outcomes reported (7 stillbirths, 1 neonatal death)

Relationship of diagnosis of IGT/GDM to fetal/maternal outcomes was not explored

The adjustments of definitions of GDM appear to have been made with the intent of reducing the number diagnosed and were not dependent on the association between abnormal values and either fetal/maternal outcomes

The aim of this paper is not really clear

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Lindsay *et al.*, 1989[242]	Atlanta, Georgia, USA All prenatal patients attending clinic were screened Inclusion criteria: Study group: abnormal screening test and only one abnormal value of GTT (C&C criteria[20]) Control group: women with normal screening selected randomly from database Mean age of study group higher than control (25.8 vs 23.0 years; $p < 0.0001$) Mean parity higher in study group (1.3 vs 0.9: $p = 0.004$). Similar gestational age of infants (39.6 vs 39.3 weeks) No details of ethnicity	Not RCT Comparative study *Aim*: to determine whether women with only one abnormal value on GTT are at greater risk than a control group with normal glucose screening in pregnancy Study group: 139 women with one abnormal value on GTT Control group: 725 women with normal GCT screening	Screening test used: 1-h PG tested after 50 g glucose solution. Screened at 24–28 weeks. No dietary restriction Threshold 135 mg/dl Diagnostic test: 3-h 100-g GTT after 3 days of CHO loading and a 12-h overnight fast GDM diagnosed using C&C criteria (fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl)	4.4% (202/4618) diagnosed as GDM (C&C) After controlling for race, maternal obesity, parity, pregnancy duration, and morbid obesity, a single abnormal value on GTT was associated with a greater risk of 1. A macrosomic infant (≥ 4000 g): OR 2.55 (95% CI, 1.77 to 5.37) 2. Pre-eclampsia/eclampsia (after controlling for nulliparity, maternal age, morbid obesity, multiple gestation and race): OR 2.55 (95% CI, 1.44 to 4.52) Study group had more CS than controls (32.4% vs 22.3%) but after controlling for confounding factors there was no significant difference: OR 1.45 (95% CI, 0.92 to 2.29) No significant differences between groups in incidence of 5-min APGAR scores < 7, preterm delivery, shoulder dystocia, congenital malformations, perinatal mortality or chronic hypertension
Sensitivity and specificity N/A *Authors' comments*: Frequency of macrosomia in study group was 18%; discusses control for confounding factors, possibility of selection bias; suggests that current diagnostic criteria for GDM need refining *Authors' conclusion*: Patients with one abnormal value on a GTT during pregnancy are at risk for delivering macrosomic infants and developing pre-eclampsia/eclampsia				
Comments Clearly presented with analysis controlling for potential confounding factors and defined fetal/maternal outcomes Sample representative of local population though ethnicity of study population was not described Assumes all those with GDM were detected with GCT ≥ 135 mg/dl at 28 weeks. May have underestimated the incidence of GDM and some women with GDM may be included in the control group Authors' conclusions are supported by the evidence in this population with GCT at 28 weeks				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Magee *et al.*, 1993[94]	USA Women from a prepaid health maintenance organisation, used for population research Possible bias as higher socio-economic status Mean age 29 (SD 5)	No randomisation, no control group Compared two sets of criteria for diagnosing GDM 886 women screened and tested for GDM with both NDDG[18] and C&C modified criteria[20] Also compared negative screen women with positive on clinical outcomes	Screening test: Fasting 50-g GCT Threshold ≥ 7.77 mmol/l Diagnostic test: 3-h fasting 100-g GTT Thresholds as defined by NDDG: any two: fasting ≥ 5.9; 1-h ≥ 10.6; 2-h ≥ 9.2; 3-h ≥ 8.1 mmol/l and C&C: any two: fasting ≥ 5.3; 1-h ≥10.1; 2-h ≥ 8.7; 3-h ≥ 7.8 mmol/l Testing at 24–32 weeks	91 positive screens refused GTT, thus not included Incidence with NDDG = 3.2% Incidence with C&C modified criteria = 5% GDM mothers significantly different from negative screenees on: FH; BMI; infant birth weight ratio; CS; resuscitation at delivery; anaemia Modified criteria women with GDM similar in terms of age, FH, infant macrosomia as women diagnosed with NDDG criteria
Sensitivity and specificity No data available for calculation of sensitivity PPV: NDDG criteria 17.9%; modified criteria 27.2%				
Comments Those screened negative not tested with GTT but compared on indices of clinical outcomes				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Maresh *et al.*, 1989[199] Study between 1976 and 1984	London, England Mean age: 31 years Multiparous: 68% White: 25% Married: 85% Weight: abnormal GTT 75 kg; controls 70 kg Study population differed from general population in age (older); marital status (more married); parity (greater parity), and ethnicity (more non-white) Mean gestation at first visit: Abnormal GTT: 29 weeks (range 8–38 weeks) Control (normal GTT): 27 weeks (range 10–38 weeks)	Not RCT Observational cohort study *Aim*: to determine which maternal factors (age, weight) contributed to the severity of abnormal glucose tolerance, infant birth weight; and neonatal morbidity among women with GDM Abnormal GTT: 213 Classified into 2 groups according to severity of DM: Class A1: FPG < 6 mmol/l (160 women) Class A2: DM: Class A1: FPG ≥ 6 mmol/l (53 women) Controls (matched for age, parity and ethnicity): 213	Screening test used: all had 1-h PG after 50-g GCT with no dietary preparation at booking clinic Abnormal if glucose > 7.7 mmol/l (139 mg/dl) Diagnostic test: standard 3-h GTT after prior fast GDM diagnosed if area under the curve > 42 'mmol/l units' If abnormal screening but normal GTT, screening test repeated at 26–28 weeks Glucose load for GTT not reported Joint diabetic antenatal clinic aimed to keep PG < 7 mmol/l by dietary regulation. If control not satisfactory on diet (FPG persistently > 6 mmol/l) insulin started. Individually tailored regimes with combination of short and medium acting insulin twice daily. Women self-monitored at home. By delivery, 20% in A1 group and 50% of the A2 group were taking insulin Obesity defined as Quetlet index ≥ 31	Reports rates of GDM as 1.5% between 1976 and 1984 (area under the curve > 42 'mmol/l units'. Gives reference as Gillmer, 1975) Birth weight not related to maternal age or severity of DM but was related to maternal obesity Neonatal morbidity indices such as admission to SCBU for longer than 48 h were significantly related to the severity of DM and not to maternal age or obesity See later for detail

Sensitivity and specificity
N/A
Birth weight was not related to maternal age or severity of DM but was related to maternal obesity
Neonatal morbidity indices such as admission to SCBU for longer than 48 h were significantly related to the severity of DM and not to maternal age or obesity
GDM, GDM severity and maternal age and weight:
GDM mothers were older than general population (31 vs 27 years; $p < 0.001$). Non-significant relationship between age and severity of DM, or between age and weight of mother. Maternal weight was related to severity of DM (51% in A1 vs 33% in A2 with obesity)
Infant birth weight and severity of DM
Macrosomia: non-significant downward trend in % macrosomic infants from A2 group to controls (19% in A2 vs 14% in A1 vs 10% in controls)
Infant birth weight and maternal obesity
No significant difference in mean birth weight in obese GDM vs non-obese (3580 vs 3280 g). Significantly more macrosomic infants in obese vs non-obese (24 vs 10%)
Neonatal morbidity
After allowing for obesity severity of DM was associated with a significantly increased incidence of admission to SCBU for > 48 h ($p < 0.05$); and neonatal hypoglycaemia ($p < 0.05$)
No significant difference in rates of 5-min APGAR scores < 7 among groups
Stillbirths: GDM = 0.9% (2/213) vs control = 0.9% (2/213)
Neonatal deaths: GDM = 0.5% (1/213) vs control = 0%
Mode of delivery
Differences in rate of CS among groups N/S (30% in A2 vs 20% in A1 vs 14% in controls)
Authors' conclusion: In treated GDM mother, infant birth weight is more a function of the weight of the mother than of severity of CHO intolerance. End points such as admission to SCBU are more related to the mother's GDM than to her obesity. Concludes that screening for GDM and subsequent treatment are worthwhile. Suggests there is a need for a therapeutic trial of GDM before a definite policy statement can be made

Comments
High proportion of non-whites in sample
Reported very low rates of GDM (1.5%). Evidence for validity of criteria used to diagnose as GDM not evaluated
Not clear why a level of 6 mmol/l was selected to classify severity of DM
No control for confounding by stage of pregnancy at which glucose intolerance developed, was detected or was adequately controlled, or changes in obstetric practice over time
Authors point out:
that community has a high proportion of non-whites, the group that predominates in this study
that sample was too small to determine differences in perinatal mortality
GDM mothers were treated and this may have removed the effect on the fetus of metabolic disturbance due to DM
Reports changes in practice over the study period with less obstetric intervention over time

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Marquette *et al.*, 1985[124]	USA Ethnic background: 84% black, 15% white, 1% other Mean age 24.4 years	Cost-effectiveness study to determine criteria for screening test Extent of cost-effectiveness was to maximise PPV. *n* = 1012 A further 22 patients failed to attend for GTT appointment and were excluded All patients given screening GCT and threshold levels and RFs examined	Screening was performed at between 26 and 30 weeks after overnight fasting Screening test: GCT PG was determined 1 h after the 50 g glucose load was taken Those rated positive (≥ 130 mg/dl) were given the diagnostic test within 2 weeks Diagnostic test: 3-h GTT (O'Sullivan and Mahan[19]; NDDG[18]). Two or more of the following values had to be obtained to confirm the diagnosis: initial ≥ 105; 1-h ≥ 190; 2-h ≥ 165; 3-h 145 mg/dl	Incidence was 2.4% (24/1012) Sharp increase in incidence of GDM with increase in glucose screen value 3 patients had a screen ≥ 200 mg/dl – all had GDM Only 1/24 had a screen level below 150 mg/dl 22/24 were ≥ 24 years of age 21/24 were ≥ 24 years and had a screen level of ≥ 150 mg/dl No adverse effects documented Costs of diagnosis using ≥ 150 mg/dl threshold and aged ≥ 24 years was 40% that of universal screening with ≥ 130 mg/dl
Sensitivity and specificity 50-g GCT: [≥ 130 mg/dl] PPV 10% (24/235); sensitivity assumed 100% sensitive since only positive screens were given diagnostic GTT [≥ 150 mg/dl] PPV 24% (23/96); sensitivity 96% (23/24) [≥ 130 mg/dl and aged ≥ 24 years]; sensitivity 92% (22/24); PPV 14% (22/153) [≥ 150 mg/dl and aged ≥ 24 years]; PPV 31% (21/68); sensitivity 88% (21/24)				
Comments Only positive screens given diagnostic test so incidence may be higher if some GDM undetected. Assumes GCT 100% sensitive The incidence of 2.4% assumes incidence zero in those not screening positive. Shown not to be the case in other studies Despite title, no costs actually given				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Martin *et al.*, 1995[60]	Australian women 40% Mediterranean and Middle Eastern and 2.4% Asian or Chinese surnames Selection of RFs – not defined Previous DM excluded	Unselected non-randomised 1371 women Comparison 75-g GTT in all women with varying diagnostic guidelines	Screening/diagnostic test: Fasting 75-g GTT Fasting and 2-h bloods Criteria: WHO[108]: 2-h ≥ 8.0 mmol/l; WHO[24]: 2-h ≥ 7.8 mmol/l Australian: fasting ≥ 5.5 mmol/l and/or 2-h ≥ 8.0 mmol/l Testing at 26–32 weeks	Incidence: 4.2% WHO[108]; 5.2% WHO[24]; 5.5% Australian No differences in neonatal outcomes (not defined) Vomiting in 0.5% women
Sensitivity and specificity No screening tests used				
Comments No cost data available				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Mello *et al.*, 1997[96]	All white Italian women Universal screening in unit Mean age 30.4 (SD 5.3) Obesity (12% to 30%); previous macrosomia (0% to 2.4%); previous GDM (0.1 to 5.3%); FH of DM (34% to 46%)	Not randomised, no control 1883 women Investigated link between glucose metabolism and RFs for fetal overgrowth Also evaluated effect of positive GCT with no GDM on obstetric complications Defined 7 groups: 1. Abnormal GCT but absence GDM 2. Normal GCT period 1 and abnormal but absence GDM at period 2 3. Abnormal GCT but absence of GDM period 1, and normal GCT period 2 4. Normal GCT (used as control) 5. GDM period 1 6. Normal GCT period 1 and GDM period 2 7. Abnormal GCT but normal OGTT period 1 and GDM period 2	Screening test: Non-fasting 50-g GCT PG level ≥ 7.5 mmol/l at two periods (16–20 and 26–30 weeks), with aim of identifying GDM without need for OGTT Diagnostic test: Fasting 3-h OGTT within 1 week 2 or more PG levels as defined by C&C[20]	Incidence GDM = 4.4% period 1; 7.6% period 2 (total 11.7%) Only glucose metabolism, PH of mothers had a significant effect on birth weight Significant increase in risk of fetal overgrowth in groups 1, 2 and 6 (not accounted for by factors such as previous GDM etc.) Also previous birth > 4000 g No differences in frequency of normal delivery, CS or vacuum extraction Minor abnormalities of glucose metabolism without GDM significant RF for fetal overgrowth and related to time of onset of abnormal glucose, such that a positive GCT at 26–30 weeks most important RF. May be due to insulin therapy of those with GDM No adverse effects of tests noted Supports the consideration of a positive 1-h GCT as a marker of abnormal glucose metabolism
Sensitivity and specificity At 16–20 weeks: sensitivity 64% (140/220); specificity 87% (1455/1663); PPV 40% (140/348)				
Comments RFs accounted for: Previous GDM, previous birth weight > 4000 g but absence GDM, maternal obesity, multiparity, gestational age Maternal age > 35 years Complications: Pregnancy-induced hypertension, intrauterine growth retardation, polyhydramnios and preterm delivery				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Meriggi *et al.*, 1988[149]	Turin, Italy Inclusion criteria: women aged ≥ 25 years with fasting plasma BG < 95 mg/dl Included those with and without RFs Those with RFs also had 2-h post-lunch PG < 120 mg/dl	Not RCT *Aim*: to examine the relationship between capillary and venous plasma glucose concentrations 418 women were included, comprised 3 groups: 1. 122 women without RFs screened 24–28 weeks 2. 116 women with RFs screened 12–16 weeks 3. 180 women with RFs screened 24–28 weeks	Screening test used: paired venous and capillary glucose levels fasting and 1 h after 50-g GCT following 12-h overnight fast Those without RFs had screening between 24 and 28 weeks. Those with RFs ideally had GCT between 12 and 16 weeks. If negative was repeated at 24–28 weeks. Those presenting late only had 24–28-week screening GCT positive if PG ≥ 135 mg/dl (C&C[20]). If positive underwent GTT Diagnostic test: 3-h 100-g GTT after 3 days of unrestricted diet (> 250 g CHO) and overnight fast GDM diagnosed if two or more values ≥: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl (C&C) CBG tested using Reflocheck strip and Reflecheck reflectance meter. Interassay variation was < 5% in testing	6.5% (27/418) diagnosed as GDM (C&C). Selected population 26.5% (111/418) had positive GCT (≥ 135 mg/dl) Differences in GDM rates between groups N/S Complications in GDM cases: toxaemia 11% (3/27); CS 15%; neonatal morbidity 19%; macrosomia 11% (3/27)

