



The
University
Of
Sheffield.

Status quo and trends within the field of quality adjusted life years

John Brazier

Professor of Health Economics

Health Economics and Decision Science

School of Health and Related Research

University of Sheffield

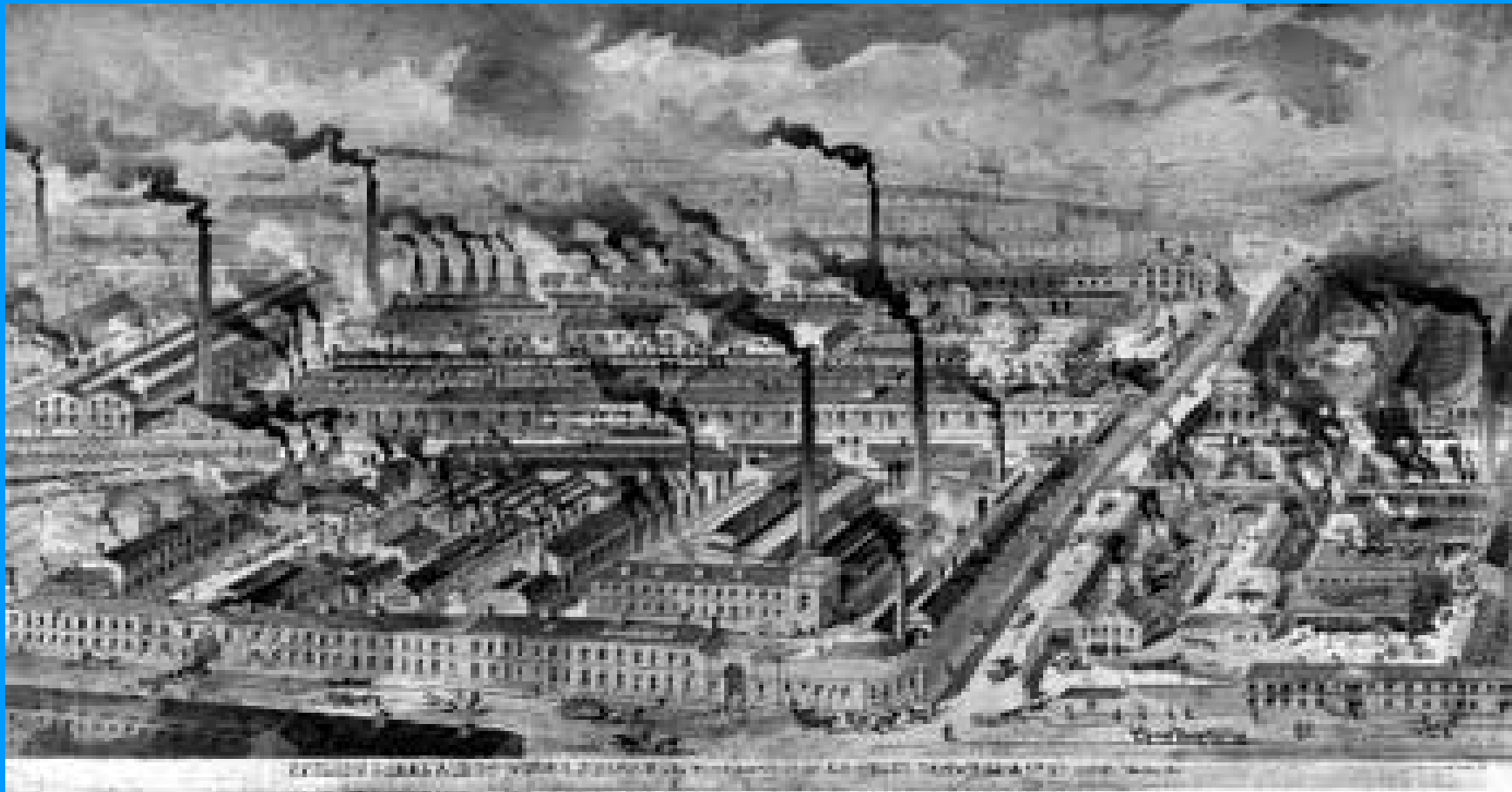
Presentation for the Symposium on Cost-effectiveness

June 17th 2008, Kaiserin-Friedrich-Stiftung, Berlin, Germany.



