The Australian Centre for Health Services Innovation

Cost-Effectiveness Analysis and Implications for Health Librarians

9th July 2015

Dr Son Nghiem Senior Research Fellow

Victoria McCreanor Research Associate

Greg Merlo Research Assistant
This booklet is **not** to be reproduced without approval.

THE AUSTRALIAN CENTRE FOR HEALTH SERVICES INNOVATION

Institute for Health & Biomedical Innovation
Queensland University of Technology
60 Musk Avenue, KELVIN GROVE Q 4059

[www.aushsi.org.au](http://www.aushsi.org.au)
[contact@aushsi.org.au](mailto:contact@aushsi.org.au)
Cost-effectiveness Analysis in Heath Care - Overview

Victoria McCreanor

Overview

• Cost-effectiveness modelling studies vs RCTs
• Overview of cost-effectiveness analysis
  – Decision trees
  – Markov models
• Using model outcomes for decision-making
  – Cost-effectiveness plane
  – Probabilistic sensitivity analysis
• Brainstorming activity

Cost-effectiveness Studies vs RCT
Clinical Study Designs – Evidence Pyramid

**RCT as the gold standard for economic evaluation?**

Very good for estimating effectiveness and causality
- Confounders are randomly assigned

**Some Limitations**
1. Often only compare one alternative with placebo or existing practice
2. Short time frame and long term outcomes
3. High internal validity but poor generalizability
4. Might be better sources for some of the information

**RCT as the gold standard for economic evaluation?**

**NO**

- Useful for informing the effectiveness of any new intervention
- One piece of the jigsaw
- Not particularly good for making a decision
Modelling Studies

Enable comparison of all relevant competing interventions
Include long-term cost and quality and length of life outcomes
Findings are generalizable compared to those achieved in a controlled experiment

RCT has developed a central role in applied cost-effectiveness studies

Goal is to decide whether to adopt a new technology/service

A framework is needed:
- Estimates all costs and effects
- Includes all relevant interventions
- Includes appropriate time horizon
- Covers specific patient groups
- Uses the best available evidence in an explicit and transparent way

“These requirements suggest that, in most circumstances, the use of a single RCT as a vehicle for economic analysis will be an inadequate and a partial basis for decision making”

“RCT evidence should be viewed as simply one of the sources of evidence, which must be placed in a broader framework of evidence synthesis and decision analysis.”

Overview of Cost-effectiveness Analysis
To understand the value for money of a new service or intervention we estimate how:

Costs will change
Benefits (QALYs) will change

\[ \Delta \]  
Delta = “change in”

An Incremental Cost-Effectiveness Ratio

\[ \text{ICER} = \frac{\text{Cost}_{\text{After}} - \text{Cost}_{\text{Before}}}{\text{QALY}_{\text{After}} - \text{QALY}_{\text{Before}}} = \frac{\Delta C}{\Delta E} \]

- Measures changes to both COSTS (C) and EFFECTS (E)

QALY = Quality-Adjusted Life-Year
- Generic measure of effectiveness of a treatment
- Allows comparison across different diseases areas
- Takes into account effects of treatment on both quality of life and length of life

Decision Trees

What is the decision  Chance events  Endpoints
Data

Chance events

The probability of the event happening for each decision
The sum of the probabilities for each chance node must equal 1

Endpoints

The values for the endpoints for each decision.
Four sets of costs and health benefits

How they work

Evaluate this way

New & lives + $150 + 10 years
New & dies + $130 + 0 years
Old & lives + $50 + 10 years
Old & dies + $30 + 0 years

Data
New surgical technique

How they work

When the patient lives:
the expected value of the costs is $150 \times 0.9 = $135
the expected value of the health benefits is $10 \times 0.9 = 9$ years

When the patient dies:
the expected value of the costs is $130 \times 0.1 = $13
the expected value of the health benefits is $0 \times 0.1 = 0$ years

$148 and 9 years
How they work

When the patient lives:
The expected value of the costs is $50 \times 0.8 = $40
The expected value of the health benefits is 10 \times 0.8 = 8\, \text{years}

When the patient dies:
The expected value of the costs is $30 \times 0.2 = $6
The expected value of the health benefits is 0 \times 0.2 = 0\, \text{years}

$46\, \text{and 8\, years}$
Add all branches and compare

<table>
<thead>
<tr>
<th>New surgical technique</th>
<th>Old technique</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected costs</td>
<td>$50</td>
<td>$40</td>
</tr>
<tr>
<td>Expected health benefits</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>

These guys need to learn about state based models

The Essentials of state based models

Often called ‘Markov Models’

The states are independent and exhaustive

Movements between them are based on transition probabilities

The model updates at regular cycles defined in weeks/months/year

Costs and health benefits are attributed to someone who spends one cycle in any state
Markov Models – States

Healthy

Sick

Dead

Dead state absorbs people
No one can leave the model

Transition to:

<table>
<thead>
<tr>
<th>Transition from</th>
<th>Healthy</th>
<th>Sick</th>
<th>Dead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>0.6</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Sick</td>
<td>0</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>Dead</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Transition to:

Cycle 1

Healthy

Sick

Dead

Cycle 2

Cycles – simulating a cohort of 1000 patients

Cycle 1

Healthy

Sick

Dead

Cycle 2

Cycles – simulating a cohort of 1000 patients

Cycle 1

Healthy

Sick

Dead

Cycle 2

Cycles – simulating a cohort of 1000 patients

Cycle 1

Healthy

Sick

Dead

Cycle 2

Cycles – simulating a cohort of 1000 patients

Cycle 1

Healthy

Sick

Dead

Cycle 2

Cycles – simulating a cohort of 1000 patients

Cycle 1

Healthy

Sick

Dead

Cycle 2

Cycles – simulating a cohort of 1000 patients

Cycle 1

Healthy

Sick

Dead

Cycle 2
To make these models useful for decision making ...
We need to be able to describe these outcomes in terms of costs and QALYs

One cycle is 12 months in our model

We have these data to add to our model

<table>
<thead>
<tr>
<th>For any cycle</th>
<th>Costs</th>
<th>Utility Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>$150</td>
<td>0.8</td>
</tr>
<tr>
<td>Sick</td>
<td>$3,150</td>
<td>0.7</td>
</tr>
<tr>
<td>Dead</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

To make these models useful for decision making ...
We need to be able to describe these outcomes in terms of costs and QALYs

One cycle is 12 months in our model

We have these data to add to our model

To understand the value for money of a new service or intervention we estimate how:

Costs will change
Benefits (QALYs) will change

\[ \text{ICER} = \frac{\Delta C}{\Delta E} \]

\[ \text{ICER} = \frac{\text{Cost}_{\text{After}} - \text{Cost}_{\text{Before}}}{\text{QALY}_{\text{After}} - \text{QALY}_{\text{Before}}} \]
23/06/2015

**NEW SERVICE ‘X’**  
Baseline comparator | After the change | The Change
---|---|---
Costs | $6,721,207 | $15,414,036 | $8,692,829
Health Benefits in QALYs | 4,043 | 4,252 | 210

A change to costs of $8,692,829 gives 210 QALYS

Our **ICER**:

\[
\frac{\Delta C}{\Delta E} = \frac{8,692,829}{210} = \$41,448
\]

The cost per QALY gained is the change to costs divided by the change to health benefits, and is $41,448.

The cost-effectiveness plane

Lower costs  
Less Health  
Higher costs  
$8,692,829

210 years of life
Using Model Outcomes for Decision-Making

A decision-making tool

- Higher Costs
- Fewer QALYs
- Lower Costs
- More QALYs

Most lie here

- X
- ✓
Are we willing to pay $41,448 per QALY?

Before

NEW SERVICE ‘X’

Higher Costs
$8,692,829

Willingness to pay threshold (WTP)

After

In Australia, we are generally willing to pay $40,000 - $60,000 per QALY gained

Fewer QALYs

More QALYs

Lower Costs

Probabilistic Sensitivity Analysis (PSA)

- Very useful for decision-making if there is uncertainty around different parameters in the model
- Run the model 1,000s of times using randomly selected values from uncertainty ranges or distributions – Monte Carlo Simulation (named after the casino)
- Lots of different point estimates
- Count how many fall under the WTP threshold of the decision-maker
- Gives probability of cost-effectiveness, depending on WTP threshold

PSA Results

WTP Threshold
Brainstorming – search terms

What search terms would you use to find papers which use this kind of analysis or modelling?

https://bubbl.us/mindmap
Some Terminology

- CEA = Cost-effectiveness Analysis
- ICER = Incremental Cost-effectiveness Ratio
- Markov Model – Health state-based model
- Monte Carlo Simulation – random draws for sensitivity analysis
- QALY = Quality-Adjusted Life-Year
- PSA = Probabilistic Sensitivity Analysis
- WTP = Willingness to Pay

Next up:
How do we get the values for all the parameters in our model?
How to measure costs

Gregory Merlo

Overview

• How to measure costs?
• How to measure health benefits?