Sensitivity and specificity
Sensitivity and specificity of various cut-off values:
Values as reported. Plasma (GCT):
≥ 135 mg/dl: sensitivity 100%; specificity 80%; PPV 21%
≥ 140 mg/dl: sensitivity 96%; specificity 84%; PPV 25%
≥ 145 mg/dl: sensitivity 82%; specificity 88%; PPV 27%
≥ 150 mg/dl: sensitivity 74%; specificity 90%; PPV 29%
Capillary (GCT):
≥ 155 mg/dl: sensitivity 100%; specificity 85%; PPV 27%
≥ 160 mg/dl: sensitivity 89%; specificity 87%; PPV 26%
≥ 165 mg/dl: sensitivity 82%; specificity 90%; PPV 31%
≥ 170 mg/dl: sensitivity 70%; specificity 91%; PPV 31%
ROC curves showed optimal cut-off value in GCT as:
plasma = 135 mg/dl and for capillary blood 155 mg/dl
Authors' conclusion: The optimal cut-off values for screening GCT were set at 135 mg/dl for plasma and 155 mg/dl for capillary blood
These cut-off values should be considered in the light of the costs and perinatal outcome

Comments
Selected population (aged ≥ 25 years and fasting PG < 95 mg/dl). Included only those with levels below specified values on initial screen. Reasons for this are not clear
Results may differ in sample with higher initial BG levels
Different regimes for groups with and without RFs
Assumes all cases of GDM detected by GCT ≥ 135 mg/dl before 28 weeks
Fetal/maternal outcomes not evaluated for criteria suggested

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Merkatz *et al.*, 1980[243]	USA 61% white and 39% non-white (representative of local population) Mean age 26 years (range 12–44) All non-diabetic patients screened, regardless of the presence of any RFs	Observational study, no randomisation or control 2225 unselected women *Aim*: to establish a suitable screening program for the local population	At 24–28 weeks (range 16–36), 2-h 75-g GCT (with commercially available cola) without regard to fasting status BG levels drawn from capillary samples. 2-h values of ≥ 120 mg/dl given a 3-h 75-g fasting GTT. Diagnosis made on modified O'Sullivan and Mahan criteria[19]: 2 or more: fasting 105; 1-h 185; 2-h 140; 3-h 125 mg/dl (2 and 3-h modified)	Significant difference mean age white (26.0) and non-white (24.0), $p < 0.0001$ Mean time of testing different between white (26.2 weeks) and non-white (24.5), $p < 0.001$ Only 52% screened in target period of 24–28 weeks Positive screen results significantly higher in white women than non-white (13.3% vs 8.7%, $p < 0.001$) A significantly higher incidence of positive screen above 24 weeks (8.1% vs 26%, $p < 0.025$) Frequency positive screen increased at various maternal ages, significant for non-whites above 24 years ($p < 0.001$) 257 women with positive GCT, 4 did not have GTT. Incidence GDM 3.1% (69 women) The efficiency of the 1-h and 2-h screening values (screening defined as positive only on 2-h value) – 505 women with positive 1-h value had normal 2-h value. 68 women had normal 1-h and elevated 2-h values, 7 of whom had GDM
Sensitivity and specificity No data available to calculate sensitivity or specificity. None of the negative screened women were given GTT, not even those with a high 1-h value				
Comments Authors suggest superiority for the 2-h value, but did not test women with just positive 1-h value No discussion as to the use of capillary blood rather than venous				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Meyer *et al.*, 1996[160]	USA 170 Hispanic (51.7%), 126 black (38.3%) and 33 white (10%) of lower to middle socio-economic status Mean ages in each study group: 23.7 (SD 5.8), 28.5 (SD 4.4) and 30.9 (SD 6.5)	Retrospective observation study of 329 women *Aim*: to determine the benefit of early screening. Looked at maternal outcome: 1. maternal age, 2. race, 3. parity, 4. pre-pregnancy weight, weight at each screening and delivery, 5. presence of RFs, 6. gestational age at birth, 7. birth weight and 8. birth trauma and RFs: FH, obesity, glycosuria and/or a prior pregnancy with stillbirth, habitual abortion, hydramnios, GDM or macrosomia	At time of delivery, records of women reviewed. At first registration (mean 15.3, SD 7.0 weeks 'early testing') all women had 1-h 50-g GCT without regard to fasting state Positive screening test defined at > 135 mg/dl These followed within 1 week, with a 3-h 100-g fasting GTT. Diagnostic thresholds based on O'Sullivan and Mahan[19] All patients with an early negative GCT or early positive GCT with negative GTT repeated the procedure at 28 weeks ('late testing') Women classified into 3 groups for early and late testing: negative GCT, positive GCT and positive GTT	Early glucose screening identified 8 GDM (2.4%), 12 more at second testing (3.7%) RFs were present in 172 (52%), 7/8 GDM in early diagnosis had RF Of the 82 patients screened twice, 2 had GDM at second testing
Sensitivity and specificity Regression analysis of maternal and neonatal data revealed that black patients had a higher incidence of abnormal early screening and tolerance tests when compared to other racial groups ($p < 0.001$) Women with GDM detected in early testing were significantly older (> 30), ($p < 0.001$) A significant association between poor pregnancy weight gain and late development of GDM was demonstrated, independent of pre-pregnancy weight, maternal age or gestation age ($p < 0.001$) Macrosomia occurred in 26 (8%) of infants, 5 in GDM patients Sensitivity based upon finding of early and repeat testing 40%; specificity 88% and PPV 17%				
Comments All those with negative screening in early testing were given repeat test No cost data available Authors suggest that early glucose screening in patients older than 30, who are black or have RFs may be useful				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria Results from survey
Mires *et al.*, 1999[1]	Senior obstetricians in UK maternity units	Not RCT Semi-structured questionnaire survey 214 units completed responses (response rate 84%) *Aim*: to determine whether and how maternity units in the UK screen for and diagnose GDM	Screening tests used in 89% (191/214) of units Screening tests used: Maternal RFs 81% (154/191) leading directly to diagnostic test in 67% (128/191); Maternal glycosuria leading directly to diagnostic test in 57% (109/191); Maternal RFs leading to biochemical test then diagnostic test if indicated in 36% (69/191); Biochemical tests in all women in 30% (57/191) 84% (160/191) of units used more than one screening method Biochemical screening tests used in 68% (129/191) units. Of these, 44% (57/129) applied the test to all women, 56% (72/129) applied the test to women with other RFs. Biochemical screening tests used (n = 129): RBG 43%; RPG 11%; 50-g GCT 10%; other 8%; combination 25%; method not indicated 5%. Cut-off values were as follows: first random blood sugar test (range 5.5–11.0 mmol/l; modal 6 and 7.0 mmol/l); second random blood sugar test (range 5.5–11.1 mmol/l; modal 7.0 mmol/l); first RPG test (range 5.8–11.0 mmol/l; modal 6.0 and 6.4 mmol/l); second RPG test (range 5.8–8.0 mmol/l; modal 5.8 and 6.0 mmol/l) 84% (88/105) of studies using biochemical screening gave details on gestation at time of test: < 12 weeks 27%; 13–19 weeks 20%; 20–27 weeks 7%; 28–34 weeks 45%. 49% (43/88) of units performed 2 tests in pregnancy Diagnostic tests used: 79% (168/214) units used a 75-g OGTT; 14% (29/214) used a 50-g OGTT; and 1% (3/214) used a test meal 66% (142/214) of units had a joint obstetric/diabetic clinic. 58% (125/214) of units had a written policy on GDM screening. 76% (162/214) of units were in favour of national guidelines; 14% (29/214) were against guidelines and 10% (23/214) were undecided
Sensitivity and specificity N/A *Authors' conclusion*: The majority of obstetric units in the UK screen for GDM but with little consensus on the appropriate screening method. National guidelines would probably be welcome			
Comments Good response rate to survey (84%) Criteria for diagnosing GDM in UK maternity units not reported Provides overview of current (1998) lack of consensus on screening for GDM			

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Moses *et al.*, 1995[119]	Australia No demographic details of women given Mean ages for 2 groups 26.4 (SD 5.2) and 28.1 (SD 5.5) years No inclusion or exclusion criteria noted	Observation study of unselected sample of women undergoing universal testing for GDM No randomisation or control *Aim*: to determine proportion of women with GDM missed if testing confined to RFs 1209 women tested, data missing about RFs in 24 cases, thus 1185 analysed RFs observed: age, parity, weight as defined by BMI, FH	All women given fasting 75-g GTT. No screening test used Thresholds used were Australasian DM in pregnancy criteria: for first 6 months of the study this was a 2-h glucose ≥ 8.0 mmol/l; for last 12 months was if fasting ≥ 5.5 mmol/l and/or 2-h ≥ 8.0 mmol/l Mean time of testing 27.8 weeks	Incidence GDM 6.7% (previous incidence in same hospital 7.2%) Women without GDM significantly younger (26.4:28.1, $p < 0.02$) and had a lower BMI (24.2:25.9, $p < 0.05$) than women with GDM No difference in parity or gestational age of testing 31 women (39.2%) with GDM had no historical RFs and would have been missed if only selective testing undertaken
Sensitivity and specificity No screening test used				
Comments High incidence rate noted, threshold for diagnosis lower than NDDG but could also be due to ethnic mix as not stated No cost data available Authors conclude that women with no individual RFs had a GDM prevalence of 4.8% and that selective screening would miss more than one-third of cases of GDM				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Moses *et al.*, 1997[244]	New South Wales, Australia Women all had GDM, no ethnic background given, mean age 28.1 years (SD 5.3)	Study examining resource use and costs of diagnosis and treatment of women with GDM *n* = 134	No screening test as such; women offered 75-g GTT after overnight fasting. Tested at beginning of third trimester av. 28.1 weeks (SD 3.7) Fasting and 2-h PG samples taken. Diagnosis confirmed if ≥ 5.5 mmol/l and/or ≥ 8.0 mmol/l, respectively	Authors note that testing is cheap but managing a woman with GDM could costs hundreds of dollars Fee for a glucose sample is listed as A$9.55 or A$11.60 for 2 samples (prenatal hospital clinics were billed at 85% of these whilst private clinics charged no more than listed price) Cost of testing every pregnant woman for GDM was around A$10 For NSW, with 85–90,000 deliveries annually the cost would be < A$1 m Costs of management are detailed as follows: all women taught home glucose monitoring, hired home glucose meter; instructed to do home testing every 2–3 days fasting and 1-h samples after meals; dietary advice; insulin training and commencement if levels exceeded. There should also be additional costs of monitoring follow-up post-delivery and glucose testing Actual resources used were as follows: initial consultation with educator and dietician lasted 2 h and subsequent consultations lasted 30 min. All women attended the initial consultation and for a mean of 1.6 additional consultations. Those requiring insulin therapy required an additional 30 min for their initial visit and 15 min for each subsequent visit Cost of an educator's time is stated as A$38.50/h, insulin at A$219.67 per prescription, and scheduled fee for 3.4 consultations is A$243 No women required hospital admission for their GDM prior to delivery No adverse effects documented
Sensitivity and specificity No data given as women previously diagnosed				
Comments Incidence stated as 6.7–8.8% in Australia, which is much higher than UK estimates Cost of GTT given to all women given as similar (variance < A$1) to GCT to all and GTT to positive 15–20% No attempt is made to calculate treatment costs				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Nahum and Huffaker, 1990[157]	Los Angeles, California, USA White 61%; non-white 39% (composed of black 15%; Filipino 14%; Asian 10%; Indonesian 1%) 18% (208/1151) gravidas registering underwent first trimester screening and were eligible. Most patients registered after first trimester 40% of those eligible were subsequently excluded: failed to attend; GDM diagnosed; premature end to pregnancy	Not RCT *Aim:* to investigate the relationship between results of first trimester and early third trimester screening for GDM 124 women had paired first and third trimester screening	Screening test used: 1-h VBG after 50 g glucose load without regard to fasting/fed state Threshold ≥ 135 mg/dl Timing: first trimester and early third trimester (26–32 weeks) Diagnostic test: 3-h GTT as soon as possible after abnormal GCT. GDM diagnosed if fasting value ≥ 95 mg/dl or if two or more values ≥: 1-h 180; 2-h 155; 3-h 140 mg/dl	7.1% (82/1151) diagnosed as GDM (see criteria in adjacent column) 4.3% (9/208) of those screened in first trimester were diagnosed as GDM 8 were diagnosed as GDM during second trimester and 65 during third trimester No statistically significant differences in infant birth weights between women with the following GCT results in first trimester: ≤ 110 mg/dl; 110–139 mg/dl; ≥ 140 mg/dl

Sensitivity and specificity

Of 1-h BG > 110 mg/dl in first trimester in predicting GDM by third trimester: sensitivity 100% (9/9); specificity 60% (69/115); PPV 16% (9/55). Assumes all GDM detected by GCT ≥ 135 mg/dl followed by abnormal GTT by 32 weeks in third trimester

Reports: PPV of first trimester GCT for third trimester GCT values > 135 mg/dl as follows:

≤ 110 mg/dl: PPV 10% (7/69). White PPV 9% (4/45); non-white PPV 12% (3/24)

≥ 135 mg/dl: PPV 65% (15/23). White PPV 79% (11/14); non-white PPV 44% (4/9)

≥ 140 mg/dl: PPV 74% (14/19). White PPV 91% (10/11); non-white PPV 50% (4/8)

Significant difference in levels of maternal obesity between white and non-white

Authors' conclusion: Patients having elevated first trimester GCT ≥ 140 mg/dl are at particularly high risk for elevated GCT values later in pregnancy and should forego repeat 1-h GCT in favour of 3-h GTT

Comments

Variable policy at clinic with the majority of first prenatal visits conducted by nurse practitioners who performed universal screening and the rest being seen by physicians only some of whom practised universal screening

Only the 18% of patients registered in first trimester were eligible. Not clear whether these patients were representative of the total population

Analysis of sub-populations rather than multivariate analysis used to account for confounding factors such as maternal obesity

Subgroup analysis based on small sample size

Sensitivity and specificity of first trimester GCT ≥ 140 mg/dl in predicting GDM later in pregnancy was not reported or calculable from the data presented

Non-white population was not homogeneous

Both screening tests performed without regard to fasting/fed state. No account taken of fasting/fed state in the analysis