The
University
Of
Sheffield.

Sheffield – 19th century





The
University
Of
Sheffield.

Sheffield post-industrialisation....





The
University
Of
Sheffield.

Sheffield now.....





The
University
Of
Sheffield.

Sheffield now (2).....





The
University
Of
Sheffield.

The problem

A finite limit to resources

Plus

demands/needs exceed current (or
future) resources

implies the necessity for choice

How should these choices be made?



Resource allocation decisions

A new intervention:

- May be less effective and more costly... ✗
- More effective and less costly..... ✓
- **More effective and more costly....?**
- Plus less effective and less costly....?

→ For resource allocation decisions in health care we need a measure of benefit that allows comparisons to be made across treatments and patient groups

Appraising new interventions

New interventions are assessed by NICE and other agencies around the world in terms of:

- Clinical effectiveness
- Cost-effectiveness



Effectiveness measure?

- *Mortality*

Yes, but what about quality of life?

- *Clinical measures* – like blood glucose or bone mineral density

These do not tell us the real benefit to patients and provide no basis for making comparisons across programmes of care

- *Quality of life measures* like SF-36 or St George respiratory Questionnaire

Scoring systems do not reflect preferences, they are multi-dimensional (and so measure of overall effectiveness), and what about mortality?

→ Quality Adjusted Life Years



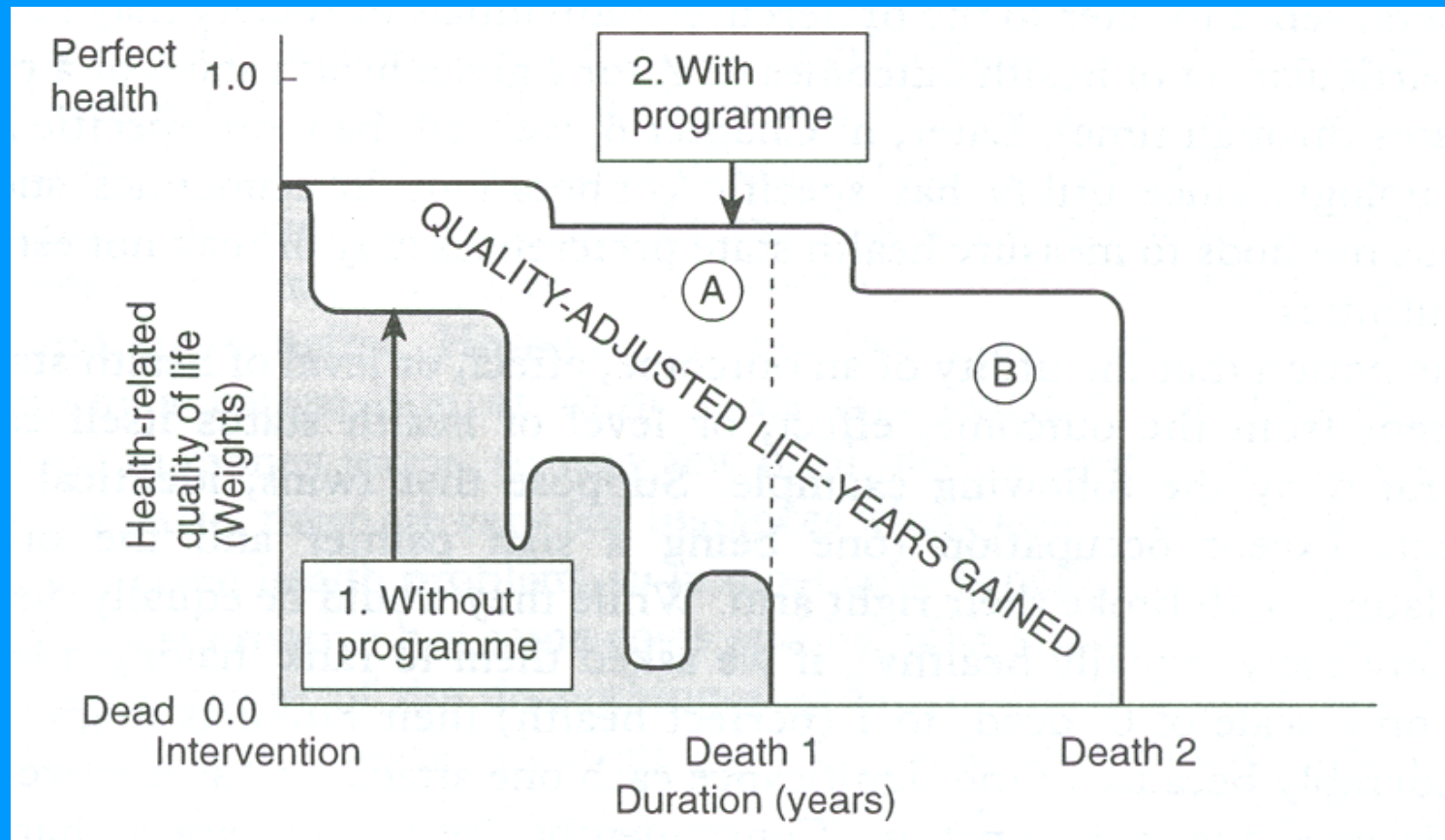
Quality Adjusted Life Years

The Quality Adjusted Life Year (QALY) combines quality of life and length of life into the single measure of benefit of a quality adjusted survival

- The 'Q' (or utility) is a value assigned to each health state from zero to one, where zero is for state equivalent to death and one for full health