The Challenge

Costing involves identifying, measuring and valuing all resource changes that occur if a health intervention is carried out.

The aim is to value the use of all scarce health care resources needed to produce a certain health effect.
Humans, Space and Things

Combining all three is normally required to supply health services

Medical resources, directly needed for the intervention

Non-medical resources needed for the intervention

Time of informal carers and other costs of informal care

Patient time including productivity changes

Time off work or normal activities

Humans, Space and Things

Time off work, or normal activities

Not being able to do something productive implies a loss or cost

Salaried or non-salaried equally valid
Resource changes occur inside and outside of the health care system and both now and in the future.

Medical resources, directly needed for the intervention

Non-medical resources needed for the intervention

Practical issues to guide costing method

A. Can they be measured with accuracy?
   - Humans
     - How much time is given up
     - Prospective time-and-motion studies
     - Individual activity diaries
     - Surveys or interviews
   
   - Space
     - Use plans or a tape measure
   
   - Things
     - Count them

B. Can they be valued?
   - Humans
     - Find out salaries
     - Include pension & leave costs
   
   - Space
     - What does the market say space is worth
   
   - Things
     - Durable things
       - Upfront investment costs are allocated
       - Maintenance costs
     - Disposable things
       - Price paid
C. Are the costs large?

Humans
People are expensive, depends on their time

Space
How much space is required and is it valuable

Things
Depends on the project (robot surgery vs. hand hygiene)

D. Will they be considered by the decision maker?

<table>
<thead>
<tr>
<th></th>
<th>Ward Manager</th>
<th>Hospital CEO</th>
<th>Minister</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical resources, directly needed for the intervention</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Non medical resource needed for the intervention</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time of informal care and other costs of informal care</td>
<td>x</td>
<td>x</td>
<td>?</td>
</tr>
<tr>
<td>Patient time including productivity changes</td>
<td>x</td>
<td>x</td>
<td>?</td>
</tr>
</tbody>
</table>

How to measure health benefits
For Debate...

Economics of coronary artery bypass grafting

**ALAN WILLIAMS**

Medical Management or CABG

"procedures should be ranked so that activities that generate more gains to health for every £ of resources take priority over those that generate less; thus the general standard of health in the community would be correspondingly higher"

---

**QALYs**

Estimate changes to the quantity of life

Expressed as survival; few problems of comparison

**ALIVE**

**DEAD**
QALYs

Estimate changes to the quality of life

Index of health utility between zero and one

The advantage of QALYs

All changes to health can be valued on an any health programme compared

Where do health utility weights come from?

<table>
<thead>
<tr>
<th>Describe the health state</th>
<th>Value it</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully mobile, no physical pain, able to engage socially</td>
<td>that's worth?</td>
</tr>
<tr>
<td>Social function restrictions, no pain, some psychological distress</td>
<td>that's worth?</td>
</tr>
<tr>
<td>Low energy, some pain, minor psychological distress</td>
<td>that's worth?</td>
</tr>
<tr>
<td>Unconscious, unable to move or function</td>
<td>that's worth?</td>
</tr>
</tbody>
</table>

Three popular approaches (that differ)

- Visual Analogue Scale: easy to use
- Time Trade off experiments: force choices
- Standard Gamble experiments
Visual Analogue Scale

How would you rate your health (or the health of a loved one)?

Time Trade Off

Which option would you prefer?

Alternative 2
- 10 years of perfect health followed by death
- 9 years of perfect health followed by death
- 8 years of perfect health followed by death
- 7 years of perfect health followed by death
- 6 years of perfect health followed by death
- 5 years of perfect health followed by death
- 4 years of perfect health followed by death
- 3 years of perfect health followed by death
- 2 years of perfect health followed by death
- 1 year of perfect health followed by death

Alternative 1
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death
- 10 years with no sight followed by death

How many years in perfect health will you trade for?

Divide years by 10 to get a utility score.

Standard Gamble

HEALTHY (10 years)

DEAD (10 years)
Who should we get this information from?

These exercises can be done on real patients with the health conditions

Advantage is it avoids need to describe the health state
Disadvantage is it adds distress to patients

An alternative is to use a Multi-Attribute Utility Scale

Standardised health state classifications with pre-existing set of utility weights
EQ-5D is popular (AQoL is an Australian one)

EQ-5D

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Levels</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>3</td>
<td>1 = no problem, 2 = some problems, 3 = extreme problems</td>
</tr>
<tr>
<td>Self-care</td>
<td>3</td>
<td>1 = no problem, 2 = some problems, 3 = extreme problems</td>
</tr>
<tr>
<td>Usual Activities</td>
<td>3</td>
<td>1 = can do all usual activities, 2 = some problems, 3 = extreme problems</td>
</tr>
<tr>
<td>Pain/Dyspareunia</td>
<td>3</td>
<td>1 = no pain or dyspareunia, 2 = moderate pain/dyspareunia, 3 = extreme pain/dyspareunia</td>
</tr>
<tr>
<td>Anxiety/Depression</td>
<td>3</td>
<td>1 = no anxiety or depression, 2 = moderate anxiety or depression, 3 = extreme anxiety or depression</td>
</tr>
</tbody>
</table>

11111 = best possible health state
Health Utility Index close to 1
Three levels for each dimension of health

1 = no problem
2 = some problems
3 = extreme problems

33333 = best possible health state

Health Utility Index close to zero

243 unique states exist

Josephine Patient occupies one of the 243 unique health states in the EQ-5D. They have all been valued (TTO or SG) by a representative sample of the population.

Her Health utility score is 0.57

EMBASE search

Research Question
– Is screening for type 2 diabetes in adults cost-effective?

Search terms (and synonyms)
– Type 2 diabetes
– Screening
– Cost effectiveness (or economic evaluation)
  • https://sites.google.com/a/york.ac.uk/issg-search-filters/resource/filters-to-find-i
Economic evaluation search filters

York university
- https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-find-

Options
- NHS EED – recommended
  - http://www.crd.york.ac.uk/crdweb/searchstrategies.asp
- SIGN – Pragmatic but untested
  - http://www.sign.ac.uk/methodology/filters.html#econ
- HIRU – Tested but not pragmatic!
What is health economics?

A branch of economics concerned with issues related to efficiency, effectiveness, value and behavior in the production and consumption of health and health care.

Adapted from G. Basu (2013) Bayesian Methods in Health Economics
What do health economists do?

- Analyze the efficiency and equity in health care financing
- Identify determinants of health care demand
- Estimate and predict trends in health care costs
- Determine the labor market in health care
- Analyzing the budget system and inventory management

What is a Systematic Review?

A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review.


Systematic review process

1. Question
   - Scoping search
   - Protocol
   - Full search
   - Title and abstract screening
   - Full-text retrieval
   - Reference management
   - Included references agreed
   - Additional searching
   - Writing up
   - Update search
   - Synthesis
   - Data extraction
## Systematic vs Traditional Review

<table>
<thead>
<tr>
<th>Feature</th>
<th>Systematic Review</th>
<th>Narrative / Traditional Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
<td>Often focused (clinical) question.</td>
<td>Sometimes based on scope.</td>
</tr>
<tr>
<td>Sources &amp; Search</td>
<td>Explicit search strategy of multiple</td>
<td>Not usually specified.</td>
</tr>
<tr>
<td></td>
<td>databases. Comprehensive sources.</td>
<td></td>
</tr>
<tr>
<td>Selection</td>
<td>Criterion-based selection, uniformly</td>
<td>Not usually specified.</td>
</tr>
<tr>
<td></td>
<td>applied.</td>
<td></td>
</tr>
<tr>
<td>Appraisal</td>
<td>Rigorous critical appraisal.</td>
<td>Varies</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Quantitative summary. Also qualitative</td>
<td>Often qualitative summary.</td>
</tr>
<tr>
<td></td>
<td>narrative.</td>
<td></td>
</tr>
<tr>
<td>Inferences</td>
<td>Based on all available evidence.</td>
<td>Based on a sample of the evidence.</td>
</tr>
<tr>
<td>Grading</td>
<td>All evidence is graded (quality)</td>
<td>May or may not be graded.</td>
</tr>
</tbody>
</table>


## Mythbusters

- Systematic reviews:
  - are the same as ordinary literature reviews, only bigger
  - include only randomized controlled trials
  - require the adoption of a biomedical model of health
  - are of no relevance outside of health / medicine
  - must involve statistical analysis / synthesis
  - must be conducted by experts
  - can be done without experienced librarian support
  - are not really research


## Review Team

- **Clinical expert**
  - Initiates, defines, selects topic.
- **Clinical expert**
  - Partners in above process, and collaborates in review to prevent bias.
- **Statistician**
  - Provides methodological oversight, ensures process quality for entire project.
- **Librarian**
  - Provides methodological oversight, ensures process quality for information search process.
- **Healthcare Consumer**
  - Provides insight into the priorities for research, information conduit for relating priorities and findings between consumers and clinicians.
Research Questions

- Questions often use the PICO(S) framework:
  - Population/Participants
  - Interventions
  - Comparisons
  - Outcomes
  - Study design

- “To assess the effects of [intervention] compared to [comparison/control] for [condition/problem] in [population] in [context] on [outcomes].”