Cost of strategy studied was not reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Narchi and Kulaylat, 1997[245]	Saudi Arabia Inclusion criteria: charts of all infants diagnosed with Down's Syndrome (using chromosomal analysis) born in author's hospital between 1987 and 1994 Characteristics of population not described	Not RCT Retrospective review of charts *Aim*: to examine the relationship between Down's Syndrome and maternal DM 22,300 infants born 1870 infants born to diabetic mothers who did not have other autoimmune/ endocrine problem (1748 with GDM and 122 with pre-GDM)	Screening test used: 1-h PG after 50-g GCT at 24 weeks/ earlier if RFs present Abnormal if ≥ 7.8 mmol/l. Fasting/fed status not specified Diagnostic test: 3-h 100-g GTT. GDM if at least two values ≥: fasting 5.8; 1-h 10; 2-h 9.1; 3-h 8.0 mmol/l	7.9% (1748/22178) infants were born to mothers with GDM. After excluding those women with pre-pregnancy DM and those with other autoimmune/endocrine 35 children with Down's Syndrome born (0.2% of births). 7 babies were born to GDM mothers, 28 babies born to non-diabetic mothers Rates of Down's Syndrome were statistically significantly greater in mothers with GDM (3.75 per 1000 infants of diabetic mothers vs 1.36 per 1000 in other infants, $p = 0.02$) RR = 2.75 (95% CI, 1.20 to 6.29) No significant difference in maternal age ($p = 0.67$) or sex distribution ($p = 0.17$) between groups No significant difference in rates of Down's Syndrome in similar age groups (table not included in paper copy though reported to be available online)
Sensitivity and specificity N/A Authors comment that lack of statistical significance in Down's Syndrome rates in different age groups may have been due to small numbers within each maternal age group *Authors' conclusions*: Down's syndrome occurs more frequently in infants of diabetic mothers				
Comments Population in Saudi Arabia studied. Results may differ in different populations Assumes all GDM detected by GCT at 24 weeks As authors comment, small number of babies with Down's Syndrome may have limited power of study to detect statistically significant differences No use of multivariate analysis to adjust for potential confounding factors No mention of whether antenatal screening with termination of Down's Syndrome fetus was available				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Neiger and Coustan, 1991[106]	USA Mean age 27 years No further population data given Women with known DM excluded	106 unselected women No randomisation *Aim*: to evaluate usefulness of repeating GTT if only one abnormal level found	Universal screening of women with 1-h 50-g GCT at 24–28 weeks (fasting status not stated) Fasting 3-h 100-g GTT to those ≥ 130 mg/dl Those with only one abnormal value: fasting ≥ 95; 1-h ≥ 180; 2-h ≥ 155 and 3-h ≥ 140 mg/dl given repeat GTT 1 month later (mean gestational age 34.7 weeks, SD 2.2)	Of 106 women given repeat test, 36 (34%) diagnosed GDM with 2 or more abnormal BG levels No significant differences between those diagnosed GDM and those 'normal' at second GTT in terms of mean gestational age The abnormal value in the first GTT was the fasting value in 7, the 1-h in 14, 2-h in 11 and 3-h in 4 patients 27 patients had again only 1 more abnormal value 3 of these underwent a third GTT, 2 of which had 2 or more abnormal values
Sensitivity and specificity No data available to calculate sensitivity or specificity as no data collected regarding numbers of women screened Comparison of degree of abnormality of first elevated values between those positive on second with those negative. This was not significant				
Comments Using O'Sullivan and Mahan criteria[19] (converted as using plasma) suggest that at least 34% women may have missed diagnosis of GDM Possible explanation offered by authors is that the GTT may not be stable over time No data examining whether fetal/maternal outcomes different between those diagnosed on first GTT and those on second				

All other studies *contd*

Reference	Study population and selection	Types of study	Tests and diagnostic criteria across obstetric centres
Nelson-Piercy and Gale, 1994[113]	North-East Thames Health Region, London, UK Subjects: Clinical Director/Chairman of each Obstetric Department Area has a population with widely differing ethnic and socio-economic characteristics	Not RCT Questionnaire survey conducted as face-to-face interview with Clinical Director/ Chairman of each Obstetric Department *Aim*: to review current practice and diagnosis of GDM 17 centres but one centre had 2 different approaches to screening so counted as 2 units giving 18 units	13/17 centres had a formal policy for screening. Further 3 centres had uniformity of practice Screening tests used: Traditional RFs used in 15 centres. 8 centres considered maternal obesity and 13 centres regarded glycosuria as an RF (7 centres stipulated this had to be present more than once) BG was routinely checked in 6/18 centres. In 3 of these centres BG screen was combined with screening for RFs. 2 centres checked BG only on women with traditional RFs. 10 centres based screening solely on presence of RFs Of the 8 centres using BG, 6 used RBG (cut-offs ranged from 5.8 to 9.0 mmol/l); 1 used a single FBG (cut-off 5.0 mmol/l) and 1 used a 50 g glucose load (cut-off 1-h BG 8.0 mmol/l). Of 6 centres using RBG, 4 checked it once (28–32 weeks), 1 checked twice and 1 checked 3 times Diagnosis: 13/17 units used 75-g GTT; 2 used a 50 g glucose load; 1 used a mixed nutrient meal; 1 used a 1-h postprandial BG. Timing of test varied from booking to after 32 weeks. Some centres repeated test later in pregnancy There was considerable variability in diagnostic criteria and in the number of cut-off levels used Most used a 75-g GTT (13/18). Cut-offs ranged from fasting (5.8–8.0 mmol/l); 1-h (9–11 mmol/l); 2-h (6–9.5 mmol/l) 2 centres used a 50-g GTT with cut-offs: fasting (5.0 and 7.8 mmol/l); 1-h (9.4 mmol/); 2-h (6.0 and 10.0 mmol/l) 1 centre used mixed meal: fasting 5.5 mmol/l; 2-h 7.0 mmol/l 1 centre used postprandial BG: 1-h 9.0 mmol/l
Sensitivity and specificity N/A *Authors' comments*: No attempt was made to verify that stated policies accurately reflected existing practices *Authors' conclusions*: There is no consensus concerning screening and diagnosis of GDM in the NE Thames region. This may reflect uncertainty about the value of screening for this condition			
Comments Demonstrates variability in practice of screening and diagnosing GDM in one region			

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Nielsen *et al*., 1988[139]	Danish women, no demographic information given	1992 women Compared predictability of Lind and Anderson's random glucose test and their clinic's standard test: fasting blood sugar No randomisation to groups, all women had both screening tests Only women who met their previous criteria on fasting glucose test were given GTT	Lind and Anderson's random glucose: Time of testing, within 2 h of eating ≥ 6.1 mmol/l More than 2 h ≥ 5.6 mmol/l All women with FBG ≥ 4.0 mmol/l given 75-g GTT at 24–28 weeks WHO thresholds[24]: GDM, 2-h > 11.0 mmol/l plus one other ≥ 11.0 IGT, 2-h ≥ 11.0 but > 8.0 mmol/l plus one other ≥ 11.0 mmol/l Time of testing between 9 and 36 weeks	22 women exceeded Lind and Anderson's cut-off levels. Of these 11 had a fasting glucose > 4.0 mmol/l and given GTT One of these 11 patients diagnosed GDM RF women also tested with fasting glucose test. 72 had fasting levels > 4.0 mmol/l, 6 of these were diagnosed GDM
Sensitivity and specificity PPV data only available for fasting glucose screen = 9%				
Comments Not stated at what intervals bloods taken for GTT Presuming that women screened positive on random test but negative on fasting test not GDM Wide range of gestational age at testing				

All other studies *contd*

Reference	Factors influencing incidence rates of post-partum DM
O'Sullivan, 1991[208]	Review article Incidence of post-partum DM varied amongst studies, factors that influence this variance: 1. Use of non-standardised diagnostic criteria. For example, the application of the WHO criteria[24] and the NDDG criteria[18] to data of a study showed that the WHO criteria gave an incidence rate 56% higher than the NDDG 2. Exclusion or inclusion of known diabetics 3. The adoption of non-pregnant diagnostic standards 4. Verification of known diabetic patients encountered during follow-up 5. Proportion of women who fail to return to normal in the immediate post-partum period 6. Number of hourly intervals required to be negative: fewer leads to increased incidence 7. Different observational periods 8. Use of self-report measures Of 12 worldwide studies reviewed, only 2 used the same diagnostic criteria. All covered different time spans. The incidences exhibit a range of ~ 19–87% for DM plus IGT and ~ 6–62% DM alone

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Owen *et al.*, 1995[93]	USA 280 clinical directors, response rate 74% (206)	Survey to directors of obstetric and maternal–fetal medicine departments *Aim*: to determine current practice of screening and diagnoses of GDM	Assessed common tests used	Universal screening was undertaken by 96.6% (199), the remainder used various historic RFs Screening: 98.5% used a 50-g oral GCT, only 9.3% required a fasting state. 98.5% obtained 1-h blood sample Cut-offs: 40% respondents used serum levels 140 mg/dl; 27% used serum 135 mg/dl; 16% whole blood 140 mg/dl; 10% whole blood 135 mg/dl Timing: 96% tested 24–28 weeks 92% reported that the presence of one RF would prompt earlier testing Diagnostic testing: 97% used 3-h 100-g GTT Thresholds: 58%: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl 77% within ± 5 of these ranges 96% used 2 or more abnormal GTT values for diagnosis GDM. However, 28% would have used fasting level > 105 mg/dl alone for GDM diagnosis
Sensitivity and specificity No data Some differences noted between the obstetric departments and the maternal–fetal departments. Obstetric directors more likely to use 140 mg/dl for GCT, maternal–fetal 135 mg/dl				
Comments GDM patients: Most used diet control alone Use of different insulin doses, and frequency of monitoring BG stated No question regarding side-effects of tests No cost data Authors conclude high concordance in many aspects, and that fasting glucose seems to be weighted differently than post-ingestion levels in the GTT				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Rajab and Mehdi, 1998[246]	Bahrain, United Arab Emirates No details of population	Not RCT Prospective cohort *Aim*: to compare the outcome of pregnancy in women with GCT screening levels > 7.7 mmol/l and ≥ 8.3 mmol/l 3400 women screened Pregnancy outcomes for the following groups: A. GCT > 7.7 and < 8.3 mmol/l (194 women) B. GCT ≥ 8.3 mmol/l (194 women) A and B were randomly selected C. GCT < 7.7 mmol/l (194 women matched as much as possible for age, parity and weight with group B)	Screening test used: BG 1 h after 50 g glucose load (GCT) given while in fasting state between 28 and 32 weeks. If BG ≥ 7.7 mmol/l underwent GTT Diagnostic test: 3-h GTT Abnormal GTT if BG > 5.5, 9.7 and 5.8 at fasting, 1 h, 2 h and 3 h respectively. (3 values at 4 time periods ??? – not clear which level belongs to which time) IGT if one value on GTT was elevated Abnormal GTT if more than one abnormal value Patients were advised on either diet or medications or both	5.8% (197/3400) considered to have abnormal GTT (criteria not clear) plus further 5.8% (199/3400) considered to have IGT (criteria not clear) Pregnancy outcome No significant difference in pregnancy-induced hypertension between groups (C = 8% vs A = 6% vs B = 13%) Pre-term (< 37 weeks) delivery was significantly more common in B (GCT ≥ 8.3) compared to control C (11% vs 4%; $p < 0.02$). A had similar rates to C (3% vs 4%) Birth weight > 4.5 kg: C = 4%; A = 6%; B = 9% APGAR > 6 at 1 min: no significant differences between groups

Sensitivity and specificity

[calculated by present authors]. Though outcomes not reported in association with GTT results
Assumes GCT ≥ 7.7 mmol/l to be 100% sensitive in detecting those with abnormal GTT
GCT ≥ 7.7 mmol/l in predicting GTT abnormal (2 abnormal values): sensitivity 100% (197/197); specificity 83.5% (2675/3203); PPV 27.2 % (197/725)
GCT ≥ 8.3 mmol/l in predicting GTT abnormal (2 abnormal values): sensitivity 84.3% (166/197); specificity 88.6% (2838/3203); PPV 31.3% (166/531)
GCT ≥ 7.7 mmol/l in predicting any abnormal GTT values: sensitivity 100% (396/396); specificity 89% (2675/3004); PPV 54.6% (396/725)
GCT ≥ 8.3 mmol/l in predicting any abnormal GTT values: sensitivity 80.1% (319/396); specificity 92.9% (2792/3004); PPV 60.1% (319/531)
Authors' conclusion: Study was on a small scale but it suggests that it is possible to raise the cut-off level requiring full GTT from 7.7 to 8.3 mmol/l without a serious adverse effect on pregnancy outcome

Comments

Compares outcome of pregnancy for two different cut-off values of screening
Criteria for classification of abnormal GTT were not evaluated
Characteristics of population were not described
Reason for selection of screening level of 8.3 mmol/l was not stated
Assumes all cases of IGT were detected by GCT screening ≥ 7.7 mmol/l by 32 weeks
Report contains some inconsistencies such as giving 3 values for 4 time periods in criteria used to classify GTT results and in table 2, % with APGAR score < 6 at 5 min is much higher than % at 1 min
Numbers with adverse events in subgroups were small, limiting the ability of the study to detect significant differences
No comment on degree of BG control of those with abnormal GTT
Costs of alternative screening strategies not reported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Reece *et al.*, 1987[187]	USA study No age or demographic details given Unselected women who were already due to undergo screening for GDM	Not randomised, no control 61 patients *Aim*: to compare standard glucose with glucose polymer for use as screening test	All women underwent the two glucose screening tests Screening: CHO diet (300 g) for 3 days. Fasted overnight 50 g glucose polymer, bloods at fasting and 1 h later Within 3 days, 50 g standard glucose, bloods at fasting and 1 h later Thresholds as defined by O'Sullivan and Mahan[19] Positive screenees given diagnostic GTT (glucose load not defined) Thresholds: O'Sullivan and Mahan Time of testing not defined	5 (8.2%) screened abnormal with both glucose 5 (8.2%) with one glucose abnormal (3 polymer, 2 glucose) At fasting mean BG similar. At 1 h BG post-polymer higher than post-standard glucose (but high level of agreement $\kappa = 0.62$, $p < 0.0001$) Incidence GDM = 1.22% (2 patients both of whom had positive screen in both glucose) Of those completing questionnaire (response rate not stated) regarding side-effects 40% experienced nausea and vomiting after standard glucose and 10% after polymer
Sensitivity and specificity No data available to calculate sensitivity or specificity. PPV 20%				
Comments Cost data stated = polymer test 35 cents per test Standard glucose = US$2.50 per test Concern as time of screening not defined, nor age ranges, RFs of women Authors conclude that both tests similar in test results, but disparity in patient satisfaction				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Rey, 1997[247]	Ethnic groups ? Population-based Selection effects (possible biases) Age ranges ? Country	RCT etc. Number of women involved	Screening test used and threshold defining as positive. Fasting or non-fasting Timing of screening test Diagnostic test and thresholds – criteria used for DM (ADA[17], WHO[24]))	Percentages with GDM (if need be give more than one % if different thresholds were used. List all clinical outcomes used and numbers, e.g. birth weight, mode of delivery 9% CS), perinatal mortality, ICU stays, etc. Included adverse effects of tests, e.g. nausea and vomiting after glucose loads
Sensitivity and specificity Record results with CIs if given PPV				
Comments Miscellaneous, e.g. any other comments on quality of study; cost data if given				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Roberts *et al.*, 1997[59]	Belfast, Northern Ireland Northern European Caucasian population Two groups: 1. Unselected. Excluded: multiple pregnancy; pre-existing DM; treatment with steroids/antihypertensive agents. Mean age 27.7 years (range 18–40); mean booking weight 64.6 kg (range 43.6–107.4). Majority booked between 6 and 16 weeks' gestation 2. Selected on positive screening: glycosuria in second fasting sample; FH of DM in first-degree relative; maternal weight > 90 kg; history of congenital malformation/unexplained stillbirth; previous baby weighing ≥ 4.5 kg	*Aim*: to compare responses to a 75-g OGTT and standard breakfast and relate glucose response to maternal morbidity and fetal outcome Group 1 (unselected): 115 women recruited by phone at around 28 weeks. 102 completed both tests Group 2: 936 'high-risk' women screened with 75-g GTT. 117 classified as glucose impaired (WHO[24]), 104 of whom had breakfast/lunch profile	No screening Group 1 (unselected): 75-g OGTT and a 300 cal standardised breakfast at 30–32 weeks' gestation. Tests 1 week apart after overnight fast with order of tests randomised. Breakfast and glucose load had same calorific value Group 2 (selected): 75-g GTT at about 30 weeks. Classified as IGT using WHO cut-off of 8.0 mmol/l Breakfast/lunch profile: VPG before and 2 h after each meal. Both meals contained 300 calories with same nutrient content Diagnostic test: 2-h IGT GTT positive if > 8.0 mmol/l (WHO) Breakfast test positive if maximum glucose > 6.8 mmol/l for meal profile	0% (0/102) in unselected group diagnosed as GDM by around 30 weeks (2-h GTT ≥ 11.0 mmol/l ? WHO) 7% (7/102) had IGT (WHO cut-off 8 mmol/l) 2% (2/102) had IGT Mentions 'a few' drop-outs due to vomiting of glucose load but no numbers given (maximum possible number 13/115 = 11%). Higher rate withdrawals with GTT first Glucose load caused greater rise in PG than standard breakfast (fasting 4.4 vs 4.4; 1-h 7.4 vs 6.2; 2-h 6.1 vs 5.2 mmol/l). Poor correlation between GTT and breakfast values within patients (r = fasting 0.53; 1-h 0.36; 2-h 0.15)