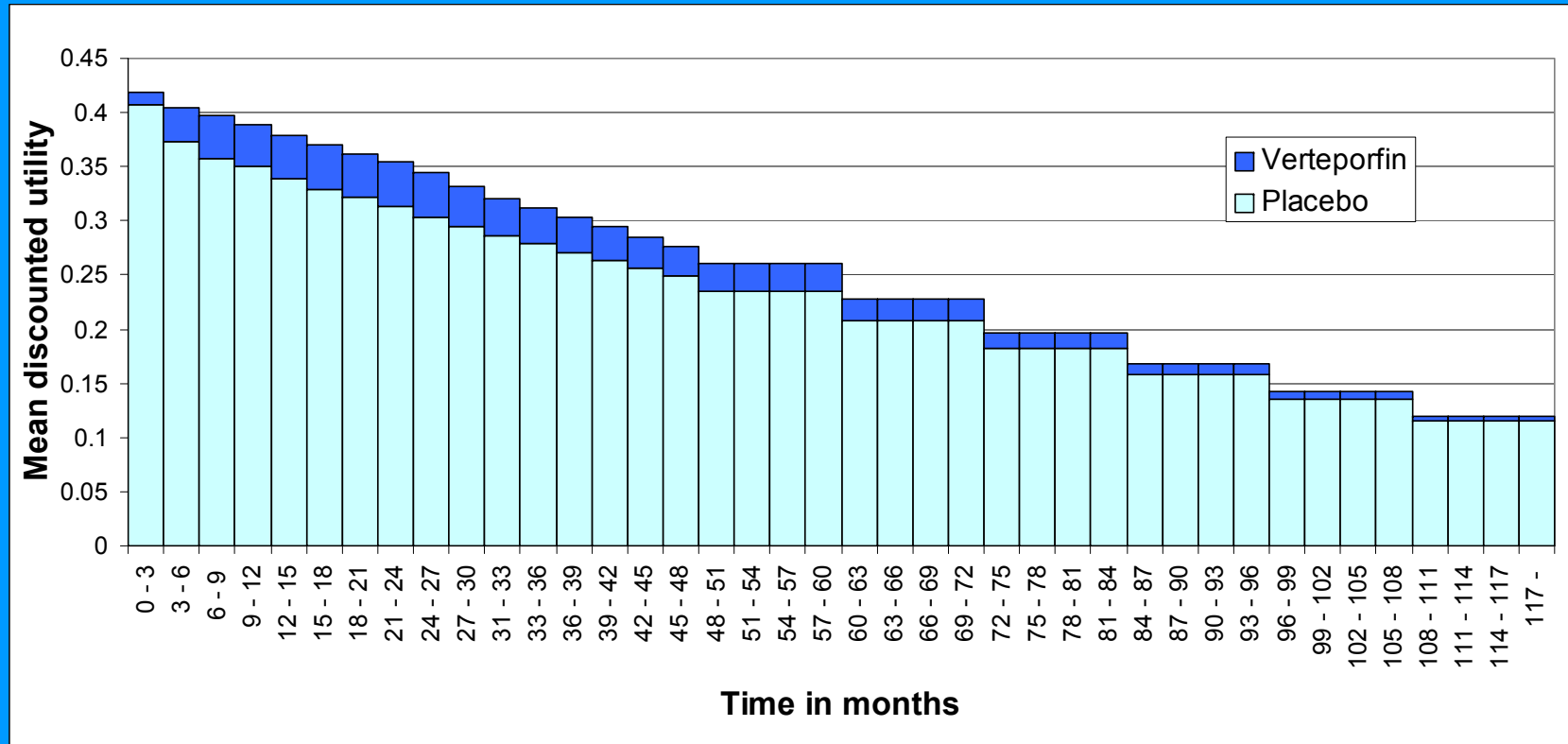
Quality-adjusted life years



Source: Drummond et al, 1997