Example Research Question

- Are antiseptic washes more effective than non-antiseptic washes at preventing nosocomial infections in patients undergoing surgery?

Mind map research question
Construct search phrases

<table>
<thead>
<tr>
<th>1st theme</th>
<th>OR</th>
<th>OR</th>
<th>OR</th>
<th>AND</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd theme</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>AND</td>
</tr>
<tr>
<td>3rd theme</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>AND</td>
</tr>
<tr>
<td>4th theme</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>AND</td>
</tr>
</tbody>
</table>

Example search phrase

(cost benefit* or cost effectiveness or economic*) AND (disability insurance[Title/Abstract] OR national health[Title/Abstract] OR health insurance[Title/Abstract]) NOT (drug* or obesity or diabetes or cancer or heart or wound* or disease or illness)

Identifying sources

- Databases?
  - Cochrane, NHS EED, HTA, DARE, Bibliomap
- Websites?
  - Google? Search by Organisation
- Hand searching?
  - Which journals are key to your work?
- Snowballing?
  - WoK or Scopus
- Books?
  - LibrarySearch or COPAC or GoogleBooks
- Grey literature?
  - Google: blog search; SIGLE: conference proceedings
- Ongoing trials?
  - Trials registers
### Reference management

<table>
<thead>
<tr>
<th>Software</th>
<th>Database connectivity</th>
<th>Import format</th>
<th>Word processor integration</th>
<th>PDF Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>BibDesk</td>
<td>Good</td>
<td>Excellent</td>
<td>Lyx</td>
<td>Good</td>
</tr>
<tr>
<td>EndNote</td>
<td>Excellent</td>
<td>Good</td>
<td>MS Word, OpenOffice, Pages</td>
<td>N/A</td>
</tr>
<tr>
<td>JabRef</td>
<td>Good</td>
<td>Excellent</td>
<td>OpenOffice, MS Word, Lyx</td>
<td>N/A</td>
</tr>
<tr>
<td>KBibTeX</td>
<td>Excellent</td>
<td>Fair</td>
<td>Lyx</td>
<td>N/A</td>
</tr>
<tr>
<td>Mendeley</td>
<td>Excellent</td>
<td>Fair</td>
<td>MS Word, OpenOffice</td>
<td>Excellent</td>
</tr>
<tr>
<td>Papers</td>
<td>Excellent</td>
<td>Poor</td>
<td>MS Word, OpenOffice, Pages</td>
<td>N/A</td>
</tr>
<tr>
<td>qIndex</td>
<td>Fair</td>
<td>Fair</td>
<td>MS Word, Lyx</td>
<td>N/A</td>
</tr>
<tr>
<td>Docear</td>
<td>Poor</td>
<td>Good</td>
<td>MS Word, Pages, TextEdit, Lyx</td>
<td>Excellent</td>
</tr>
<tr>
<td>Zotero</td>
<td>Good</td>
<td>Good</td>
<td>MS Word, OpenOffice, GoogleDocs</td>
<td>Good</td>
</tr>
</tbody>
</table>

### Meta-analysis

- The analysis of other analyses
- Uses data from randomized controlled trials
- Aggregates and combines the results of comparable studies into a coherent account to discover main effects
- Often uses statistical processes
- Looks at effect size, not only statistical significance
- Combines the results of small-scale studies
- Uses transparent means to draw conclusions

### Synthesis: complete pooling

G. Baio (2013). Bayesian Methods in Health Economics
**Synthesis: No pooling**

![Diagram of Synthesis: No pooling](image)

**Synthesis: mixed**

![Diagram of Synthesis: mixed](image)

**Bias in meta-analysis**

- **Publication bias:**
  - Studies never published
  - Studies with no beneficial effect
  - Studies sponsored by pharmaceutical industry
  - Studies from a single center versus multiple centers

- **Language bias:**
  - Positive findings published in a foreign journal
  - Negative findings published in a local journal

- **Database bias:**
  - Journals not indexed in major databases

*"Doing a meta-analysis is easy, doing one well is hard"*

-Ingram Olkin
Example paper

“Different strategies for screening and prevention of type 2 diabetes in adults: cost effectiveness analysis”

Questions

What type of economic evaluation does the paper describe?
What interventions are compared in the evaluation?
Where can you find the sources for the parameters of the economic evaluation?
Do you think screening for type II diabetes is cost effective? Why?
Different strategies for screening and prevention of type 2 diabetes in adults: cost effectiveness analysis

Clare L Gillies, lecturer in medical statistics, 1 Paul C Lambert, senior lecturer in medical statistics,1 Keith R Abrams, professor of medical statistics,1 Alex J Sutton, reader in medical statistics,1 Nicola J Cooper, MRC training fellow in health services research,1 Ron T Hsu, senior clinical teaching fellow in epidemiology and public health,7 Melanie J Davies, professor of diabetes medicine,2 Kamlesh Khunti, professor of primary care diabetes and vascular medicine3

ABSTRACT
Objective To compare four potential screening strategies, and subsequent interventions, for the prevention and treatment of type 2 diabetes: (a) screening for type 2 diabetes to enable early detection and treatment, (b) screening for type 2 diabetes and impaired glucose tolerance, intervening with lifestyle interventions in those with a diagnosis of impaired glucose tolerance to delay or prevent diabetes, (c) as for (b) but with pharmacological interventions, and (d) no screening.
Design Cost effectiveness analysis based on development and evaluation of probabilistic, comprehensive economic decision analytic model, from screening to death.
Setting A hypothetical population, aged 45 at time of screening, with above average risk of diabetes.
Data sources Published clinical trials and epidemiological studies retrieved from electronic bibliographic databases; supplementary data obtained from the Department of Health statistics for England and Wales, the screening those at risk (STAR) study, and the Leicester division of the ADDITION study.
Methods A hybrid decision tree/Markov model was developed to simulate the long term effects of each screening strategy, in terms of both clinical and cost effectiveness outcomes. The base case model assumed a 50 year time horizon with discounting of both costs and benefits at 3.5%. Sensitivity analyses were carried out to investigate assumptions of the model and to identify which model inputs had most impact on the results.
Results Estimated costs for each quality adjusted life year (QALY) gained (discounted at 3.5% a year for both costs and benefits) were £14 150 (£17 560; $27 860) for screening for type 2 diabetes, £6242 for screening for diabetes and impaired glucose tolerance followed by lifestyle interventions, and £7023 for screening for diabetes and impaired glucose tolerance followed by pharmacological interventions, all compared with no screening. At a willingness-to-pay threshold of £20 000 the probability of the intervention being cost effective was 49%, 93%, and 85% for each of the active screening strategies respectively.

Conclusions Screening for type 2 diabetes and impaired glucose tolerance, with appropriate intervention for those with impaired glucose tolerance, in an above average risk population aged 45, seems to be cost effective. The cost effectiveness of a policy of screening for diabetes alone, which offered no intervention to those with impaired glucose tolerance, is still uncertain.

INTRODUCTION
In 2000, an estimated 171 million people worldwide had diabetes and numbers are projected to double by 2030.1 Life expectancy in people with diabetes might be shortened by as much as 15 years.2 Currently there is no systematic or structured screening policy for type 2 diabetes in the United Kingdom, though some general guidance has recently been issued by the National Screening Committee.3 One approach to screening would be to screen only for type 2 diabetes, which will allow for early diagnosis and treatment. This might be important as early detection and treatment could prevent future associated microvascular and macrovascular complications. An estimated 50% of people with diabetes are currently undiagnosed,4 and at presentation around 20-30% have already developed complications.5 An alternative screening approach would be to lower the threshold of the screening test and to screen for impaired glucose tolerance and type 2 diabetes together. As well as allowing for earlier diagnosis of type 2 diabetes, interventions can be administered to those identified with impaired glucose tolerance to attempt to delay the onset of type 2 diabetes. A recent systematic review and meta-analysis of intervention trials for prevention of type 2 diabetes6 found both lifestyle and pharmacological interventions significantly reduced the risk of type 2 diabetes in people with impaired glucose tolerance.