Sensitivity and specificity

Group 1 (102 unselected women):

Maternal complications: 5 had UTI, 7 had pregnancy-induced hypertension; 2 had polyhydramnios. None of these women had IGT/abnormal breakfast tolerance. No significant difference between normal and IGT/normal vs abnormal breakfast profile in gestation at delivery; onset of labour or mode of delivery

Fetal outcome: 1 stillbirth, 2 major congenital malformations. Mothers had normal GTT and breakfast tolerance. 11/11 mothers whose infants were admitted to SCBU had normal breakfast test, and 10/11 had normal GTT

Tests were normal in all mothers whose infants had: transient tachypnoea of newborn (5); hyperbilirubinaemia (10); birth weight > 4.5 kg (3)

Group 2 (104 selected 'high-risk' women):

2-h post-breakfast glucose was highest of 4 breakfast/lunch profiles in 66% of women

No significant correlation between 2-h GTT and maximum 2-h post-breakfast/lunch profile (r = 0.35)

Fetal outcome:

No perinatal deaths. 1 congenital malformation whose mother had normal profile. Mean birth weights in those with maximum profile values above and below 8.0 mmol/l were not significantly different (3634 vs 3706 g)

Authors' comments: Small study. Neither test (GTT/breakfast) was predictive of maternal morbidity/poor fetal outcome

Comments

Small sample size with small number of bad fetal/maternal outcomes

Northern European Caucasian population

Some patients withdrew after 75-g GTT due to vomiting but number not stated

Assumes all GDM and IGT diagnosed by GTT/breakfast around 30 weeks

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Rust *et al.*, 1996[248]	USA No demographic data given Various mean ages given for different groups ranging from 22.7 to 26.7 years (no SD given) No inclusion or exclusion criteria defined	Retrospective study of pregnancy outcome in GDM women as defined by three different diagnostic criteria: standard NDDG[18], Sacks[166], C&C[20] Women also stratified into overweight and non-overweight according to pre-pregnancy body habitus 463 pregnancies identified, data available on 434 Demographic variables investigated included age, race, gravidity and parity. RFs included BMI, total weight gain, time of screening and history of substance abuse (all factors defined in text)	Early in third trimester all women underwent a 1-h 50-g GCT (fasting status not defined). Those with a 1-h glucose of ≥ 140 mg/dl underwent a 3-h diagnostic GTT, as recommended by ACOG guidelines[4] (fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl) Patients divided into 2 groups of four: GDM as defined by NDDG criteria Those who would have had diagnosis of GDM as defined by (a) Sacks and (b) C&C Those with one abnormal value by any criteria Normal values All reviews of outcomes blinded to the study group Sacks criteria[166]: fasting 96; 1-h 173; 2-h 152; 3-h 131 mg/dl C&C criteria[20]: fasting 95; 1-h 180; 2-h 155; 3-h 140 mg/dl	Incidence by NDDG 23.5% (102 women) Incidence by Sacks further 16.4% (71 women) Incidence by C&C further 11.3% (49 women) Incidence of those undergoing GTT, not those screened NDDG women, significantly older, delivered earlier, gained less percentage weight, and gave birth to more infants with hypoglycaemia when compared to group 4 Sacks criteria only significantly older than group 4, C&C criteria only significantly older and greater pre-pregnancy BMI (*p*-values at least 0.01) Total weight gain, birth weight, CS, macrosomia, maternal and neonatal morbidity not significant In groups stratified overweight, non-overweight: overweight women older, gained less weight during pregnancy, had more CS and greater maternal morbidity than non-overweight
Sensitivity and specificity Unable to calculate as numbers screened not given				
Comments Authors conclude that lowering the threshold for diagnosis GDM does not have any discernible effect on perinatal outcome Maternal habitus appears to be a major contributor to macrosomia and the associated risk in morbidity				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Rust *et al.*, 1998[178]	USA Ethnicity of women: 82% black, 16% white Mean age 23.7 years (SD 6.1)	Prospective study, no randomisation All patients underwent both test conditions 475 patients enrolled, 27 failed to complete both testing procedures. 448 analysed *Aim*: to compare a 2-h postprandial screening test with the standard 50-g GCT	Screening test: All women asked to eat 100 g CHO diet and then blood sample taken 2 h after Following this all women also given a 50 g glucose load and blood sample taken 1 h after Those with glucose ≥ 140 mg/dl (ACOG criteria[4]) given 3-h 100-g GTT Diagnosis GDM made 2 or more abnormal values as defined by ACOG criteria[4]	39 women had results on either screening test > 140 mg/dl (2 on 2-h test alone, 29 on 1-h alone and 8 on both) GDM in 16 (3.6%), none from 2-h alone Mean glucose on 2-h 87.1 mg/dl (SD 19.5), on 1-h 102.6 (SD 28.0) Analysis of ROC curve for each test showed the area under the curve (± SEM) for the 2-h at 0.524 ± 0.097 (not significant) and for the 1-h 0.746 ± 0.086 ($p < 0.005$) Comparison of the predictive ability of each test demonstrated that only the 1-h test is an adequate screen for GDM
Sensitivity and specificity For 1-h test: Sensitivity 100%; specificity 95.1%; PPV 43.2%; NPV 100% For 2-h test: Sensitivity 25%; specificity 98%; PPV 40%; NPV 97.3%				
Comments Assumes those screened negative not GDM Time of testing not clearly defined, only states after 20 weeks Although gives details of 100 g CHO diet, does not state details of whether this was adhered to, any problems with it, etc. Describes cost of 1-h 50-g glucola test is 67 cents per bottle with a 12% rate of vomiting. Calculated using 2-h meal would save US$500 in their clinic a year (no calculation data given)				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Sacks *et al.*, 1987[89]	USA Mean age 16.4 years (no range given) 77% had at least one RF: FH, obesity, aged > 25 years, previous pregnancy complicated by macrosomia, congenital malformation, stillbirth or neonatal death	Retrospective study, no randomisation or control 4149 patients, presence or absence of RFs ascertained in 4116 Of these 4116, 3180 had at least one RF *Aim*: to assess effects of RFs on prevalence of GDM	RF patients: first trimester 1-h 50-g (cola) GCT without regard to last meal. If ≥ 135 mg/dl given 3-h fasting 100-g GTT following 3 days high CHO diet. If < 135 mg/dl retested at 24 weeks No RF patients: 24 weeks given 1-h 50-g GCT. If ≥ 135 mg/dl given 3-h 100-g GTT GDM diagnosed on 100-g GTT by Second International Workshop-Conference on GDM criteria[216] with one modification; if fasting glucose ≥ 120 mg/dl test discontinued. If a subsequent fasting glucose ≥ 105 mg/dl diagnosed GDM	Incidence: No RFs 4 patients (0.4%) RF 134 patients (4.2%) 8 patients were diagnosed GDM on the basis of a fasting glucose ≥ 120 mg/dl A linear increase was observed in the proportion of GDM as values on GCT increased 8% patients GDM had screening results 135–139 mg/dl Of 202 patients with positive GCT and normal GTT who were retested at 24 weeks, 33 (16.3%) GDM Of 2040 with normal GCT repeated at 24 weeks, 2.6% (53) GDM Significantly greater yield of positive GCT and GTT in patients with at least one RF Only FH, obesity, macrosomia, maternal age > 25 were significantly associated with GDM
Sensitivity and specificity PPV for RF patients 15% PPV for non-RF patients 34% Unable to calculate sensitivity and specificity because of missing data				
Comments Authors conclude that a threshold of 135 mg/dl required and that selective RFs may be considered in designing a cost-effective screening programme				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Sacks *et al.*, 1989[166]	California 2 groups: 1. Volunteers, predominantly Hispanic women not known to have glucose intolerance. 4 diagnosed with GDM and included in second group for analysis 2. Women with GDM first diagnosed during current pregnancy. Identified by GCT ≥ 135 mg/dl on one occasion followed by diagnostic GTT Total group: Mean age 28 years; nulliparous 39%; first testing at 27 weeks; mean time interval between tests 24 h	Not RCT *Aim*: to evaluate the reproducibility of the 1-h 50-g GCT under clinical conditions Total of 110 women 80 women without GDM 30 women with GDM. None had yet started dietary and/or insulin therapy	Screening test used: 1-h PG after 50-g GCT using commercially prepared cola. All women tested on 2 consecutive days Timing: between 24 and 28 weeks' gestation. Tested without regard to time or time elapsed since last meal. Women advised to have same food and activity level on both days Volunteers with values ≥ 135 mg/dl on either or both tests underwent diagnostic GTT GDM diagnosed on GTT according to Second International Workshop-Conference on GDM criteria[216] Women classified as GDM or control Subgroup classified as straddlers (test results ≥ 135 mg/dl on one day only)	4.8% (4/84) predominantly Hispanic women diagnosed as GDM (Second International Workshop-Conference on GDM) GDM 30 women Control 80 women Subgroup: straddlers 21 women (10 with GDM, and 11 without GDM) GCT values were significantly higher in those with GDM compared to those without GDM ($p < 0.001$ on both days) There was no significant difference in mean values in the straddlers group between those with and without GDM Difference in day-to-day difference in GCT values between GDM and no GDM N/S (24.2 mg/dl vs 17.2 mg/dl, $p = 0.054$) Difference in day-to-day difference in GCT values in straddlers group between GDM and no GDM N/S (31.3 mg/dl vs 47.2 mg/dl)

Sensitivity and specificity
Not calculable
10 women with GDM and 10 without GDM had GCT ≥ 135 mg/dl on only one day
Reliance on a single GCT ≥ 135 mg/dl would have missed 8 (27%) of the 30 cases with GDM. 3 cases of GDM would have been missed on both days
Day-to-day difference in GCT values:
GDM 24.2 mg/dl (range 96–227);
Control 17.2 mg/dl (range 60–189);
Straddlers GDM: 31.3 mg/dl (range 99–169)
Straddlers without GDM: 47.2 mg/dl (range 78–184)
Association of fasting hyperglycaemia (PG ≥ 105 mg/dl) with GCT results in 30 women with GDM: fasting hyperglycaemia present in approximately equal frequency in patients with 2 GCT < threshold (1/3 = 33%); straddled threshold (3/10 = 30%); and exceeds threshold (6/17 = 35%)
Authors' comments: Selection of a 135 mg/dl threshold was somewhat arbitrary; compliance with instruction to have same meals and activity on both days was not closely monitored; reason for variability in results in same subject is not clear
Authors' conclusion: The 1-h GCT is moderately reproducible. Reliance should not be placed on a single normal test result, particularly among patients with RFs

Comments
Predominantly Hispanic population
No baseline comparison of characteristics between groups with and without GDM
GCT without regard to time elapsed since last meal
Assumes all cases of GDM will be detected by GCT ≥ 135 mg/dl on 2 successive days by 28 weeks
Small numbers analysed, particularly in the straddlers group, leading to lack of statistical power to detect any difference
A considerable range was found in the differences between GCT results on consecutive days. No exploration of factors associated with individuals showing marked differences. See 'day-to-day differences' above
The conclusion that results of 1-h GCT is moderately reproducible does not seem supported by the evidence (large range of values of differences in values over the 2 days and reliance on a single GCT ≥ 135 mg/dl would have missed 8/30 cases with GDM), though the advice not to place reliance upon a single normal test result is supported

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Sermer *et al.*, 1994[31]	Toronto, Canada Women aged ≥ 24 years Exclusion criteria: history of DM; seen by physician after 24 weeks; delivered before 20 weeks' gestation Race (3154 women – no data collected at one centre): white 82%; black 5%; Oriental 8%; other 4% Mean age 31 years; mean gravidity 2; mean parity 0.6; FH of DM 15%	Prospective cohort study *Aim*: to assess the impact of time since last meal on GCT; find GCT threshold values that best predict normal/abnormal GTT as defined by NDDG[18] 4274 women had initial screening test (GCT) 3836 (90%) had GTT Grouped according to time since last meal (< 1 h; 1–2 h; 2–3 h; 3–4 h; > 4 h) Groups comparable on baseline characteristics	Screening at 26 weeks with 50 g glucose load, followed by PG test 1 h later (GCT). Time of last meal recorded 90% underwent 100-g GTT at 28 weeks GTT: performed after overnight fast (> 8 h, < 14 h) after 3 days of a 150 g CHO diet and unrestricted exercise GCT positive threshold of 7.8 mmol/l Diagnostic test: GTT positive if at least two blood sugar values exceeded: fasting 5.8; 1-h 10.5; 2-h 9.1; 3-h 8.0 mmol/l (NDDG[18]); fasting 5.3; 1-h 10.0; 2-h 8.6; 3-h 7.8 mmol/l (C&C[20])	3.8% (??/3836) diagnosed with GDM (100-g GTT at 28 weeks) (NDDG criteria) 6.9% (??/3836) diagnosed with GDM (100-g GTT at 28 weeks (C&C criteria) Significant difference in GCT glucose levels depending on timing of last meal ($p < 0.0001$) 20% (769/3836) screened positive (1-h GCT ≥ 7.8 mmol/l)

Sensitivity and specificity

NDDG criteria for GTT at 28 weeks

Overall sensitivity and specificity of screening GCT in diagnosis of GDM at 28 weeks (using 7.8 mmol/l threshold for GCT)
Sensitivity 76.6%; specificity 82.2%; PPV 14.4%
Misclassification: false-positives 17.1%; false-negatives 0.9%; overall misclassification 18.0%
ROC analysis gave optimised cut-offs (maximum combined sensitivity and specificity) as: < 2 h 8.2 mmol/l; 2–3 h 7.9 mmol/l; > 3 h 8.3 mmol/l. 14.9% (572/3836) screened positive (1-h GCT cut-offs as above)
Overall sensitivity and specificity of GCT in diagnosis of GDM at 28 weeks using above optimal cut-offs from ROC curves according to time since last meal and NDDG criteria
Sensitivity 73.8%; specificity 87.4%; PPV 18.7%
Misclassification: false-positives 12.1%; false-negatives 1.0%; overall misclassification 13.1%

C&C criteria for GTT

Overall sensitivity and specificity of screening GCT in diagnosis of GDM at 28 weeks (using 7.8 mmol/l threshold for GCT)
Sensitivity 67.5%; specificity 83.5%; PPV 23.3%
Misclassification: false-positives 15.4%; false-negatives 2.2%; overall misclassification 17.6%
ROC analysis gave optimised cut-offs (C&C criteria) as: < 1 h 7.3 mmol/l; 1–2 h 7.1 mmol/l; 2–3 h 6.9 mmol/l; > 3 h 7.6 mmol/l. Best overall GCT threshold using ROC curves is 7.3 mmol/l
Overall sensitivity and specificity of GCT in diagnosis of GDM at 28 weeks:
Using GCT threshold 7.8 mmol/l: sensitivity 67.5%; specificity 83.5%; PPV 23.3%
Misclassification: false-positives 15.4%; false-negatives 2.2%; overall misclassification 17.6%
Using optimal cut-offs from ROC curves according to time since last meal (C&C criteria): sensitivity 86.4%; specificity 72.7%; PPV 19.1%
Misclassification: false-positives 25.5%; false-negatives 0.9%; overall misclassification 26.4%