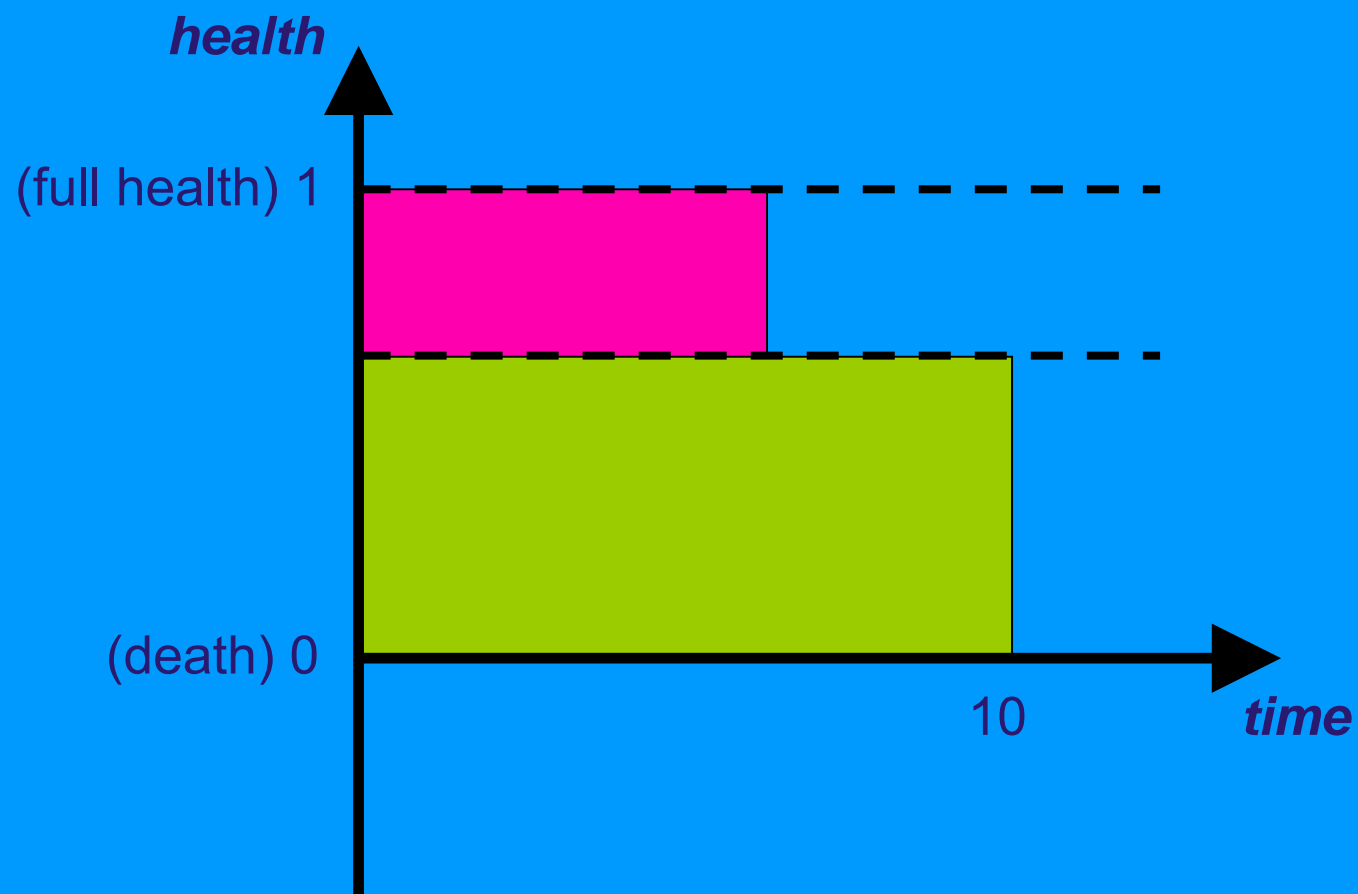


Calculating QALY gain of PDT on Macular Degeneration



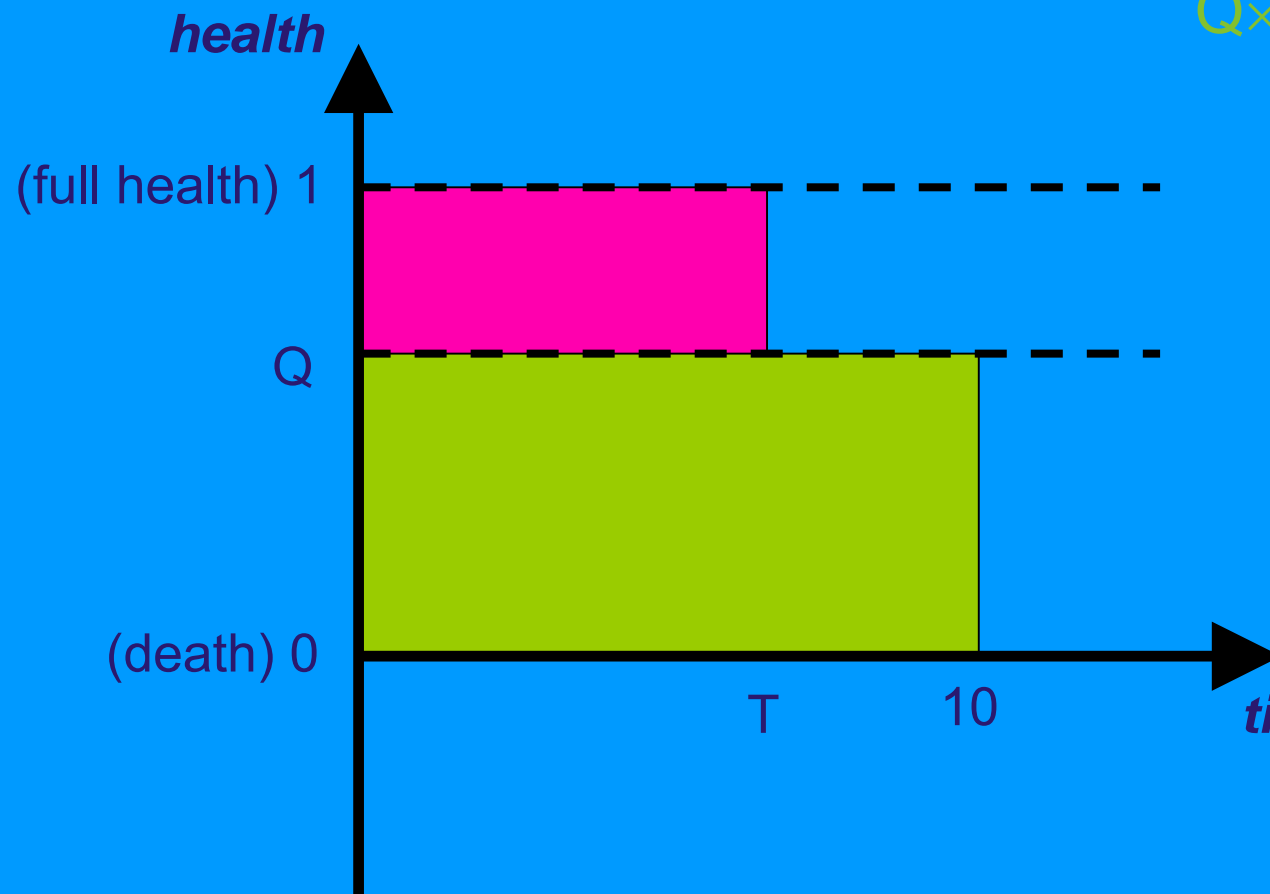