As no definitive trials have examined the effectiveness of screening for type 2 diabetes or impaired glucose tolerance,7,8 assessment of such policies has so far been conducted through simulation studies. Several decision models have been compiled that have developed to simulate the long term effects of each screening strategy, in terms of both clinical and cost effectiveness outcomes. The base case model assumed a 50 year time horizon with discounting of both costs and benefits at 3.5%. Sensitivity analyses were carried out to investigate assumptions of the model and to identify which model inputs had most impact on the results.

Results Estimated costs for each quality adjusted life year (QALY) gained (discounted at 3.5% a year for both costs and benefits) were £14 150 (£17 560; $27 860) for screening for type 2 diabetes, £6242 for screening for diabetes and impaired glucose tolerance followed by lifestyle interventions, and £7023 for screening for diabetes and impaired glucose tolerance followed by pharmacological interventions, all compared with no screening. At a willingness-to-pay threshold of £20 000 the probability of the intervention being cost effective was 49%, 93%, and 85% for each of the active screening strategies respectively.

Conclusions Screening for type 2 diabetes and impaired glucose tolerance, with appropriate intervention for those with impaired glucose tolerance, in an above average risk population aged 45, seems to be cost effective. The cost effectiveness of a policy of screening for diabetes alone, which offered no intervention to those with impaired glucose tolerance, is still uncertain.
interventions to prevent type 2 diabetes\textsuperscript{9-16} or strategies for screening and early detection of diabetes.\textsuperscript{7,17,20} Previous models of screening for type 2 diabetes alone have generally assessed the impact of early treatment on cardiovascular events, though some additionally included microvascular events such as retinopathy. Overall most of the models produced favourable results for screening, but cost effectiveness varied with age group screened and the population targeted for screening. Only two studies reported costs for a UK setting,\textsuperscript{7,19} one of which had a limited time horizon of five years.\textsuperscript{19} Both of these studies concluded there was still uncertainty concerning the cost effectiveness of screening for diabetes.

Of the eight models assessing cost effectiveness of interventions for prevention of diabetes, only three included costs of identifying individuals with impaired glucose tolerance.\textsuperscript{10,12,16} The time horizon over which the models were run ranged from just three years after the intervention up to the expected lifetime of the population. Models used data from various sources from published trials, epidemiological studies, and national statistics. In general data were limited to a few sources. All models compared a strategy of interventions against no interventions, rather than screening for impaired glucose tolerance followed by interventions, with no screening. All but one model simulated populations where all individuals had impaired glucose tolerance at the start of the model and the end state was development of diabetes, or death, hence only a limited section of the disease pathway was modelled. Also the models did not take into account that screening for impaired glucose tolerance will at the same time allow individuals with undiagnosed diabetes to be identified, thus allowing for early treatment and possibly reducing rates of complications. Hence, while these studies offer an assessment of the cost effectiveness of interventions for prevention of diabetes, none assessed the impact of screening followed by interventions on the whole disease pathway. In 2007 Waugh et al assessed screening or intervention strategies for type 2 diabetes in a thorough review of previous decision models.\textsuperscript{7}

We compared three active screening strategies: (a) a one-off screening for type 2 diabetes; (b) screening for impaired glucose tolerance and type 2 diabetes and intervening with lifestyle interventions in those with impaired glucose tolerance; and (c) as for (b) but with pharmacological interventions. We compared these three active screening strategies against a fourth strategy of no screening (current practice). The full pathway from screening, to interventions and treatment for type 2 diabetes, all the way through to death, was modelled. This model directly compares the two alternative approaches of screening for type 2 diabetes alone or screening for impaired glucose tolerance and type 2 diabetes together. When modelling the effectiveness of interventions, we used all data from relevant randomised controlled trials\textsuperscript{8} and included uncertainty around model inputs when appropriate. By carrying out several sensitivity analyses we investigated the essential elements that affect the cost and clinical effectiveness of different screening policies.

**METHODS**

The hybrid model consists of a decision tree and a Markov model (fig 1). The decision tree comprises three main arms, representing no screening, screening for undiagnosed type 2 diabetes, and screening for impaired glucose tolerance and undiagnosed diabetes, with either lifestyle or pharmacological interventions applied in those with impaired glucose tolerance.
Individuals who have already been identified as having type 2 diabetes are excluded from the screening process. The decision tree uses prevalence of impaired glucose tolerance and undiagnosed type 2 diabetes and estimates sensitivity and specificity of a screening test to determine how many individuals from the population start in each state of the Markov model. The Markov model consists of seven states: normal glucose tolerance, impaired glucose tolerance, type 2 diabetes, and undetected type 2 diabetes.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Value (SE)</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalences</td>
<td>Dirichlet</td>
<td>Normal glucose tolerance 83%; impaired glucose tolerance 12%; type 2 diabetes 5%</td>
<td>STAR study 24</td>
</tr>
<tr>
<td>Screening test efficiency</td>
<td>Multi-nominal</td>
<td>For type 2 diabetes: sensitivity 89.5%, specificity 91.3%; for impaired glucose tolerance and type 2 diabetes: sensitivity 59.4%, specificity 88.0%</td>
<td>STAR study 24</td>
</tr>
</tbody>
</table>

### Transition rates (per 100 person years)

- Normal to impaired glucose tolerance:
  - <65 years: Log normal 1.66 (0.08)
  - ≥65 years: Log normal 2.49 (0.11)
- Impaired glucose tolerance to type 2 diabetes: Log normal 1.96 (0.25)
- Time spent with undetected diabetes (years): Log normal 1.65 (0.68)

### Mortality rates (per 100 person years)

- 45-54 years: 0.32
- 55-64 years: 0.84
- 65-74 years: 2.36
- 75-84 years: 6.09
- ≥85 years: 15.68

### Increased risk of death with diabetes (hazard ratio)

- Log normal 0.756 (0.087)
- Log normal 0.104 (0.039)

### Intervention effects on risk of developing type 2 diabetes (hazard ratio)

- Lifestyle v standard treatment: Log normal −0.646 (0.099)
- Antidiabetic drugs v placebo: Log normal −0.425 (0.141)
- Hba1c:
  - Undiagnosed diabetes Normal 9.0% (0.056)
  - Screen detected diabetes Normal 7.0% (0.028)
  - Clinically detected diabetes Normal 7.9% (0.042)
- Costs:
  - Screening tests:
    - FPG test: £0.40/person
    - OGTT test: £1.30/person
  - Nurse cost: £26/hour
  - Metformin intervention: £16.10/year
  - Lifestyle intervention:
    - Year 1: £398/year
    - Subsequent years: £280/year
  - Undiagnosed diabetes:
    - Year before diagnosis: £114/year
    - Years 2-5 before diagnosis: £22/year
  - Diagnosed diabetes:
    - Screen detected: £2490 (53.3)/year
    - Clinically detected: £2756 (63.1)/year

FPG = fasting plasma glucose; OGTT = oral glucose tolerance test.
*Costs are standardised to 2006.
†Constant for all time spent with undetected type 2 diabetes.
‡Starting utility, which was then decreased for each year spent with diabetes because of predicted increases in complications, based on UKPDS data.