Screening

Using GCT threshold of 7.8 mmol/l: 18.5% classified as positive GCT
Using optimised GCT threshold (?NDDG criteria): 13.7% classified as positive
Authors recommend GCT thresholds according to time of last meal of: < 2 h 8.2 mmol/l; 2–3 h 7.9 mmol/l; > 3 h 8.3 mmol/l
Need to be reconsidered if NDDG change criteria for positive GTT

Comments

Prior sample size estimation. Large sample
Bonferroni correction for comparison of multiple baseline characteristics between groups
Follow-up only till 28 weeks – may have missed some who developed GDM later. Assumes all cases of GDM detected by GTT at 28 weeks
Excluded women aged < 24 years/presentation after 24 weeks
19/438 who did not have GTT had GCT glucose > 7.8 mmol/l (significantly higher than those who did have GTT)
Type of meal not taken into account
No costs. Comments that will be reduction in GTT following positive screening test using new thresholds compared to 7.8 mmol/l threshold

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Shah *et al.*, 1982[249]	USA No details of ethnicity, no mean age given	Cohort study 90 women RF women: FH, previous macrosomic infant (> 4000 g), stillbirth, congenital malformation, aged ≥ 25 years, obese (≥ 200 lbs), recurrent vaginitis, glycosuria, polyhydramnios, and an infant believed to be LGA	All women with historical or clinical RFs for GDM given a 50-g GCT. Prior to ingestion blood taken for measurement of HbA_{1C} Those screened ≥ 140 mg/dl given 3-h 100-g fasting GTT, plus 8 negative screenees who had a strong clinical history GTT abnormal if any 2 values >: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl	Incidence of GDM 16.6% (high rate due to selection of high risk women) HbA_{1C} not statistically different between those with GDM and those without
Sensitivity and specificity Sensitivity of HbA_{1C} 27%, specificity 86.5%				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Shivvers and Lucas, 1999[161]	USA population (of those screened > 200 mg/dl) Hispanic 69%, black 19%, white 7%, Middle Eastern 5%. No mean age given, only defines age (of those screened > 200 mg/dl) in terms of > 30 years (47%) Total numbers of women screened not given	Retrospective, descriptive study of previously selectively screened women with historical RFs (FH, previous macrosomia, stillbirth or malformed infant) No randomisation, no controls *Aim*: to evaluate the diagnosis of GDM on a 50-g GCT result > 200 mg/dl Outcomes measured included gestation age at screen, race, age and weight Also birth weight, gestational age at delivery	Screening at 24–28 weeks, with 1-h 50-g GCT. Those with ≥ 140 mg/dl (NDDG criteria[18]) given 3-h 100-g fasting GTT. NDDG criteria, 2 or more: fasting ≥ 105; 1-h ≥ 190; 2-h ≥ 165 and 3-h ≥ 145 mg/dl All women with glucose ≥ 200 mg/dl referred to GDM clinic to have fasting serum glucose screening. If this < 105 mg/dl 3-h GTT performed, if ≥ 105 mg/dl admitted for glycaemic evaluation and given dietary counselling. If persistent ≥ 105 mg/dl prescribed insulin	69 patients had screening test value ≥ 200 mg/dl. Of these medical records were unavailable, 2 were excluded as previously had DM and 4 did not complete a 100-g GTT Of the 59 women left, all had fasting level < 105 mg/dl and classified as normal (n = 11), class A1 GDM (n = 35) (those GDM on GTT), or A2 GDM (n = 13), those treated with insulin (6 of these screened earlier) There was no 50 g screening result that precluded a normal diagnosis The relationship between maternal diagnosis and large for gestation age significantly greater in GDM groups than controls (p < 0.05)
Sensitivity and specificity N/A				
Comments Numbers of women originally screened not given Different times of testing for some of the A2 GDM women A screening result of > 200 mg/dl does not preclude a diagnostic test				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Solomon *et al.*, 1996[250]	USA Nurses between ages 25 and 42 No demographic details included Inclusion: 422 (3.4%) women with reported diagnosis GDM, medical records of 120 of these observed (not stated how selected) = GDM group 100 women control group (randomised) = GDM screening group	Cross-sectional survey, not randomised, no controls Participants from National Health study (116,678 nurses) *Aim*: to see to what extent recommendations are followed in USA and whether NDDG criteria[18] used consistently	Questionnaire regarding use of different tests during their previous pregnancies Defined responders as: 1. Definite GDM if documented 3-h OGTT with NIDDM criteria 2. Probable GDM if physician diagnosis or OGTT modified criteria 3. Possible GDM if elevated screening or other abnormal but non-diagnostic glucose tolerance	3.4% self-reported diagnosis of GDM in previous pregnancy Of 100 GDM screening group – 93 included (3 non-responders, 2 not in defined time period, one in fact had GDM and one could not recall). All reported 2 or more urine tests, 83% reported a 1-h 50-g GCT. Of the 16 who did not, 69% had one or more RFs GDM group, 120 selected, 3 medical records not available and 3 excluded due to molar pregnancy in 2 and multiple birth in 1 64% (73) definite GDM (5 required insulin during pregnancy); 39% (34) probable; 6% (7) possible 93 (82%) had documented 3-h GTT 16% self-report GDM with no GTT GTT in 22 (25%) women with physician diagnosis of GDM failed to meet NDDG criteria
Sensitivity and specificity Survey design, no testing undertaken Reported results defined in descriptive statistics alone				
Comments Authors suggest screening not universal, even > 30 years 17% not routinely screened in manner defined by NDDG, despite presence of RFs in more than 2/3				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Solomon *et al.*, 1997[251]	USA Participants in nurses' health study II (cohort study of 1,116,678 nurses aged between 25 and 42) No demographic details included Inclusion: women with a singleton pregnancy during 1990–94	Not RCT Cross-sectional survey 14,613 women without previous GDM *Aim*: to assess whether RFs for DM may be RFs for GDM	Of 722 women reporting GDM medical records requested for subset of 114 to validate self-reported GDM 100 women not reporting a pregnancy with GDM also used to assess use of screening for GDM	722 (4.9%) reported a diagnosis of GDM. 94% of medical records reviewed confirmed this with a physician diagnosis Relative risks for GDM increased significantly with increasing maternal age, with women > 40 having a 2x increased risk for GDM as compared to women 25–29 years. Unchanged in multivariate analysis adjusting for BMI, FH of DM, ethnicity, parity and pregravid activity Relative risk also increased amongst women with a FH. In multivariate analyses having a mother with DM was associated with a significantly increased GDM risk Women with African-American, Hispanic or Asian ethnicity all had significantly increased age-adjusted relative risks for GDM as compared to white women Average BMI was higher amongst African-American and Hispanic women when compared to white women but this difference did not account entirely for their higher risks of GDM Risks for GDM also increased in women who were current smokers (significant from non-smokers) and in women with increased weight gain in early adulthood Pregravid exercise was associated with a non-significant reduction in risk for GDM
Comments Selective population (health professionals and responders) Small numbers of non-white population Women at high risk may have been screened more in first place Conclusion – these observations may facilitate the identification of women at particular risk for GDM				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Super *et al.*, 1991[159]	Cleveland, Ohio, USA High-risk pregnant patients with one or more of the following: FH of DM in first/second-degree relative; history of GDM; previous neonate > 4.1 kg; maternal weight > 90 kg; maternal age ≥ 28 years; unexplained intrauterine death Sample: age ≥ 28 years 41%; obese 25%; previous GDM 19%	Not RCT Observational cohort *Aim*: to determine whether the diagnosis of GDM can be made early in pregnancy in a high-risk population, to determine the best cut-off values for a 50 g screening test, and to determine whether plasma insulin values or insulin-glucose molar ratios can enhance the properties of this test 43 women enrolled in first trimester (mean 11 weeks); 23 returned for second 50-g GCT in second trimester (mean 20 weeks); 22 returned for third 50-g GCT in third trimester (mean 31 weeks) + 5 had GTT only 32 women enrolled early in second trimester (mean 20 weeks); 16 returned in third trimester (mean 31 weeks) + 5 had GTT only Overall 16% dropped out	Screening test: 50-g oral GCT after overnight fast and 3 days with minimum of 300 g CHO. Fasting, 1-h and 2-h glucose measured. Dietary compliance assessed Abnormal if 1-h PG ≥ 7.5 mmol/l GCT repeated every 10 weeks till delivery/abnormal If abnormal, underwent diagnostic test: 100-g OGTT after same dietary preparation GDM diagnosed if any 2 glucose values from 50-g GCT or 100-g GTT ≥: fasting 5.28; 1-h 10; 2-h 8.62; 3-h 7.78 mm (C&C[20] interpretation of O'Sullivan and Mahan[19])	20% (15/75) GDM in high-risk population (C&C criteria after 50-g GCT/100-g GTT)

Sensitivity and specificity

First trimester screening

ROC curves gave best cut-off value of 7.2 mmol/l for 1-h glucose after 50-g GCT

Sensitivity and specificity of 50-g GCT in first trimester with cut-off 7.2 mmol/l: sensitivity 91% (95% CI, 74 to 100); specificity 88% (95% CI, 77 to 99)

Sensitivity and specificity of 50-g GCT in first trimester with cut-off 7.5 mmol/l: sensitivity 70% (95% CI, 43 to 97); specificity 91% (95% CI, 81 to 100)

Second trimester screening

ROC curves gave best cut-off value of 6.86 mm for 1-h glucose after 50-g GCT

Sensitivity and specificity of 50-g GCT in first trimester with cut-off 6.86 mmol/l: sensitivity 85% (95% CI, 56 to 100); specificity 79% (95% CI, 68 to 90)

Sensitivity and specificity of 50-g GCT in first trimester with cut-off 7.5 mmol/l: sensitivity 71% (95% CI, 35 to 97); specificity 83% (95% CI, 73 to 93)

First trimester: 43 underwent GCT; 12 (28%) were abnormal (≥ 7.5 mmol/l); 6 (14%) diagnosed as GDM. Second trimester 23 underwent GCT; 3 (13%) were abnormal (≥ 7.5 mmol/l); further 2 diagnosed as GDM. Third trimester: 22 underwent GCT; 2 (9%) were abnormal (≥ 7.5 mmol/l); further 3 diagnosed as GDM

Second trimester: 32 underwent GCT; 10 (31%) were abnormal (≥ 7.5 mmol/l); 2 diagnosed as GDM. Third trimester: 16 underwent GCT; 4 (25%) were abnormal (≥ 7.5 mmol/l); further 2 diagnosed as GDM

The ability to diagnose GDM was not enhanced by the use of insulin or insulin-molar ratios

Comments

Ethnicity not reported

Pilot with small sample size

High-risk population

Authors report study limitations as:

Possibility of missing patients with GDM by not testing all women with 100-g GTT. In this study only those with screening values ≥ 7.5 mmol/l underwent GTT

Possibility that some women who dropped out (16%) may have become GDM in third trimester. This would lower the sensitivity with little change in specificity

This was a pilot study in a high-risk population. Results need to be replicated in general population

Authors' conclusion: The diagnosis of GDM can be made in a high-risk population during the first half of pregnancy. Future studies are needed to address the cost-effectiveness and usefulness of such a program in a general obstetrical population and to determine whether earlier diagnosis coupled with various intervention strategies can reduce the persistent morbidity associated with GDM

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Sutherland *et al.*, 1989[63]	Aberdeen, Scotland, UK Subjects: mean age 27 years (range 19–37); mean height 162 cm; mean weight at 26 weeks 68 kg Exclusion criteria: known DM; use of corticosteroids	Not RCT Observational study *Aim*: to compare BG levels after a 75-g OGTT with BG level after a BTT 101 women studied	No screening tests used Both diagnostic tests in all women after overnight fast: 1. 75-g GTT near 26 weeks 2. BTT (453 cal) near 27 weeks CBG levels taken at: fasting, 30, 60, 90 and 120 minutes Based on reported incidence of GDM of 3% Classified BG levels as abnormal if > 97th percentile value 97th percentile levels after 75-g GTT in mmol/l: fasting 5.1; 30 min 10.6; 60 min 11.0; 90 min 9.4; 120 min 8.2 97th percentile levels after BTT: fasting 5.1; 30 min 7.8; 60 min 7.3; 90 min 6.4; 120 min 6.1	3% (3/101) had values > 97th centile on both GTT and BTT 4% (4/101) had FBG ≥ 97th percentile The above values do not appear consistent but difficult to extract results 8.9% (9/101) ≥ 97th percentile on GTT 7.9% (8/101) ≥ 97th percentile on BTT

Sensitivity and specificity
N/A
3 babies suspected of showing clinical features of diabetic fetopathy (these features were not defined): one had high FBG, none had abnormal 2-h GTT value (WHO criteria[24]). All had abnormal BG during BTT
Correlations between maternal characteristics (waist, hip and waist:hip ratio, age, height, weight and BMI after delivery) vs BG values. Significant correlations noted between BG and maternal age throughout the GTT but only with the 60 and 90 min BG during the BTT; height inversely correlate with BTT but not with GTT; maternal weight with BTT but not GTT; FH of DM with 2-h BTT (p = 0.043)
Authors' comments: Consider there is a need for a radical reappraisal of the approaches to evaluating maternal metabolism in human pregnancy
Data suggest that OGTT result may be oversensitive to the effects of maternal age
The 2-h GTT value failed to identify any of the 3 women with possible diabetogenic fetopathy
If BTT is to be used maternal height may have to be taken into consideration
Authors' conclusion: The BTT has a valuable place in the investigation of maternal metabolism and merits fuller investigation to clarify its place as a routine investigative procedure

Comments
Authors comment that 'many pregnant women in the fasting state find the solution (75-g GTT) unpalatable' and that this results in a high default rate at appointments for re-testing. Rates of intolerance not reported. Reports that 101 women participated and 75 women refused to participate
The outcomes associated with BG levels ≥ 97th percentile in each test were not clearly presented
Criteria used to classify as 'significantly impaired CHO metabolism' assumed incidence of GDM to be approximately 3%
It was not reported whether women with abnormal test results received any treatment for the remainder of the pregnancy
Small sample size with consequently very small number of adverse fetal/maternal outcomes reported thus limiting power of study
The statistical significance of difference in rates of adverse outcome associated with abnormal BG values was not reported
Aberdeen test meal (BTT): protein 15.6%; fat 17.8%; CHO 61.3%; fibre 5.3%; energy 453 kcal