Trading off health and life





The time trade-off



$$Q \times 10 = 1 \times T$$

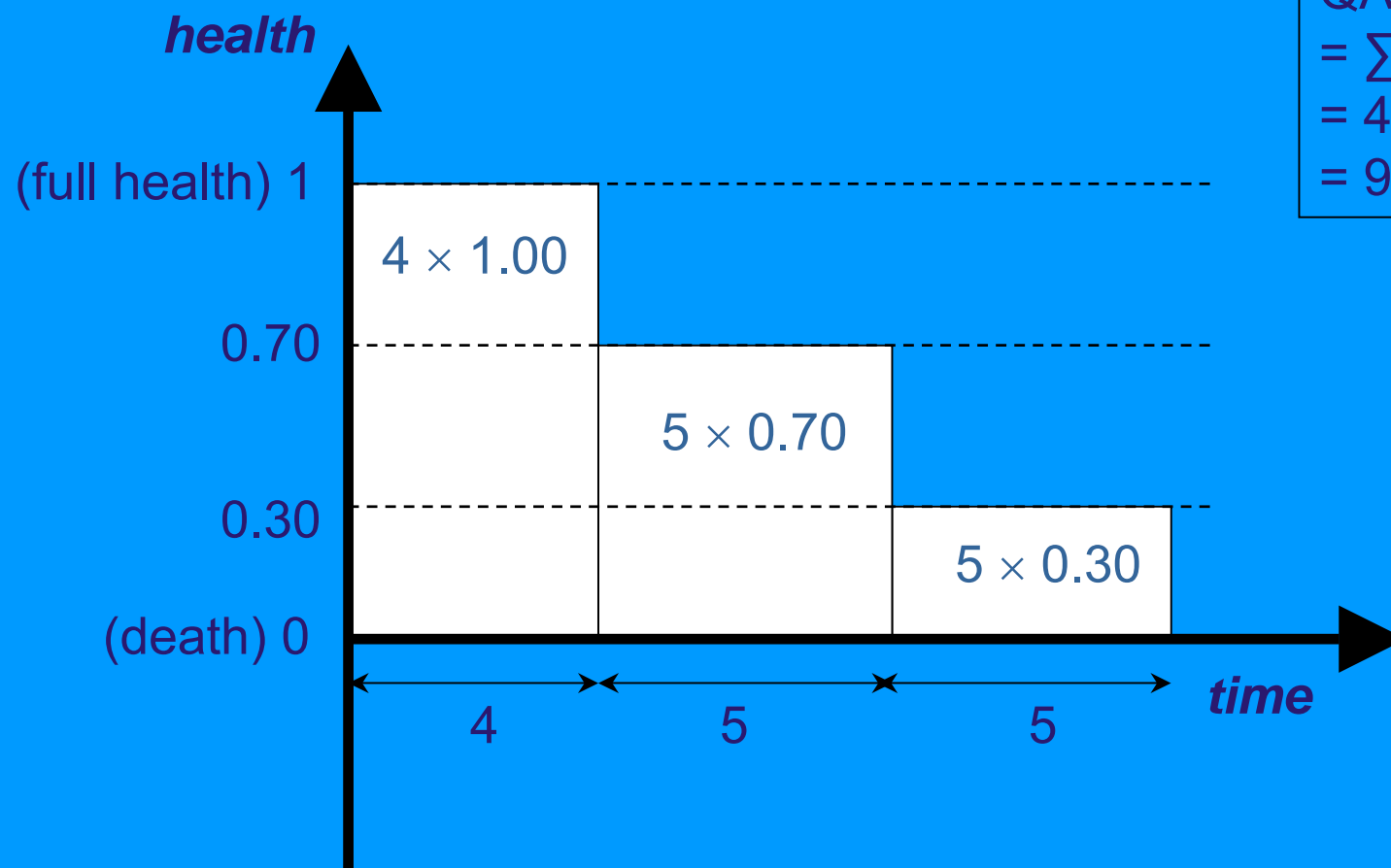
$$Q = T/10$$

If $T = 6$, then $Q = 0.6$,
and both options
provide 6 QALYs.