Table 1 | Estimates used to determine parameters for decision model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Value (SE)</th>
<th>Source(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data for decision tree</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalences</td>
<td>Dirichlet</td>
<td>Normal glucose tolerance 83%; impaired glucose tolerance 12%; type 2 diabetes 5%</td>
<td>STAR study 24</td>
</tr>
<tr>
<td>Screening test efficiency</td>
<td>Multi-nominal</td>
<td>For type 2 diabetes: sensitivity 89.5%, specificity 91.3%; for impaired glucose tolerance and type 2 diabetes: sensitivity 59.4%, specificity 88.0%</td>
<td>STAR study 24</td>
</tr>
<tr>
<td>Transition rates (per 100 person years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal to impaired glucose tolerance:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65 years</td>
<td>Log normal</td>
<td>1.66 (0.08)</td>
<td>Baltimore study</td>
</tr>
<tr>
<td>≥65 years</td>
<td>Log normal</td>
<td>2.49 (0.11)</td>
<td></td>
</tr>
<tr>
<td>Impaired glucose tolerance to type 2 diabetes</td>
<td>Log normal</td>
<td>1.96 (0.25)</td>
<td>12 studies</td>
</tr>
<tr>
<td>Time spent with undetected diabetes (years)</td>
<td>Log normal</td>
<td>1.65 (0.68)</td>
<td>Harris</td>
</tr>
<tr>
<td>Mortality rates (per 100 person years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45-54 years</td>
<td>—</td>
<td>0.32</td>
<td>DoH statistics (2000)</td>
</tr>
<tr>
<td>55-64 years</td>
<td>—</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>65-74 years</td>
<td>—</td>
<td>2.36</td>
<td></td>
</tr>
<tr>
<td>75-84 years</td>
<td>—</td>
<td>6.09</td>
<td></td>
</tr>
<tr>
<td>≥85 years</td>
<td>—</td>
<td>15.68</td>
<td></td>
</tr>
<tr>
<td>Increased risk of death with diabetes (hazard ratio)</td>
<td>Log normal</td>
<td>0.756 (0.087)</td>
<td>DECODE</td>
</tr>
<tr>
<td>Increased risk of death for 1% increase in Hba1c (hazard ratio)</td>
<td>Log normal</td>
<td>0.104 (0.039)</td>
<td>Rossing</td>
</tr>
<tr>
<td>Intervention effects on risk of developing type 2 diabetes (hazard ratio)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifestyle v standard treatment</td>
<td>Log normal</td>
<td>−0.646 (0.099)</td>
<td>12 studies</td>
</tr>
<tr>
<td>Antidiabetic drugs v placebo</td>
<td>Log normal</td>
<td>−0.425 (0.141)</td>
<td>9 studies</td>
</tr>
<tr>
<td>Hba1c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed diabetes</td>
<td>Normal</td>
<td>9.0% (0.056)</td>
<td>UKPDS</td>
</tr>
<tr>
<td>Screen detected diabetes</td>
<td>Normal</td>
<td>7.0% (0.028)</td>
<td>UKPDS</td>
</tr>
<tr>
<td>Clinically detected diabetes</td>
<td>Normal</td>
<td>7.9% (0.042)</td>
<td>UKPDS</td>
</tr>
<tr>
<td>Utilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed diabetes</td>
<td>Normal</td>
<td>0.788 (0.020)†</td>
<td>ADDITION</td>
</tr>
<tr>
<td>Screen detected diabetes</td>
<td>Normal</td>
<td>0.788 (0.020)‡</td>
<td>ADDITION</td>
</tr>
<tr>
<td>Clinically detected diabetes</td>
<td>Normal</td>
<td>0.771 (0.035)‡</td>
<td>UKPDS</td>
</tr>
<tr>
<td>Costs*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening tests:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FPG test</td>
<td>—</td>
<td>£0.40/person</td>
<td>NHS (2006)</td>
</tr>
<tr>
<td>OGTT test</td>
<td>—</td>
<td>£1.30/person</td>
<td></td>
</tr>
<tr>
<td>Nurse cost</td>
<td>—</td>
<td>£26/hour</td>
<td>Curtis</td>
</tr>
<tr>
<td>Lifestyle intervention:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1</td>
<td>—</td>
<td>£398/year</td>
<td>Avenell</td>
</tr>
<tr>
<td>Subsequent years</td>
<td>—</td>
<td>£280/year</td>
<td></td>
</tr>
<tr>
<td>Undiagnosed diabetes:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year before diagnosis</td>
<td>—</td>
<td>£114/year</td>
<td>Gulliford</td>
</tr>
<tr>
<td>Years 2-5 before diagnosis</td>
<td>—</td>
<td>£22/year</td>
<td>Curtis</td>
</tr>
<tr>
<td>Diagnosed diabetes:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screen detected</td>
<td>Normal</td>
<td>£2490 (53.3)/year</td>
<td>UKPDS</td>
</tr>
<tr>
<td>Clinically detected</td>
<td>Normal</td>
<td>£2756 (63.1)/year</td>
<td>UKPDS</td>
</tr>
</tbody>
</table>

*Costs are standardised to 2006.
†Constant for all time spent with undetected type 2 diabetes.
‡Starting utility, which was then decreased for each year spent with diabetes because of predicted increases in complications, based on UKPDS data.
tolerance, undiagnosed impaired glucose tolerance, diagnosed impaired glucose tolerance, death, and three states for people with diabetes (undiagnosed, diagnosed clinically, or diagnosed through screening, either from a screening test or because they are diagnosed with impaired glucose tolerance initially and hence enter a surveillance programme). We ran four Markov models simultaneously, one for each of the screening strategies. Whether type 2 diabetes and impaired glucose tolerance are diagnosed or undiagnosed determines whether the patients receive relevant treatments or interventions, and they are modelled accordingly in terms of transition rates to other states. For example, individuals identified with impaired glucose tolerance receive an intervention and the estimated intervention effect slows their progression to development of diabetes. Each model cycle represents one year and the model is run for a time horizon of 50 years. Table 1 summarised all the model inputs. When more than one estimate was available for a parameter, we pooled estimates using a Bayesian random effects meta-analysis within the comprehensive decision model. Model results include both clinical and cost effectiveness outcomes, with cost per quality adjusted life year (QALY) being the primary outcome. We investigated both an undiscounted model and a model with costs and benefits discounted at 3.5% annually, as recommended by the National Institute for Health and Clinical Excellence.21

The hybrid model was implemented within WinBUGS using a Bayesian comprehensive decision modelling approach.22 We adopted this approach because of its flexibility in terms of statistical modelling and it enabled us to include and propagate all uncertainty in parameters throughout the model.22 We assumed non-informative prior distributions for all model parameters. Model parameters were estimated by using Markov chain Monte Carlo simulation methods.23 Results are based on a sample of 20 000 simulations, following a “burn in” of 10 000, and we assessed convergence of the Markov chain by visually inspecting trace plots and by running multiple chains with different initial values.23 We have reported the results from the decision model with 95% credibility intervals, which are analogous to confidence intervals.

Data for the decision tree
The base case scenario for the model was a one-off screening for a population aged 45, in whom type 2 diabetes had not previously been diagnosed. Data for the decision tree—that is, test sensitivity and specificity and prevalence of impaired glucose tolerance and type 2 diabetes—were taken from the screening those at risk (STAR) study.24 For this study, individuals aged 40-75

| Table 2 | Clinical and cost outcomes from decision model, where prevalence of impaired glucose tolerance was 15% and type 2 diabetes 7.5%, and sensitivity and specificity of screening tests was 85% and 80%, respectively. Figures are mean values per person (95% credible intervals) for no screening and mean difference from or compared with no screening (95% credible intervals) for all other strategies |
| --- | --- | --- | --- | --- |
| Undiscounted | No screening | Screening for diabetes only | Lifestyle interventions | Pharmacological interventions |
| Total life years | 30.34 (27.75 to 32.86) | 0.06 (0.02 to 0.12) | 0.15 (0.08 to 0.22) | 0.13 (0.06 to 0.20) |
| QALYs | 28.06 (23.49 to 32.01) | 0.07 (<0.03 to 0.18) | 0.22 (0.08 to 0.36) | 0.17 (0.03 to 0.32) |
| Years spent without diabetes | 20.85 (19.63 to 29.45) | — | 0.33 (0.21 to 0.43) | 0.20 (0.10 to 0.37) |
| Lifetime risk of diabetes (%) | 64.55 (18.02 to 91.83) | — | -0.98 (<0.50 to -1.42) | -0.54 (<0.21 to -1.17) |
| Total cost | £17 290 (5746 to 39580) | 730 (<9 to 2341) | 610 (<373 to 2693) | 579 (<428 to 2658) |
| Cost per life year gained | — | 11 460 | 4179 | 4768 |
| Cost per QALY gained | — | 8681 | 2863 | 3429 |
| Cost per case prevented | — | 62 810 | 105 000 |
| Probability of cost effectiveness at willingness to pay threshold per QALY (%): £20 000 | 68.1 | 98.6 | 94.7 |
| £30 000 | 76.5 | 99.6 | 97.3 |
| Discounted at 3.5% a year for both costs and benefits | Total life years | 18.19 (17.25 to 18.98) | 0.02 (<0.01 to 0.05) | 0.05 (0.03 to 0.08) | 0.05 (0.02 to 0.07) |
| QALYs | 17.13 (15.02 to 18.49) | 0.03 (<0.02 to 0.09) | 0.09 (0.03 to 0.17) | 0.07 (0.01 to 0.15) |
| Years spent diabetes free | 13.69 (7.99 to 17.08) | — | 0.17 (0.11 to 0.23) | 0.11 (0.06 to 0.19) |
| Total cost | £7636 (2636 to 19 370) | 587 (61 to 1525) | 580 (<103 to 1760) | 528 (<163 to 1719) |
| Cost per life year gained | — | 23 710 | 10 900 | 11 690 |
| Cost per QALY gained | — | 14 150 | 6242 | 7023 |
| Probability of cost effectiveness at willingness to pay threshold per QALY (%): £20 000 | 48.6 | 91.0 | 85.0 |
| £30 000 | 60.8 | 97.4 | 91.6 |
(white) or 25-75 (non-white) from 15 general practices in Leicestershire who had at least one recognised risk factor for type 2 diabetes were invited for screening. Risk factors included a known history of coronary heart disease, hypertension, dyslipidaemia, cerebrovascular disease, a first degree relative with type 2 diabetes, and a body mass index (BMI) >25. Therefore the screening data included in the primary model were from a population considered to be “at risk” of type 2 diabetes. For the base case model we used data only from white patients, though we used the data on South Asians for sensitivity analyses to assess results for different ethnic groups.