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Swinker, 1983[191]	West Virginia, USA Family practice centre No characteristics of the study population given	Comparison of RF screening with GCT including cost. (However, RF sensitivity was only considered retrospectively) Original sample of 60 pregnancies; 50 underwent the GCT screening (dropouts were 8 who did not take the test, one diabetic woman and one with a history of hydramnios who was given the GTT directly: all were excluded from the analysis). The remaining 50 were given the GCT. Of 16 positive screens only 15 were given the diagnostic test; no mention of missing woman	Patients were screened (non-fasting) using the 1-h 50-g GCT; those with values > 130 mg/dl were referred for a GTT. GTT criteria for GDM were one or more of the following values: > 100 mg/dl initially, > 200 mg/dl at 1 h, > 150 mg/dl at 2 h, and > 130 mg/dl at 3 h The results of the GCT were compared against women with RFs retrospectively; those women with RFs were not given the GTT	Incidence of GDM in the study group was 6% (3/50) There were 32% (16/50) positive screens and 3 women diagnosed with GDM 485 (24/50) of the women had at least one RF Of the 3 women with GDM, only two had RFs The cost of screening and diagnosis (with the GCT and GTT) used in this study was given as US$520 (or US$10 per patient). Cost to detect one case of GDM was given as US$173. (GCT glucose load at US$1 each and lab tests at US$4 each; GTTs at US$18 each) This compares with 24 GTTs for those with RFs at US$430 Author concludes that the GCT-based screening was only slightly more expensive but RF-based would have missed one of the women with GDM No adverse effects documented
Sensitivity and specificity 50-g GCT [≥ 130 mg/dl] PPV 20% (3/12); sensitivity 100%; specificity 74% (34/46) Notes: Figures assuming there were no women with GDM below the 130 mg/dl threshold. Also skewed by one drop-out after screening stage: woman had positive screen but was not given diagnostic test				
Comments Low numbers – only 50 women screened Only positive screens given diagnostic test so incidence may be higher if some GDM undetected. Assumes combination of both tests = 100% sensitivity Women with RFs not given diagnostic GTT so incidence in this group uncertain; assumes no further GDM women would be identified (and that none had been missed with the GCT-based approach) Sensitivity etc. skewed by one woman who screened positive but was not given GTT so presence of GDM unknown Costs significantly outdated and in US dollars; study run between 1981 and 1982, no date given for costs Costs appear to only include materials and lab tests – no administration or other staffing				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Tardioli *et al.*, 1993[98]	Italy Demographic details of women not given, study states that population representative Median age 26 years (range 25–30) for GDM women and 29 (range 19–41) for non-GDM women	Pilot study to establish guidelines for future screening of women for GDM, unselected consecutive women, 161 women	No screening test used, 141 women had a 3-h 100-g GTT between 24 and 28 weeks. Diagnosis made by O'Sullivan and Mahan criteria[19] (no further description given). Reasons for 20 women not having GTT not given Data also taken on delivery and newborns	8 women had GDM (reported 5.6% incidence but this from 141 not total sample of 161) GDM women had more RFs than non-GDM women: FH of DM, BMI > 25 and previous obstetric complication however, some data appears to be missing in the tables as only 7 GDM women are given and 97 non-GDM women Newborn weight and rate of CS were also reported to be higher amongst GDM women, again not all data appears to be reported. None of these differences were tested statistically
Sensitivity and specificity No screening test used				
Comments The reporting of the data in the text is not consistent The paper also states that newborn complications are more apparent in GDM women but does not report what these complications are				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Thaisz *et al.*, 1993[252]	Budapest, Hungary Attendees at antenatal clinic between 1989 and December 1992 No details of sample reported	Not RCT *Aim*: to screen all pregnant women for GDM and properly care for women with IDDM 4676 women screened for GDM Study reports on 162 women with diabetic pregnancies (69 with GDM plus 55 with pre-existing DM)	Screening test used: test breakfast with 75 g CHO on initial visit. If 2-h postprandial BG < 7.5 mmol/l screening repeated at 24–28 weeks. If BG between 7.5 and 9.0 mmol/l underwent GTT Not stated what happened to those with BG > 9 (? classified as GDM) Diagnostic test: 3-h 100-g GTT till Jan 1992, then 75-g GTT used Those with GDM and IDDM had weekly medical checks, fortnightly obstetric checks, monthly ultrasounds, ophthalmic checks and urine culture every 3 months Overweight patients advised restricted energy intake (25 kcal/kg ideal body weight) in 5 portions /day Insulin started if diet not satisfactory	2.67% (124/4676) were diagnosed as GDM (criteria not specified) Normoglycaemia (HbA_{1C} 5.4 ± 0.4%) was achieved in all cases of GDM Maternal complications in diabetics (GDM + pre-existing IDDM) Hypertension 22% (35/162) Vaginal infections 13% (21/162) Bacteruria 10% (17/162) IDDM Retinopathy in 13% (7/55) of those with IDDM Neuropathy in 4% (2/55) Infant morbidity Macrosomia 9.6% in GDM and 10.5% in IDDM Prematurity rate 3.75% No malformations or fetopathies were found
Sensitivity and specificity N/A *Authors' conclusion*: This screening and treatment method seems to meet international standards and could be adopted nationwide The validity of this screening method was not evaluated				
Comments More a description of management of GDM/IDDM than an evaluation of screening. Outcomes of women with negative screening/positive screening and negative GTT were not reported Criteria used to diagnose GDM were not stated Characteristics of population not described Distribution of gestational age at which first screening was performed was not reported Results from two different glucose loads for GTT were combined Full details of treatment of diabetics were not described, e.g. indications to start insulin therapy Outcomes such as prematurity and macrosomia were not defined				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Truscello *et al.*, 1988[122]	USA Teenage pregnant women, white 32%, Mexican-American 39%, black 22% and other (mainly Indochinese) 7%, mostly of low socio-economic family	Cohort observational study, no randomisation or control 137 consecutive teenage women between 12 and 18 years *Aim*: to determine the frequency of GDM in pregnant teenagers	24–34 weeks all had 1-h 50-g GCT with glucola. A value ≥ 140 mg/dl led to 3-h 100-g fasting GTT. Diagnosis GDM was based on the O'Sullivan and Mahan criteria[19] Women split into 3 groups on mean glucose value of GCT: Group A > 1 SD below mean (60–82 mg/dl), Group B ± 1 SD of the mean (83–125 mg/dl) and Group C > 1 SD above mean (126–180 mg/dl). The 1-h screening values were correlated with maternal pre-pregnancy weight, BMI, weight gain during pregnancy, maternal morbidity, placental weight, infant birth weight and infant morbidity	8 teenagers (5.8%) had positive screening tests and 2 had GDM (incidence 1.4%) Maternal variables between 3 groups: no differences in age or gestational age, but patients in group A had the lowest pre-pregnancy weight ($p < 0.01$) and a lower BMI ($p < 0.01$) than other two groups. Group C teenagers gained the most weight during pregnancy ($p < 0.05$) Maternal and neonatal outcomes: no difference in maternal morbidity, gestational age, infant birth weight, APGAR scores, or infant morbidity between 3 groups
Sensitivity and specificity No data to calculate sensitivity or specificity. PPV 25%				
Comments The authors report different distributions of the mean value of glucose in each of the 4 racial groups All teenagers gained more than the recommended weight during pregnancy, which may be associated with GDM or large birth weight of babies There was a high incidence of macrosomia in all 3 groups Authors conclude that GDM screening should be given to all pregnant teenagers				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
van Turnhout *et al*., 1994[253]	Rotterdam, The Netherlands Inclusion criteria: pregnant women with one or more of the following RFs: PH GDM; previous macrosomic or hypoglycaemic infant; positive FH; age ≥ 35 years; obesity; recurrent glycosuria; accelerated growth in present pregnancy Excluded: known type 1 or type 2 DM; women with multiple pregnancy Subjects had median gestational age = 24 weeks (range 20–35)	Not RCT Observational study *Aim*: to evaluate and compare the 50 g glucose load (GCT) and home-monitoring glucose profile (HGP) 415 women	Both screening tests used in all subjects using capillary blood: 1. 1-h 50-g GCT after 20 weeks' gestation without prior glucose loading or fasting 2. Samples for HGP collected the next day (1 h after breakfast, lunch and dinner) Women classified into 3 groups: 1. Both GCT and maximal HGP ≤ 7.0 mmol/l (302 women): normal 2. Both GCT and HGP > 7.0 mmol/l (19 women): dietary measures alone/ combination with insulin. Treatment aimed at keeping postprandial BG ≤ 7.0 mmol/l 3. Either GCT or maximal HGP > 7.0 mmol/l underwent second HGP2. Treated if HGP2 > 7.0 mmol/l (94 women)	4.6% (19/415) diagnosed as GDM (both GCT and HGP > 7.0 mmol/l) 2.4% (10/415) diagnosed as GDM (GCT > 7.8 mmol/l and HGP > 7.0 mmol/l) Poor correlation between GCT and HGP (r = 0.37) Poor correlation between maximal values of first and second HGP (r = 0.32) Fetal outcomes related to glucose test results: see below

Sensitivity and specificity
(Results as reported)
At GST threshold of 7.8 mmol/l and HGP threshold of 7.0 mmol/l: sensitivity of GST 27%; specificity 89%. Likelihood ratio for positive test 2.4; ratio for negative test 0.8
Sensitivity and specificity of first HGP for HGP2: sensitivity 26%; specificity 84%. Likelihood ratio for positive test 1.6; ratio for negative test 0.9
Fetal outcomes related to glucose test results (statistical significance not reported)
1. GCT normal, HGP normal (n = 302):
infant > 90th centile: 8%; instrumental delivery: 17%; 5-min APGAR < 7: 2%; perinatal deaths: 1%
2. GCT normal, HGP normal (n = 18):
infant > 90th centile: 6%; instrumental delivery: 22%; 5-min APGAR < 7: 0%; perinatal deaths: 0%
3. GCT normal, HGP normal (n = 76):
infant > 90th centile: 8%; instrumental delivery: 18%; 5-min APGAR < 7: 3%; perinatal deaths: 0%
4. GCT normal, HGP normal (n = 19):
infant > 90th centile: 16%; instrumental delivery: 21%; 5-min APGAR < 7: 0%; perinatal deaths: 0%
Authors' conclusion: The 1-h 50-g glucose screening test discriminates poorly between pregnant women with and without postprandial hyperglycaemia
Apparently neither a single GCT nor a single HGP is a reliable indicator of hyperglycaemia in women at risk for GDM

Comments
Subjects all had at least one traditional RF for GDM. Other characteristics of subjects not described
No costs
Statistical significance across groups for fetal outcomes not reported and small event rates in these groups limiting the evidence

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Verma *et al.*, 1997[254]	Alberta, Canada Population mostly young, single gravidas of North American aboriginal ethnicity Inclusion criteria: term (≥ 36 weeks) newborns with recorded birth weight, recruited within 24–48 h of delivery; singleton pregnancy; documented results of maternal GCT at 24–28 weeks, absence of congenital anomalies or infections Exclusion criteria: women with pre-existing medical conditions known to have effect on fetal growth (hypertensive disorders, DM, other chronic medical illness)	Not RCT Retrospective case records review with controls *Aim*: to examine the relationship between newborn macrosomia and PG profile in both GCT pos/GTT neg group and GDM group 209 macrosomic newborns (cases) and 791 non-macrosomic (controls) Association between birth weight and the following was examined: GTT glucose; GTT fasting, 1-h, 2-h and 3-h; plus for GDM group: average of fasting (AF); and average of postprandial (APP)	Screening tests used: 50-g GCT at 24–28 weeks. No further details Diagnostic tests: two abnormal values on 100-g GTT (NDDG criteria[18]) Therapy for GDM according to SOCG and ACOG.[4] Diet initially. Insulin used if FBG > 5.5 mmol/l and/or 2-h postprandial > 6.5 mmol/l on 205 of occasions Cases classified as: GCT pos/GTT neg (113 cases) GDM (50 cases)	GDM rates N/A 26% (13/50) of GDM received insulin Statistically significant correlation between birth weight as a continuous variable and GCT, AF and APP values for the GDM group No significant correlation for GCT or any GTT values in the GCT pos/GTT neg group Birth weight of infants were similar in both groups (3706 g vs 3515 g)
Sensitivity and specificity N/A Difference in AF and APP values in third trimester of GDM subjects between macrosomic and non-macrosomic was statistically significant but the authors consider the clinical significance of this difference (0.6 mmol/l) to be doubtful Those in the diet + insulin group were significantly heavier pre-pregnancy than subjects in the diet only GDM group (80.4 kg vs 64.4 kg: $p < 0.05$) *Authors' conclusion*: Data failed to support a significant relationship between macrosomia and PG concentrations				
Comments Nature of relationship between birth weight and glucose levels was not presented graphically (i.e. whether linear or not) No investigation of/adjustment for potential confounding factors was undertaken Threshold for positive GCT was not reported Assumes all cases of GDM diagnosed by positive GCT and GTT by 28 weeks The value of using various glucose values as predictors of fetal outcome was not evaluated				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Watson, 1989[155]	New York, USA Consecutive prenatal registrants from military dependent population. No further details Excluded: late entry (after 20 weeks); known insulin dependent diabetics; 8 patients on terbutaline for tocolysis. Ultrasound scanning early in second trimester to confirm gestation	Cohort study *Aim*: to quantify the effect of advancing pregnancy on the 50 g oral glucose screening test for GDM All 550 women screened	Screening test: 50 g oral glucose load at 20, 28 and 34 weeks' gestation regardless of previous dietary intake. If test positive (1-h serum glucose ≥ 140), proceed to follow-up 3-h 100-g GTT. If initial GTT negative, repeat GTT in third trimester GTT after high CHO diet for 3 days followed by overnight fast before the test Diagnostic test: 3-h GTT defined as abnormal if any two values of PG ≥: fasting 105; 1-h 190; 2-h 165; and 3-h 145 mg/dl (modified from whole blood values of O'Sullivan and Mahan[19])	27/550 (4.9%) diagnosed with GDM by 3-h GTT (O'Sullivan and Mahan) Perinatal outcomes reported for 3 women who screened negative at 28 weeks (GCT values between 120 and 128 mg/dl at 28 weeks) but positive at 34 weeks with positive GTT (one LGA 4340 g and one hypoglycaemic in immediate neonatal period) Linear increase in PG test value with advancing pregnancy No adverse effects documented
Sensitivity and specificity Screening (1-h 50-g GCT ≥ 140) Predictive value of screening GCT at 20 weeks (assuming follow-up negative initial GTT with further GTT in third trimester): sensitivity 33% (9/27); specificity 95% (498/523); PPV 26% (9/34) Predictive value of screening GCT at 20, 28 and 34 weeks (assuming follow-up negative initial GTT with further GTT in third trimester): sensitivity 100% (reported as 19.6%); specificity 79% (412/523); PPV 19.6% (27/138) ASSUMES SCREENING DETECTS 100% OF GDM Predictive value of screening GCT at 20 and 28 weeks: sensitivity 89% (24/27); specificity 87% (453/523); PPV 26% (24/94) Positive screening tests in 25% (138/550) of population at any time in pregnancy 33% (9/27) cases of GDM detected by 20-week GCT (3 immediate positive GTT, 6 delayed) 56% (15/27) cases of GDM positive screen at 28 weeks after negative screen at 20 weeks (9 immediate positive GTT, 6 delayed) 11% (3/27) cases of GDM positive screen at 34 weeks after negative screen at 28 weeks (3 immediate positive GTT) Authors suggest the following regime for screening: screening all patients at 24–28 weeks without regard to previous dietary history taking 1-h serum BG threshold post-50 g glucose load of 140 mg/dl or greater. In patients with post-screening glucose of 120–139 mg/dl at 28 weeks, rescreening later in third trimester will reveal some previously undiagnosed GDM. In all patients with positive screening test and immediate negative GTT a follow-up GTT later in pregnancy is justified			
Comments No details of characteristics of population Only those with serum glucose 140 mg/dl or more on screening at 20, 28 or 34 weeks underwent diagnostic GTT. Assumes screening (GCT ≥ 140 mmol/l) detected all cases of GDM Screened on 3 occasions Follow-up till 34 weeks. May underestimate prevalence if some developed GDM after 34 weeks Those with initially negative GTT following positive screen had further GTT in third trimester No costs Estimates that 16% of population would require rescreening using guidelines suggested NB: Note rates of delayed GDM diagnosis in those with initially negative GTT following positive screening test Authors suggest that disregarding dietary history for screening leads to greater convenience and acceptability to patients that outweigh advantages of testing in the fasting state			

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Watson, 1990[135]	Germany Military dependants, unrestricted access to medical care without monetary cost No further demographic details given Those with previous DM excluded	500 unselected women, no randomisation *Aim*: to compare random urine glucose measurements with the GCT as a predictor for GDM	All women given random urinalysis for glucose at each antenatal visit (mean 10.8, SD 2.6). Diagnosis glycosuria if trace, 1+, 2+ or 3+ found on at least 2 visits. Severe glycosuria if ≥ 2+ on two visits At 28 weeks (no range given) 50-g GCT without regard to ingestion state. Threshold ≥ 140 mg/dl Diagnostic test fasting 100-g GTT, after 3 days high CHO diet Thresholds 2 or more values: fasting 105; 1-h 190; 2-h 165 and 3-h 145 mg/dl	22 (4.4%) incidence of GDM 85 (17%) showed glycosuria and 19 (3.8%) severe glycosuria 10 patients with glycosuria with GDM (6 glycosuria, 4 severe glycosuria) No adverse effects documented
Sensitivity and specificity Any degree of glycosuria: sensitivity 27.3%; specificity 83.5%; PPV 7.1% Severe glycosuria: sensitivity 18.2%; specificity 96.9%; PPV 21.1% Combined: sensitivity 0.45%; specificity 81%, PPV 9.6%				
Comments Severe glycosuria more common in women with GDM ($p < 0.05$) but only accounts for 18% Most women with GDM did not have detectable glycosuria Higher than average incidence GDM Authors conclude that routine random urine testing is a poor screening method but recommend that those classed as severe glycosuria before 24 weeks should have an earlier 50-g GCT				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Wein *et al.*, 1997[83]	Mixed race Australians (Northern European 52%; Asian 29%; Mediterranean 9%; Middle Eastern 5%; Indian 4% and African 1%) Mean age 31–33 years All women diagnosed with GDM between 1981 and 1996	Cohort study Not randomised Compared women without postnatal DM and those with Screening GTT in 57,563 women (71% of all pregnancies) Postnatal GTT in 2957 (69.7%) of all those with GDM	Antenatal testing at 26–30 weeks with 50-g fasting GTT Testing 0, 1 and 2 h Defined as 1-h ≥ 9.0 mmol/l and 2-h ≥ 7.0 mmol/l Graded into 1: as above 2: 1-h ≥ 10.0 + 2-h ≥ 7.8 3: 2-h 8.9–12.1 4: fasting ≥ 7.8 or 2-h ≥ 12.2 Postnatal Tested at 6–8 weeks Fasting 75-g GTT DM diagnosed if fasting ≥ 7.8 mmol/l or 2-h ≥ 12.2 mmol/l (WHO[24])	Incidence: 5.3% of all confinements and 7.4% of those in which glucose tolerance was tested NIDDM diagnosed postnatal in 55 women (+ 2 ketoacidotic and 2 ongoing hyperglycaemia postnatal) = prevalence of 2.0% Postnatal DM in: 0.1% women grade 1; 2.0% grade 2; 4.1% grade 3; 40% grade 4 Significant predictors DM: Severity of GDM, Asian origin, the 1-h glucose value at antenatal GTT. For every 1 mmol/l ↑ odds of DM ↑ by 47% 10.6% who had insulin therapy had postnatal DM but not significant predictor
Sensitivity and specificity No screening test used				
Comments All patients received GTT only Not all GDM women followed up No cost data available Authors conclude that postnatal GTT indicated in all GDM cases				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Weiner *et al.*, 1986[167]	Iowa, USA Mean age of women 25.1 years (SD 5.2) 94% of the women were white, 80% were married and 72% completed to at least year 12 of school	Study examining the cost-effectiveness of screening in terms of per case of GDM identified by comparing 1-h and 2-h GCT values *n* = 798	All women were given the 1-h 50-g GCT 98% were tested between 27 and 29 weeks Women were told to fast for 3 h before the test Women were asked to remain for a second hour to obtain a 2-h reading for the GCT If the 1-h GCT exceeded 139 mg/dl, patients were given the 3-h 100-g diagnostic GTT	22.3% (176) women had a positive 1-h screen at > 139 mg/dl. Of these, 23 had positive GTTs giving an incidence of GDM of 2.9% (8 further positive screens were not given GTTs) 43% of women (342/790) complied with the request to have a 2-h GCT 5 women with positive screens had to be excluded as they were not given a GTT. There were 69 positive 2-h screens at > 115 mg/dl and 55 at > 117 mg/dl. In the same subgroup there would have been 83 using the 1-h > 139 mg/dl threshold. 9 women with GDM were detected by each method. (An incidence of 2.7% in the screened population.) None of the women with GDM identified in the whole population group had a 2-h GCT below 118 mg/dl in this subgroup. There were no significant differences between the subgroup and original population. There was a statistically significant reduction in the number of GTTs required using the 2-h GCT compared to the 1-h GCT Cost of a 1-h screen was given as US$7.25 whilst the cost of a GTT was US$64.00 (including US$24 lab charge and US$40 clinic charge for nursing time) Cost per case of GDM detected was US$662 using a 2-h GCT and a cut-off of 117 mg/dl. Using a cut-off of 114 mg/dl it was US$762. Cost per case using the 1-h GCT values was US$866 for > 139 mg/dl and US$699 for > 149 mg/dl. Using RFs, 77% of women identified with GDM would have been missed and the cost per case of GDM detected would have been US$1805 To test the validity of the results, a second screening procedure was initiated using 190 separate women. The same number of women was again detected by both 1-h and 2-h screening tests. Costs per case of GDM detected were US$1215 using the 1-h GCT and US$831 using the 2-h GCT No adverse effects documented
Sensitivity and specificity 1-h GCT [≥ 140 mg/dl]: sensitivity 83% (19/23); specificity 87% (669/767); PPV 16% (19/117) 2-h GCT [≥ 118 mg/dl]: sensitivity 100% (9/9); specificity 85% (282/333); PPV 15% (9/60) Validity test (*n* = 185) 1-h GCT [≥ 140 mg/dl]: sensitivity 75% (3/4); specificity 82% (148/181); PPV 8.3% (3/36) 2-h GCT [≥ 118 mg/dl]: sensitivity 75% (3/4); specificity 92% (166/181); PPV 16.7% (3/18) All calculations assume all women with GDM detected				
Comments 2-h GCT cut-off levels used (114 and 117 mg/dl) selected to identify all women with GDM Incidence uncertain as only those women with a 1-h GCT above 139 mg/dl were given the diagnostic GTT (applies to both groups) Validity tests led to higher costs per case of GDM detected and lower PPV for 1-h GCT				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Williams *et al.*, 1999[255]	USA Mixed ethnic groups: white, Hispanics, native Americans, Asians and African-Americans No detail of what proportion of each group RFs based on ADA: non-white, BMI > 27, no FH of DM and older than 25 years *Aim*: to investigate what proportion of women would have missed screening if ADA selective screening guidelines used (Group 1) and the percentage of women with GDM who may have been undiagnosed using the same criteria (Group 2)	Review of medical records from a selected sample without RFs and those with GDM Of 25,118 deliveries, a sample of 250 records of white women aged under 25 were randomly chosen from a total of 4629 of this classification. Of these 250, 224 also had BMI < 27 and no FH and were used for the analysis (Group 1) Second group of records of 200 women with GDM also reviewed	All women screened with 1-h 50-g GCT between 24 and 28 weeks without regard to last meal If glucose ≥ 140 mg/dl given a 3-h fasting 100-g GTT. Diagnostic thresholds: fasting 105; 1-h 190; 2-h 165 and 3-h 145 mg/dl (NDDG criteria[18]) No detail confirming that the sample selected had the standard screen and diagnostic test	Group 1: Proportion of women with all four low RFs was 11.1%. Therefore these women would not have been offered GDM screening if the ADA selective screening guidelines were used Group 2: 141 medical records had details on all RFs (59 excluded) The prevalence of women with none of the four RFs, yet with a diagnosis of GDM = 5 (4%). Therefore these women would not have been screened by ADA selective screening guidelines
Sensitivity and specificity No data to calculate sensitivity or specificity				
Comments The authors point out that the time involved in applying the new criteria outweighs the small cost benefit of not testing 11% of the population In an addendum, the authors also assessed, from another sample using similar design, the adoption of selective screening by the same four factors along with new criteria from the Fourth International Workshop-Conference on GDM,[21] which were: history of GDM or previous poor obstetric outcome. Suggest that fewer women will be missed, however likely that less women will have none of the six criteria				

All other studies *contd*

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Young *et al.*, 2000[256]	Temple, Texas, USA Inclusion criteria: women who had 2 deliveries between 1994 and 1997 with records regarding GDM screening in both pregnancies Population predominantly white (60.6%); Hispanic (21.5%); African-American (17.3%) Exclusion criteria: multiple pregnancies or history of DM Most (92%) had at least one RF for GDM Index pregnancy: mean age 21.6 years; BMI 24; 1-h post-GCT glucose 108 mg/dl Subsequent pregnancy: mean age 24.2 years; BMI 25; 1-h post-GCT glucose 111 mg/dl	Not RCT Retrospective review of case records *Aim*: to evaluate the likelihood of a woman without GDM in an index pregnancy developing abnormal CHO metabolism in a subsequent pregnancy Records from 381 patients examined	Screening test used: 1-h post 50-g GCT with threshold ≥ 140 mg/dl in all women Fasting/fed status and timing not reported Diagnostic test: 3-h 100-g GTT. GDM diagnosed if two or more values ≥: fasting 105; 1-h 190; 2-h 165; 3-h 145 mg/dl (ADA criteria[17])	Quotes local overall rates of GDM as 4% None of the 381 women with normal glucose screening in initial pregnancy had GDM in subsequent pregnancy (0/381) 12% (45/381) women with normal glucose screening in initial pregnancy had abnormal GCT (≥ 140 mg/dl after 50 g glucose) in subsequent pregnancy
Sensitivity and specificity N/A Potentially predictive values for an abnormal 50-g GCT in subsequent pregnancy evaluated using multiple regression (337 women): only the serum glucose for the index pregnancy was predicative ($p = 0.001$). The following were not significantly associated: race, historical RFs, age or weight gain at index pregnancy; age or BMI in second pregnancy; or rate of change of BMI between pregnancies *Authors' comments*: Absence of GDM in subsequent pregnancies may be due to relatively short 4-year study period *Authors' conclusion*: Despite a high rate of RFs for GDM, women in this population without GDM in an index pregnancy have a minimal risk (< 1%) that GDM will occur in a subsequent singleton pregnancy within 4 years. This may help in determining whether women should undergo screening for GDM				
Comments Timing and fasting/fed status when screening undertaken was not reported Timing of diagnosis of 'no GDM' in index pregnancy was not reported. Only those with abnormal GCT had GTT ? Will these results hold for all subsequent pregnancies independent of the time interval since first screened				

Reference	Study population and selection	Types of study	Test used and diagnostic criteria	Outcome
Zhang *et al.*, 1995[189]	China Chinese population, mean age 29.1 (SD 7.5) years	Prospective study, women randomly chosen 220 women	Third trimester testing (mean 31.8, SD 2.7 weeks) 50-g fasting GCT, threshold ≥ 7.8 mmol/l Diagnostic test: 100-g fasting GTT. GDM as defined by Second International Workshop-Conference on GDM[216]	Incidence 5% No clinical outcomes used No reports of adverse effects of test
Sensitivity and specificity No data to calculate sensitivity or specificity Also tested CPG and correlated this with venous plasma ($r = 0.87$, $p < 0.01$)				
Comments No cost data given Paper written in Chinese				

Appendix 3

Summary of GDM costing in the literature (US$ unless stated)

Study	Base year (published)	GCT threshold (1-h values, mmol/l)	Population	Cost/ 50-g GCT	Cost/ 100-g GTT	Cost/case of GDM detected	Costs include
Lemen *et al.*[192]	?? (1998)	7.8		25	50	2733	Lab. costs
		7.8	RF			1258	
Moses *et al.*[244]	?? (1997)	N/A			A$9.90 (75 g)		?? Materials and lab. only
Neilson *et al.*[165]	?? (1991)	8.3		17.75	59.15	722.31	?? Direct costs only
Hong *et al.*[103]	?? (1989)	7.2	All	3.5	15	184	?? Materials only
		7.8	All			158	
		7.2	Aged ≥ 25 years			122	
		7.8	Aged ≥ 25 years			106	
		7.2	Aged ≥ 30 years			120	
		7.8	Aged ≥ 30 years			102	
Coustan *et al.*[4]	?? 1985 (1989)	7.2		from Marquette *et al.*, 1985[124]	from Marquette *et al.*, 1985[124]	249	Used Marquette *et al.*, 1985[124] costs
		7.8				222	?? Direct costs
		8.3				722	
		7.2	Aged ≥ 25 years (younger if RF)			215	
		7.8	Aged ≥ 25 years (younger if RF)			192	
		??	Aged ≥ 30 years (younger if RF)			190	
Kirkpatrick *et al.*[154]	?? (1988)	7.75 or RF			50-g GTT	256	Lab. costs, glucose loads, phlebotomist's salary
Weiner *et al.*[167]	?? (1986)	7.8		7.25	64	866/1215	Direct + indirect costs
		8.3				699	
		2-h values:					
		6.4				762	
		6.5				662/831	
		N/A	RF			1805	
Lavin[148]	1985 (1985)	8.3		4.75		328.96	Direct costs
Marquette *et al.*[116]	?? (1985)	7.2		2.45	11		?? Direct costs
Swinker[191]	?? (1983)	7.2		5	18	173	Materials and lab. only

Appendix 4

Studies that gave a diagnostic test to all patients (in reverse chronological order)