Transition rates and HbA1c concentrations
To estimate annual transition rates we used several sources, including epidemiological studies and clinical trials.25-36 To estimate the annual transition rate from undiagnosed to clinically diagnosed diabetes, we used the estimated average time people have diabetes before being diagnosed.37 We estimated the effects of interventions on the transition from impaired glucose tolerance to diabetes using studies identified in a recent meta-analysis of lifestyle and pharmacological intervention trials.6 Death rates were taken from Department of Health statistics for England and Wales for 2000 and were increased for people with diabetes compared with those without.38 For the three diabetic states (undiagnosed, clinically diagnosed, and screen detected) death rates varied depending on predicted HbA1c (haemoglobin A1c) concentrations.39 HbA1c was predicted to be highest in people with undiagnosed diabetes, as they are yet to receive any interventions, and was estimated by using HbA1c concentrations at entry to the UK prospective diabetes study40 before treatment began. We expected HbA1c concentrations to be the best controlled in people with diabetes detected by screening because of early detection, and estimated levels using the 10 year average from the intensively treated group in the UK prospective diabetes study.41 For people with clinically diagnosed diabetes, we used the HbA1c concentrations of the group receiving conventional treatment in the UK prospective diabetes study.41

Quality of life variables
For the states of normal glucose tolerance, undiagnosed impaired glucose tolerance, and diagnosed impaired glucose tolerance, we assumed the utility value to be that of full health and set at 1. We calculated

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Results (undiscounted) of sensitivity analyses for varying prevalence rates of impaired glucose tolerance, normal glucose tolerance, and type 2 diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence*</td>
<td>No screening</td>
</tr>
<tr>
<td>QALY</td>
<td>28.06 (23.49 to 32.01)</td>
</tr>
<tr>
<td></td>
<td>70/20/10</td>
</tr>
<tr>
<td></td>
<td>10/60/30</td>
</tr>
<tr>
<td>Total cost (£)</td>
<td>83/12/5</td>
</tr>
<tr>
<td></td>
<td>70/20/10</td>
</tr>
<tr>
<td></td>
<td>10/60/30</td>
</tr>
<tr>
<td>Cost per QALY gained (£)</td>
<td>83/12/5</td>
</tr>
<tr>
<td></td>
<td>70/20/10</td>
</tr>
<tr>
<td></td>
<td>10/60/30</td>
</tr>
<tr>
<td>Probability (%) of being cost effective at willingness to pay threshold of £20 000/£30 000 per QALY</td>
<td>83/12/5</td>
</tr>
<tr>
<td></td>
<td>70/20/10</td>
</tr>
<tr>
<td></td>
<td>10/60/30</td>
</tr>
</tbody>
</table>

*Impaired glucose tolerance/normal glucose tolerance/type 2 diabetes.
utilities for those with undiagnosed and screen detected diabetes from EQ-5D data, using data on individual patients made available by the Leicester arm of the ADDITION study.42 The data were of a screen detected sample population with type 2 diabetes at baseline. For people with clinically diagnosed diabetes, utilities were taken from those reported by the UK prospective diabetes study as this comprised a clinically detected sample.43 The utility for undiagnosed diabetes was kept constant for the whole duration spent in this state as we assumed that if complications developed, which reduced the quality of life, then a diagnosis would be made. For the states of clinically and screen detected diabetes we needed to account for the fact that duration of diabetes would lead to an increased number of complications and hence a reduction in the utility value. This was done by using reported complication rates, modelled for duration of diabetes and adjusted for estimated HbA1c concentrations in each group and their estimated effect on utility values.44,45 Hence, utilities decreased for each year of duration of diabetes, to reflect increasing incidence of complications. Because of a higher predicted HbA1c concentration, the utility value was lower at diagnosis and decreased marginally more rapidly in individuals clinically diagnosed compared with those who were screen detected.

Economic variables
We estimated costs from various sources. Screening costs included the costs of an initial screening test of fasting plasma glucose and a confirmatory oral glucose tolerance test in those who tested positive. We estimated the cost of nurse time of 5 minutes for the screening test and 25 minutes for the oral glucose tolerance test.45 People with undiagnosed diabetes incur costs before diagnosis because of increased visits to the general practitioner and prescriptions,46 with a

<table>
<thead>
<tr>
<th>Compliance with screening (%)</th>
<th>No screening</th>
<th>Screening for type 2 diabetes only</th>
<th>Screening for type 2 diabetes and impaired glucose tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>QALY:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.12 (23.58 to 32.08)</td>
<td>28.26 (23.74 to 32.23)</td>
</tr>
<tr>
<td>70</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.07 (23.52 to 32.05)</td>
<td>28.17 (23.64 to 32.16)</td>
</tr>
<tr>
<td>50</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.04 (23.51 to 32.04)</td>
<td>28.13 (23.61 to 32.13)</td>
</tr>
<tr>
<td>Total cost (£):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 040 (7083 to 39 970)</td>
<td>17 910 (7124 to 39 740)</td>
</tr>
<tr>
<td>70</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 070 (6777 to 39 800)</td>
<td>18 080 (6957 to 39 620)</td>
</tr>
<tr>
<td>50</td>
<td>17 290 (5746 to 39 580)</td>
<td>17 870 (6409 to 39 750)</td>
<td>17 930 (6705 to 39 680)</td>
</tr>
<tr>
<td>Cost (£) per QALY gained:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>—</td>
<td>8681</td>
<td>2863</td>
</tr>
<tr>
<td>70</td>
<td>—</td>
<td>8732</td>
<td>3112</td>
</tr>
<tr>
<td>50</td>
<td>—</td>
<td>8743</td>
<td>3515</td>
</tr>
<tr>
<td>Probability of being cost effective at willingness to pay threshold of £20 000/£30 000 per QALY (%):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>—</td>
<td>68/76</td>
<td>99/100</td>
</tr>
<tr>
<td>70</td>
<td>—</td>
<td>69/77</td>
<td>98/99</td>
</tr>
<tr>
<td>50</td>
<td>—</td>
<td>68/77</td>
<td>97/98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compliance with interventions (%)</th>
<th>No screening</th>
<th>Screening for type 2 diabetes only</th>
<th>Screening for type 2 diabetes and impaired glucose tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>QALY:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.12 (23.58 to 32.08)</td>
<td>28.26 (23.74 to 32.23)</td>
</tr>
<tr>
<td>70</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.12 (23.58 to 32.08)</td>
<td>28.22 (23.69 to 32.18)</td>
</tr>
<tr>
<td>50</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.12 (23.58 to 32.08)</td>
<td>28.19 (23.66 to 32.15)</td>
</tr>
<tr>
<td>Total cost (£):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 040 (7083 to 39 970)</td>
<td>17 910 (7124 to 39 740)</td>
</tr>
<tr>
<td>70</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 040 (7083 to 39 970)</td>
<td>18 140 (7343 to 39 950)</td>
</tr>
<tr>
<td>50</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 040 (7083 to 39 970)</td>
<td>18 261 (7455 to 39 050)</td>
</tr>
<tr>
<td>Cost (£) per QALY gained:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>—</td>
<td>2863</td>
<td>3429</td>
</tr>
<tr>
<td>70</td>
<td>—</td>
<td>4947</td>
<td>5039</td>
</tr>
<tr>
<td>50</td>
<td>—</td>
<td>6634</td>
<td>6243</td>
</tr>
<tr>
<td>Probability of being cost effective at willingness to pay threshold of £20 000/£30 000 per QALY (%):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>—</td>
<td>99/100</td>
<td>95/97</td>
</tr>
<tr>
<td>70</td>
<td>—</td>
<td>94/97</td>
<td>89/94</td>
</tr>
<tr>
<td>50</td>
<td>—</td>
<td>88/93</td>
<td>84/90</td>
</tr>
</tbody>
</table>
reported average of three additional visits the year before diagnosis and an average of 1.4 additional visits in the two to five years before diagnosis. An estimation of these costs was included.\(^4\) For lifestyle interventions we included dietitian costs and costs of twice weekly group exercise sessions, as detailed in a previous study.\(^9\) Costs of pharmacological interventions were based on 250 mg of metformin three times a day, the standard dose used by most intervention studies. For people with diagnosed diabetes, we took average annual costs of antidiabetic treatment, implementation of treatment, and costs of complications from the UK prospective diabetes study.\(^47\) For the people with diabetes detected at screening, in whom we would expect costs of complications to be lower, we used costs from the intensively treated arm of the UK prospective diabetes study.\(^9\) For the people with diabetes who are diagnosed currently, we used the reported costs of the conventionally treated group. All costs are reported in 2006 UK £, standardised by using inflation indices.\(^45\)

### Sensitivity analyses and model extensions

We carried out sensitivity analyses using a range of values of prevalence of disease, as well as compliance levels to both screening and interventions. Changing prevalence allows us to assess the effectiveness of the screening strategies for different “at risk” populations. The effects of compliance to both screening and interventions were also important as we assumed 100% compliance to both in the base case model, which could never be achieved in practice.

To evaluate the robustness of the model we also carried out sensitivity analyses on model inputs, particularly those that were estimated from only one or two sources or were thought to be important drivers in the model. These were sensitivities of screening tests, costs of interventions, costs of diabetes, effectiveness of interventions, previous distributions on the standard deviations between studies of the four meta-analyses run within the model, and the time horizon the model was run for.

For the base case scenario we considered only a one-off screening at age 45. The model was extended further to assess the impact of having one or two additional screenings, at age 50 and 60. This was done by applying the test sensitivities from the STAR study to the numbers in the states of undiagnosed impaired glucose tolerance and type 2 diabetes at the corresponding model cycle and moving the individuals to the relevant diagnosed state.

Though the base case model used prevalences and test sensitivities and specificities of a white population, the effect of screening a South Asian or a mixed race population is also relevant in the UK. South Asians are thought to have a greater risk of type 2 diabetes, with a greater prevalence of impaired glucose tolerance and a higher transition rate to type 2 diabetes. We extended this model with data from the STAR study and estimated the transition rate from impaired glucose tolerance to type 2 diabetes from the Indian diabetes prevention programme.\(^48\)

### RESULTS

Table 2 shows clinical and cost effectiveness outcomes for an undiscounted model and a model discounted for both costs and benefits at 3.5% a year. Discounted costs for each QALY gained, compared with no screening, were £14150 (€17560; $27860) for type 2 diabetes screening, £6242 for screening for diabetes and impaired glucose tolerance with lifestyle interventions, and £7023 for screening for both diabetes and impaired glucose tolerance with pharmacological interventions. Costs were lower in the undiscounted model: £8681,

### Table 5 | Results of model extensions for number of screens (undiscounted)

<table>
<thead>
<tr>
<th>No of screens</th>
<th>No screening</th>
<th>Screening for type 2 diabetes only</th>
<th>Screening for type 2 diabetes and impaired glucose tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lifestyle interventions</td>
</tr>
<tr>
<td>QALY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.12 (23.58 to 32.08)</td>
<td>28.26 (23.74 to 32.23)</td>
</tr>
<tr>
<td>2</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.13 (23.74 to 32.06)</td>
<td>28.56 (24.74 to 32.30)</td>
</tr>
<tr>
<td>3</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.15 (23.86 to 32.16)</td>
<td>28.80 (25.04 to 32.32)</td>
</tr>
<tr>
<td>Total cost (£)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 040 (7083 to 39 970)</td>
<td>17 910 (7124 to 39 740)</td>
</tr>
<tr>
<td>2</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 500 (7491 to 40 980)</td>
<td>19 300 (7570 to 41 160)</td>
</tr>
<tr>
<td>3</td>
<td>17 290 (5746 to 39 580)</td>
<td>19 670 (7735 to 42 110)</td>
<td>20 220 (7740 to 42 210)</td>
</tr>
<tr>
<td>Cost per QALY gained (£)</td>
<td>8681</td>
<td>2863</td>
<td>3429</td>
</tr>
<tr>
<td>2</td>
<td>9544</td>
<td>2777</td>
<td>3317</td>
</tr>
<tr>
<td>3</td>
<td>10360</td>
<td>2966</td>
<td>3517</td>
</tr>
<tr>
<td>Probability of being cost effective at willingness to pay threshold of £20 000/€30 000 per QALY (%)</td>
<td>68/76</td>
<td>99/100</td>
<td>95/97</td>
</tr>
<tr>
<td>2</td>
<td>57/66</td>
<td>99/100</td>
<td>96/98</td>
</tr>
<tr>
<td>3</td>
<td>54/64</td>
<td>99/100</td>
<td>97/99</td>
</tr>
</tbody>
</table>
The probability that these strategies were cost effective compared with no screening still remained high, with an estimated probability of 88% for screening with lifestyle interventions and 84% for screening with pharmacological interventions at the willingness to pay threshold of £20 000.

Other sensitivity analyses did not change the results enough to alter the conclusions of the model. Increasing the costs of both lifestyle and pharmacological interventions by a factor of 10 reduced the probabilities of cost effectiveness of their respective screening strategies to 73% and 93%, at the willingness to pay threshold of £20 000. Increasing the costs of diabetes by a factor of two reduced the probability of cost effectiveness to 49% for screening for type 2 diabetes only, 93% for screening with lifestyle interventions, and 85% for screening with pharmacological interventions at the willingness to pay threshold of £20 000.

Table 6 | Results of model extensions for different ethnic groups (undiscounted)

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>No screening</th>
<th>Screening for type 2 diabetes only</th>
<th>Lifestyle interventions</th>
<th>Pharmacological interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>QALY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>28.06 (23.49 to 32.01)</td>
<td>28.12 (23.58 to 32.08)</td>
<td>28.26 (23.74 to 32.23)</td>
<td>28.22 (23.69 to 32.18)</td>
</tr>
<tr>
<td>South Asian</td>
<td>25.24 (20.65 to 30.79)</td>
<td>25.35 (20.83 to 30.91)</td>
<td>25.47 (20.96 to 31.02)</td>
<td>25.63 (20.92 to 30.98)</td>
</tr>
<tr>
<td>Mixed*</td>
<td>27.10 (23.79 to 30.31)</td>
<td>27.18 (23.88 to 30.39)</td>
<td>27.32 (24.02 to 30.53)</td>
<td>27.27 (23.99 to 30.53)</td>
</tr>
<tr>
<td>Total cost (££)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>17 290 (5746 to 39 580)</td>
<td>18 040 (7083 to 39 970)</td>
<td>17 910 (7124 to 39 740)</td>
<td>17 900 (7061 to 39 710)</td>
</tr>
<tr>
<td>South Asian</td>
<td>28 250 (10 170 to 55 120)</td>
<td>29 390 (12 270 to 55 490)</td>
<td>29 420 (12 500 to 55 220)</td>
<td>29 480 (12 550 to 55 270)</td>
</tr>
<tr>
<td>Mixed*</td>
<td>22 145 (8345 to 41 657)</td>
<td>23 051 (9820 to 42 131)</td>
<td>22 973 (9809 to 41 962)</td>
<td>22 976 (11 885 to 42 006)</td>
</tr>
<tr>
<td>Cost per QALY gained (££)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>—</td>
<td>8681</td>
<td>2863</td>
<td>3429</td>
</tr>
<tr>
<td>South Asian</td>
<td>—</td>
<td>8168</td>
<td>4657</td>
<td>5643</td>
</tr>
<tr>
<td>Mixed*</td>
<td>—</td>
<td>8523</td>
<td>3555</td>
<td>4497</td>
</tr>
<tr>
<td>Probability of being cost effective at willingness to pay threshold of £20 000/£30 000 per QALY (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>—</td>
<td>68/76</td>
<td>99/100</td>
<td>95/97</td>
</tr>
<tr>
<td>South Asian</td>
<td>—</td>
<td>68/75</td>
<td>89/94</td>
<td>83/88</td>
</tr>
<tr>
<td>Mixed*</td>
<td>—</td>
<td>69/77</td>
<td>98/99</td>
<td>96/98</td>
</tr>
</tbody>
</table>

*Modelled as 30% South Asian and 70% white.