Study	*n*	Test (GCT, FPG, etc.)	Threshold ≥ mmol/l	All patients, by age group or RF?	Fasting?	Sensitivity (%)	Specificity (%)	PPV (%)	Comments
Perucchini *et al.*, 1999[104]	558	50-g GCT	7.8	All	No regard	59	91		By ROC analysis
			7.5	All	No regard	61	88		By ROC analysis
			7.0	All	No regard	68	82		By ROC analysis
		FPG	4.8	All	Fasting	81	76		By ROC analysis
			4.4	All	Fasting	100	39		By ROC analysis
Cetin and Cetin, 1997[153]	274	50-g GCT	7.8	All	Those eaten within 2 h	75	86	27	
			8.2	All	Those eaten within 2 h	63	91	33	
			7.8	All	Those eaten between 2 and 3 h	60	89	30	
			7.9	All	Those eaten between 2 and 3 h	60	92	30	
			7.8	All	Those eaten > 3 h	50	89	25	
			8.3	All	Those eaten > 3 h	50	92	33	
			7.8	All	No regard	65	88	27	
Schwartz *et al.*, 1994[97]	132	50-g GCT modified glucose	7.22	All	Fasting	100	36.4		
			7.78	All	Fasting	96	52.3		
			8.33	All	Fasting	96	61.5		
			8.89	All	Fasting	96	78.5		
			9.44	All	Fasting	84	86.9		
		50-g GCT	7.22	All	Fasting	92	43		
			7.78	All	Fasting	92	52.3		
			8.33	All	Fasting	84	66.4		
			8.89	All	Fasting	80	78.5		
			9.44	All	Fasting	76	85		
Murphy *et al.*, 1994[188]	44	50-g GCT polymer	7.5	All	No regard	100	92.8	49	
	41	50-g GCT standard	7.5	All	No regard	33.3	73.6	9	
	85	50-g GCT	7.5	All	No regard	60	84	16	

continued

Study	*n*	Test (GCT, FPG, etc.)	Threshold ≥ mmol/l	All patients, by age group or RF?	Fasting?	Sensitivity (%)	Specificity (%)	PPV (%)	Comments
Mathai *et al.*, 1994[184]	232	50-g GCT	5.3	All	No regard	82	20		Approximately half sample had RFs for GDM
			5.8	All	No regard	73	39		
			6.4	All	No regard	63	55		
			6.9	All	No regard	55	55		
			7.2	All	No regard	45	74		
			7.5	All	No regard	45	81		
			7.8	All	No regard	36	82		
			8.3	All	No regard	36	89		
		RPG	4.4	All	No regard	100	38		
			4.7	All	No regard	82	52		
			5.0	All	No regard	64	66		
			5.3	All	No regard	64	74		
			5.5	All	No regard	64	81		
			5.8	All	No regard	27	86		
			6.0	All	No regard	18	90		
			6.4	All	No regard	18	93		
Jirapinyo *et al.*, 1993[164]	396	50-g GCT	7.77	RF women	No regard	86	65	23	
			8.3	RF women	No regard	83	79	32	
Benjamin *et al.*, 1986[158]	101	50-g GCT	7.77	At second trimester	No regard	25	74		
			8.3	At second trimester	No regard	25	83		
			7.77	At third trimester	No regard	88	73		
			8.3	At third trimester	No regard	82	88		
Court *et al.*, 1985[123]	100	100-g GCT standard	8	All	No regard	Not reported			
		100-g GCT polymer	8	All	No regard	88	80	25	
O'Sullivan *et al.*, 1973[100]	752	50-g GCT (converted from whole blood)	8.2	All	No regard	79	87	15	
				Clinical history	No regard	63	56		
				Both	No regard	53	93		
			8.2	Aged ≥ 25	No regard	88	82	19	
				Aged ≥ 25 + clinical history	No regard	69	35		
				Both	No regard	63	80	19	
RPG									
Nasrat *et al.*, 1988[61]	276	RPG	7.0 or 6.4	All	Eaten < 2 h	16	96	47	
					Eaten > 2 h				
Jowett *et al.*, 1987[140]	110	RPG	5.6	All	8 am	66	94		Authors used inappropriate sensitivity and specificity calculation – these are our figures
					12 pm	50	86		
					3 pm	58	67		
					5 pm	58	84		
					10 pm	50	72		
				6.0	8 am	58	96		
					12 pm	41	94		
					3 pm	50	74		
					5 pm	41	86		
					10 pm	41	87		

continued

Study	*n*	Test (GCT, FPG, etc.)	Threshold ≥ mmol/l	All patients, by age group or RF?	Fasting?	Sensitivity (%)	Specificity (%)	PPV (%)	Comments
Fasting									
Agarwal *et al.*, 2000[142]	430	FPG (converted FBG)	6.0	All	Fasting	48	97.5		
			4.9	All	Fasting	93	38.5		
Reichelt *et al.*, 1998[143]	5010	FPG	4.5	All	Fasting	94	51	0.6	
			4.6	All	Fasting	94	58	0.7	
			4.7	All	Fasting	94	66	0.9	
			4.8	All	Fasting	88	72	1	
			4.9	All	Fasting	88	78	1.3	
Fuhrmann, 1989[185]	2510	FPG (converted FBG)	5.0	All	Fasting	100	74	4	
			5.3	All	Fasting	93	83	6	
			5.7	All	Fasting	86	89	7.5	
			6.3	All	Fasting	75	95	15	
Fructosamine									
Uncu *et al.*, 1995[182]	42	Fructosamine	2.85	All	Not stated	71	44	40	
		50-g GCT	7.8	All	No regard	75	53	45	
Begley *et al.*, 1992[217]	36	Fructosamine	2.4	All	Fasting	0	100		

Note: Studies reported are those that gave all women a diagnostic GTT regardless of screening result, except those who found no incidence of GDM

Health Technology Assessment Programme

Prioritisation Strategy Group

Members

Chair,
Professor Kent Woods,
Director,
NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Professor Bruce Campbell,
Consultant Vascular
& General Surgeon,
Royal Devon & Exeter Hospital

Professor Shah Ebrahim,
Professor in Epidemiology
of Ageing,
University of Bristol

Dr John Reynolds,
Clinical Director,
Acute General Medicine SDU,
Oxford Radcliffe Hospital

Dr Ron Zimmern,
Director, Public Health
Genetics Unit,
Strangeways Research
Laboratories, Cambridge

HTA Commissioning Board

Members

Programme Director,
Professor Kent Woods,
Director,
NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Chair,
Professor Shah Ebrahim,
Professor in Epidemiology
of Ageing,
University of Bristol

Deputy Chair,
Professor Jon Nicholl,
Director, Medical Care
Research Unit,
University of Sheffield

Professor Douglas Altman,
Director, ICRF Medical
Statistics Group,
University of Oxford

Professor John Bond,
Director, Centre for Health
Services Research, University
of Newcastle-upon-Tyne

Professor John Brazier,
Director of Health Economics,
University of Sheffield

Dr Andrew Briggs,
Research Fellow,
Institute of Health Sciences,
University of Oxford

Ms Christine Clark,
Freelance Medical Writer,
Bury, Lancs

Professor Martin Eccles,
Professor of
Clinical Effectiveness,
University of Newcastle-
upon-Tyne

Dr Andrew Farmer,
General Practitioner &
NHS R&D Clinical Scientist,
Institute of Health Sciences,
University of Oxford

Professor Adrian Grant,
Director, Health Services
Research Unit,
University of Aberdeen

Dr Alastair Gray,
Director, Health Economics
Research Centre,
Institute of Health Sciences,
University of Oxford

Professor Mark Haggard,
Director, MRC Institute
of Hearing Research,
University of Nottingham

Professor Jenny Hewison,
Academic Unit of Psychiatry
& Behavioural Sciences,
University of Leeds

Professor Peter Jones,
University Department
of Psychiatry,
University of Cambridge

Professor Alison Kitson,
Director, Royal College of
Nursing Institute, London

Professor Sarah Lamb,
Research Professor
in Physiotherapy,
University of Coventry

Dr Donna Lamping,
Head, Health Services
Research Unit,
London School of Hygiene
& Tropical Medicine

Professor David Neal,
Department of Surgery,
University of Newcastle-
upon-Tyne

Professor Tim Peters,
Social Medicine,
University of Bristol

Professor Martin Severs,
Professor in Elderly
Health Care,
University of Portsmouth

Dr Jonathan Shapiro,
Senior Fellow, Health Services
Management Centre,
Birmingham

Dr Sarah Stewart-Brown,
Director, Health Services
Research Unit,
University of Oxford

Dr Gillian Vivian, Consultant
in Nuclear Medicine &
Radiology, Royal Cornwall
Hospitals Trust, Truro

continued

continued

Diagnostic Technologies & Screening Panel

Members

Chair,
Dr Ron Zimmern,
Director, Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

Mrs Stella Burnside,
Chief Executive,
Altnagelvin Hospitals Health & Social Services Trust,
Londonderry

Dr Paul O Collinson,
Consultant Chemical Pathologist &
Senior Lecturer,
St George's Hospital,
London

Dr Barry Cookson,
Director, Laboratory of Hospital Infection, Public Health Laboratory Service,
London

Professor Howard Cuckle,
Professor of Reproductive Epidemiology,
University of Leeds

Professor Adrian K Dixon,
Professor of Radiology,
Addenbrooke's Hospital,
Cambridge

Dr David Elliman,
Consultant in Community Child Health,
St. George's Hospital, London

Dr Tom Fahey,
Senior Lecturer in General Practice,
University of Bristol

Dr Andrew Farmer,
General Practitioner & NHS R&D Clinical Scientist,
Institute of Health Sciences,
University of Oxford

Professor Jane Franklyn,
Professor of Medicine,
University of Birmingham

Dr Antony J Franks,
Deputy Medical Director,
The Leeds Teaching Hospitals NHS Trust

Dr J A Muir Gray,
Programmes Director,
National Screening Committee,
NHS Executive, Oxford

Dr Peter Howlett,
Executive Director – Planning,
Portsmouth Hospitals NHS Trust

Dr S M Ludgate,
Medical Director,
Medical Devices Agency,
London

Professor Jennie Popay,
Professor of Sociology & Public Health,
Institute for Health Research,
University of Lancaster

Dr Susan Schonfield,
CPHM Specialist Commissioning,
Public Health Directorate,
Croydon Primary Care Trust

Mrs Kathlyn Slack,
Professional Support,
Diagnostic Imaging & Radiation Protection Team,
Department of Health,
London

Mr Tony Tester,
Chief Officer,
South Bedfordshire Community Health Council,
Luton

Dr Andrew Walker,
Senior Lecturer in Health Economics,
University of Glasgow

Professor Martin J Whittle,
Head of Division of Reproductive & Child Health,
University of Birmingham

Pharmaceuticals Panel

Members

Chair,
Dr John Reynolds,
Clinical Director,
Acute General Medicine SDU,
Oxford Radcliffe Hospital

Professor Tony Avery,
Professor of Primary Health Care,
University of Nottingham

Professor Iain T Cameron,
Professor of Obstetrics & Gynaecology,
University of Southampton

Mr Peter Cardy,
Chief Executive,
Macmillan Cancer Relief,
London

Dr Christopher Cates,
GP & Cochrane Editor,
Bushey Health Centre,
Bushey, Herts

Dr Karen A Fitzgerald,
Pharmaceutical Adviser,
Bro Taf Health Authority,
Cardiff

Dr Felicity J Gabbay,
Managing Director,
Transcrip Ltd,
Milford-on-Sea, Hants

Mr Peter Golightly,
Director, Trent Medicines Information Services,
Leicester Royal Infirmary

Dr Alastair Gray,
Director, Health Economics Research Centre,
Institute of Health Sciences,
University of Oxford

Mrs Sharon Hart,
Managing Editor, *Drug & Therapeutics Bulletin*,
London

Dr Christine Hine,
Consultant in Public Health Medicine,
Bristol South & West Primary Care Trust

Mrs Jeannette Howe,
Deputy Chief Pharmacist,
Department of Health,
London

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Dr Frances Rotblat,
CPMP Delegate,
Medicines Control Agency,
London

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, St James's University Hospital, Leeds

Mrs Katrina Simister,
New Products Manager,
National Prescribing Centre,
Liverpool

Professor Terence Stephenson,
Professor of Child Health,
University of Nottingham

Dr Richard Tiner,
Medical Director,
Association of the British Pharmaceutical Industry,
London

Professor Jenifer Wilson-Barnett, Head of Florence Nightingale School of Nursing & Midwifery, King's College,
London

Therapeutic Procedures Panel

Members

Chair,
Professor Bruce Campbell,
Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital

Professor John Bond,
Professor of Health Services Research, Centre for Health Services Research, University of Newcastle-upon-Tyne

Ms Judith Brodie,
Head of Cancer Support Service, Cancer BACUP, London

Ms Tracy Bury,
Head of Research & Development, Chartered Society of Physiotherapy, London

Mr Michael Clancy,
Consultant in A & E Medicine, Southampton General Hospital

Professor Collette Clifford,
Professor of Nursing & Head of Research, School of Health Sciences, University of Birmingham

Dr Carl E Counsell,
Senior Lecturer in Neurology, University of Aberdeen

Dr Keith Dodd,
Consultant Paediatrician, Derbyshire Children's Hospital, Derby

Mr John Dunning,
Consultant Cardiothoracic Surgeon, Papworth Hospital NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon, Gloucestershire Royal Hospital, Gloucester

Professor Gene Feder,
Professor of Primary Care R&D, St Bartholomew's & the London, Queen Mary's School of Medicine & Dentistry, University of London

Professor Richard Johanson,
Consultant & Senior Lecturer, North Staffordshire Infirmary NHS Trust, Stoke-on-Trent (deceased Feb 2002)

Dr Duncan Keeley,
General Practitioner, Thame, Oxon

Dr Phillip Leech,
Principal Medical Officer for Primary Care, Department of Health, London

Mr George Levvy,
Chief Executive, Motor Neurone Disease Association, Northampton

Professor James Lindesay,
Professor of Psychiatry for the Elderly, University of Leicester

Professor Rajan Madhok,
Medical Director & Director of Public Health, North & East Yorkshire & Northern Lincolnshire Strategic Health Authority, York

Dr Mike McGovern,
Senior Medical Officer, Heart Team, Department of Health, London

Dr John C Pounsford,
Consultant Physician, Frenchay Healthcare Trust, Bristol

Professor Mark Sculpher,
Professor of Health Economics, Institute for Research in the Social Services, University of York

Dr Ken Stein,
Senior Lecturer in Public Health, Peninsular Technology Assessment Group, University of Exeter

continued

Expert Advisory Network

Members

Mr Gordon Aylward,
Chief Executive,
Association of British
Health-Care Industries,
London

Mr Shaun Brogan,
Chief Executive,
Ridgeway Primary Care Group,
Aylesbury, Bucks

Mr John A Cairns,
Reader in Health Economics,
Health Economics
Research Unit,
University of Aberdeen

Professor Nicky Cullum,
Director of Centre for
Evidence-Based Nursing,
University of York

Dr Katherine Darton,
Information Unit,
MIND – The Mental
Health Charity, London

Professor Carol Dezateux,
Professor of
Paediatric Epidemiology,
Institute of Child Health,
London

Professor Pam Enderby,
Dean of Faculty of Medicine
Institute of General Practice
& Primary Care,
University of Sheffield

Mr Leonard R Fenwick,
Chief Executive,
Freeman Hospital,
Newcastle-upon-Tyne

Professor David Field,
Professor of
Neonatal Medicine,
The Leicester Royal
Infirmary NHS Trust

Mrs Gillian Fletcher,
Antenatal Teacher &
Tutor & President,
National Childbirth
Trust, Henfield,
West Sussex

Ms Grace Gibbs,
Deputy Chief Executive
Director for Nursing,
Midwifery & Clinical
Support Services,
West Middlesex
University Hospital,
Isleworth, Middlesex

Dr Neville Goodman,
Consultant Anaesthetist,
Southmead Hospital, Bristol

Professor Robert E Hawkins,
CRC Professor & Director
of Medical Oncology,
Christie Hospital NHS Trust,
Manchester

Professor F D Richard Hobbs,
Professor of Primary Care
& General Practice,
University of Birmingham

Professor Allen Hutchinson,
Director of Public Health &
Deputy Dean of ScHARR,
University of Sheffield

Professor David Mant,
Professor of General Practice,
Institute of Health Sciences,
University of Oxford

Professor Alexander Markham,
Director,
Molecular Medicine Unit,
St James's University Hospital,
Leeds

Dr Chris McCall,
General Practitioner,
The Hadleigh Practice,
Corfe Mullen, Dorset

Professor Alistair McGuire,
Professor of Health Economics,
London School of Economics,
University of London

Dr Peter Moore,
Freelance Science Writer,
Ashtead, Surrey

Dr Andrew Mortimore,
Consultant in Public
Health Medicine,
Southampton City Primary
Care Trust

Dr Sue Moss,
Associate Director,
Cancer Screening
Evaluation Unit,
Institute of Cancer Research,
Sutton, Surrey

Mrs Julietta Patnick,
National Coordinator,
NHS Cancer
Screening Programmes,
Sheffield

Professor Chris Price,
Director of Clinical Research,
Bayer Diagnostics Europe,
Stoke Poges, Berks

Ms Marianne Rigge,
Director, College of Health,
London

Dr William Rosenberg,
Senior Lecturer &
Consultant in Medicine,
University of Southampton

Professor Ala Szczepura,
Director, Centre for
Health Services Studies,
University of Warwick

Dr Ross Taylor,
Senior Lecturer,
Department of General
Practice & Primary Care,
University of Aberdeen

Mrs Joan Webster,
Consumer member,
HTA – Expert
Advisory Network

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.ncchta.org) 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.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278