£2863, and £3429 for every QALY gained, respectively. At a willingness to pay threshold of £20 000 per QALY the probability of each strategy being cost effective was 49% for screening for type 2 diabetes only, 93% for screening for both diabetes and impaired glucose tolerance and lifestyle interventions, and 85% for screening for both diabetes and impaired glucose tolerance and pharmacological intervention. Figure 2 shows cost effectiveness acceptability curves, illustrating the probability of cost effectiveness over a range of willingness to pay thresholds.

Discounted QALYs gained compared with no screening were 0.03 (−0.02 to 0.09) for diabetes screening, 0.09 (0.03 to 0.17) for screening and lifestyle interventions, and 0.07 (0.01 to 0.15) for screening with pharmacological interventions. Both the intervention strategies showed potential benefits in terms of average years spent without diabetes and cases of diabetes prevented. Although clinical effects seem small, it must be remembered they are average gains across a population, in which only 17% had either impaired glucose tolerance or undiagnosed type 2 diabetes at the time of screening.

Tables 3 and 4 show the results of the more important sensitivity analyses (undiscounted). Increasing the prevalence of impaired glucose tolerance and type 2 diabetes decreased the QALYs and increased total costs of each screening strategy. The comparisons of the three active screening/intervention strategies compared with no screening remained fairly constant in terms of costs per QALY and probability of cost effectiveness (table 3). When we lowered compliance with screening, the impact on results was also minimal (table 4). Reducing compliance with interventions, however, had a greater impact in that the total costs and cost per QALY gained increased for both the screening/intervention strategies. The probability that these strategies were cost effective compared with no screening still remained high, with an estimated probability of 88% for screening with lifestyle interventions and 84% for screening with pharmacological interventions at the willingness to pay threshold of £20 000.

Other sensitivity analyses did not change the results enough to alter the conclusions of the model. Increasing the costs of both lifestyle and pharmacological interventions by a factor of 10 reduced the probabilities of cost effectiveness of their respective screening strategies to 73% and 93%, at the willingness to pay threshold of £20 000. Increasing the costs of diabetes by a factor of two reduced the probability of cost effectiveness to 49% for screening for type 2 diabetes only, 93% for screening with lifestyle interventions, and 85% for screening with pharmacological interventions at the same threshold. As we increased the time horizon the model was run for, the probability of the three active screening strategies being cost effective compared with no screening increased. This is because the benefits of screening or interventions are not all immediate and most occur in later years of the model, when type 2 diabetes is either delayed or complications are reduced through early diagnosis and treatment. The intervention strategies became cost effective when we considered a time horizon of at least 30 years (probability of being cost effective of 0.97 for lifestyle and 0.91 for pharmacological interventions at the willingness to pay threshold of £20 000). Overall, the model’s conclusions were robust to changes made to the sensitivity analyses, giving strength to the conclusions.

Tables 5 and 6 give the results of the model extensions as undiscounted estimates. Increasing the
Screening populations with a higher prevalence of glucose intolerance might result in better clinical outcomes, although cost effectiveness seems unaffected.

WHAT IS ALREADY KNOWN ON THIS TOPIC

In people with impaired glucose tolerance interventions are clinically and cost effective. Screening for type 2 diabetes to allow early detection might be cost effective in certain groups.

WHAT THIS STUDY ADDS

Modelling the whole screening and intervention pathway from screening to death shows that screening for type 2 diabetes and impaired glucose tolerance, followed by interventions, seems to be cost effective compared with no screening. Uncertainty still exists concerning the cost effectiveness of screening for type 2 diabetes alone.

Modelling the whole screening and intervention pathway from screening to death shows that screening for type 2 diabetes and impaired glucose tolerance, followed by interventions, seems to be cost effective compared with no screening. Uncertainty still exists concerning the cost effectiveness of screening for type 2 diabetes alone.

Screening populations with a higher prevalence of glucose intolerance might result in better clinical outcomes, although cost effectiveness seems unaffected.

Our model makes several assumptions. No transition was allowed from normal glucose tolerance to diabetes without first passing through impaired glucose tolerance. This is because it is clinically unlikely that an individual would change from normal glucose tolerance to diabetes within a year, which is one model cycle. No transition was allowed from diabetes back to impaired glucose tolerance or from impaired to normal glucose tolerance. This is clinically accurate because once an individual has a diagnosis of type 2 diabetes, even if their glucose tolerance improves, they are still clinically defined as having diabetes. Also once an individual has had impaired glucose tolerance, even if their glucose tolerance improves their future risk of diabetes is probably more similar to that in individuals with impaired glucose tolerance rather than those who have always had normal glucose tolerance.

Another assumption was that the HbA1c concentrations of those with diabetes who were clinically diagnosed would be similar to the 10 year average of an intensively treated group of people with diabetes from the UK prospective diabetes study. This assumption was made in the absence of long term clinical data on individuals whose diabetes was detected by screening. Although 10 year averages of HbA1c concentrations were used for people with diabetes, when we ran our model for longer time horizons the HbA1c concentrations were potentially underestimated, which means complication rates and their effects on utilities and mortality might also be moderately underestimated. Further data are needed on how HbA1c concentration could be expected to increase over time to allow more accurate modelling.

Screening costs incorporated within the model included only costs of the test and the nurse’s time, therefore representing the costs of opportunistic screening. We did not include further costs of establishing systematic screening, such as the identification of eligible patients, the issuing of invitations to screening, and the chasing up of non-attenders. In practice, these additional costs would be small for each individual screened, particularly if screening was incorporated into current health checks. When modelling costs of treatment and complications associated with diabetes, we used the average yearly costs taken from the UK prospective diabetes study. As costs would be expected to start off low and then increase, this means that costs of diabetes might be initially underestimated when an individual receives the diagnosis and eventually underestimated by this model. In addition, as average costs were used, we did not account for issues of competing risks of complications associated with diabetes. Unfortunately, yearly data on costs of diabetes, or how the occurrence of complications impacted on the probability of other complications occurring, were not available to enable us to model costs more accurately. The issue of competing risks arises not just for costs but also for the annual probabilities of complications. Ideally, we need data on individual patients to enable the correlation structure...
in both the probabilities and costs to be appropriately accounted for.

As we ran the model for a time horizon of 50 years, the screened population (aged 45 at the start) aged with each cycle of the model, thus, when possible, we incorporated time dependent model parameters. For some parameters, such as the treatment intervention effects, however, we assumed that the effect was constant over time. Additionally, although compliance was high in the intervention trials from which estimates of their effectiveness were obtained, it is still to be determined whether compliance could be maintained outside a trial setting. Therefore long term compliance with interventions is an important consideration. Sensitivity analyses of compliance with interventions found that even with compliance rates as low as 50%, the screening strategies involving either lifestyle or pharmacological interventions were still cost effective when compared with a strategy of no screening.

Conclusions

A policy of a one-off screening for type 2 diabetes and impaired glucose tolerance, with appropriate intervention for those identified with impaired glucose tolerance, seems to be cost effective in an “at risk” population. Changing compliance with screening or interventions or increasing the number of screenings did not change the conclusions of the model. Given the uncertainty in the results presented here, particularly for the assessment of screening for type 2 diabetes, further research is needed on the long term clinical effects of early diagnosis. Furthermore, to model the two strategies that involved interventions more accurately, we require additional information on long term compliance with interventions and their potential harms and benefits.

We thank the STAR study, in particular Jenny Tringham, and the Leicester arm of the ADDITION study, for providing the data on individual patients that were used for the analyses. We also thank Philip Clarke for advice on the UKPDS outcomes model.

Contributors: CLG performed the data extraction and analyses, wrote the first draft of the article, and is guarantor. KRA and PCL gave detailed advice at all stages of the analyses. All authors contributed to the writing of the paper and gave substantial advice and input into the study. KRA and KK had the initial idea for this project.

Funding: CLG is funded jointly by the UK Medical Research Council and the Economic and Social Research Council, under an interdisciplinary postgraduate research studentship in the social and medical sciences. NJC is funded by a Medical Research Council training fellowship in health services research.

Competing interests: MD and KK have received sponsorship for attending conferences and small honorariums and funding for research from pharmaceutical companies that manufacture hypoglycaemic and anti-obesity drugs. KRA has also received funding for research from pharmaceutical companies that manufacture hypoglycaemic and anti-obesity drugs and has acted as a paid consultant to consultancy companies who undertake work for the healthcare industry generally.

Ethical approval: Not required.

Provenance and peer review: Not commissioned; externally peer reviewed.


Accepted: 25 March 2008