QALYs = total area



Health profiles



$$\begin{aligned} \text{QALYs} &= \sum_t Q(t) \\ &= 4 + 3.5 + 1.5 \\ &= 9 \end{aligned}$$



The
University
Of
Sheffield.

Valuing health: putting the 'q' into the QALY



The
University
Of
Sheffield.





The big questions

- 1. How should health be described?**
- 2. How should it to be valued?**
- 3. Who should value it?**



Here are some simple questions about your health in general. By ticking one answer in each group below, please indicate which statements best describe your own health state TODAY.

Please tick one

1. Mobility

I have no problems in walking about

I have some problems in walking about

I am confined to bed

2. Self-care

I have no problems with self-care

I have some problems washing or dressing myself

I am unable to wash or dress myself

3. Usual Activities

I have no problems with performing my usual activities
(e.g. work, study, housework, family or leisure activities)

I have some problems with performing my usual activities

I am unable to perform my usual activities

4. Pain/Discomfort

I have no pain or discomfort

I have moderate pain or discomfort

I have extreme pain or discomfort

5. Anxiety/Depression

I am not anxious or depressed

I am moderately anxious or depressed

I am extremely anxious or depressed



UK Valuation of EQ-5D health state

- 3395 interviews were conducted in respondents own home (response rate = 64%)
- highly representative sample of UK population)
- Each respondent valued 12 health states (out of 43) by time trade-off and visual analogue scale
- Statistical modeling used to value all 243 EQ-5D health states (from the 43)

Source: Dolan, 1997



UK Time trade-off value set

Dimension	Level	Coefficient
Constant		.081
Mobility	2	.069
	3	.314
Self-care	2	.104
	3	.214
Usual activity	2	.036
	3	.094
Pain/discomfort	2	.123
	3	.386
Anxiety/depression	2	.071
	3	.236
N3		.269

e.g. 11223 = $1.0 - (0.081 + .036 + .123 + .236 + .269) = .255$

NB. There is a German value set

Rationale for generic measures (e.g. EQ-5D)

Advantages

- **'Off-the-shelf'**
- **Cheap**
- **Convenient and comparatively easy to use in clinical trials (all self-completed except the QWB)**
- **Accepted by NICE and other agencies in 'reference case' analyses**

Disadvantages

- **May not be relevant or sensitive to the condition or treatment effects**

Generic measures: descriptive systems

	Dimension	Levels
QWB	Mobility, physical activity, social functioning 27 symptoms/problems	3 2
HUI3	Vision, Hearing, Speech, Ambulation, Dexterity, Emotion, Cognition, Pain	5-6
EQ-5D	Mobility, Self-care, Usual Activities, Pain/discomfort, Anxiety/depression	3
SF-6D	Physical functioning, role limitation, social functioning, pain, mental health and vitality	4-6
AQoL1	Independent living (self-care, household tasks, mobility), social relationships (intimacy, friendships, family role), physical senses (seeing, hearing, communication), psychological wellbeing (sleep, anxiety and depression, pain)	4
15D	Mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort/symptoms, depression, distress, vitality, sexual activity	4-5



Generic measures: valuation

	Valuation Technique	Countries
QWB	VAS	USA (San Diego)
HUI2	VAS mapped to SG	Canadian (Hamilton parents), UK
HUI3	VAS mapped to SG	Canada (Hamilton) France
EQ-5D	TTO and VAS	UK, USA, Germany, Spain, Japan and many others
SF-6D	SG	UK, Japan, Hong Kong, Portugal, Brazil
AQoL1	TTO	Australia (Victoria)
15D	VAS	Finland



How do they compare?

Naïve view:

They value health states on the same scale (where zero is for dead and one for full health), so they should generate the same values for the same patients

Reality:

There are major differences in terms of their descriptive systems (coverage, range and sensitivity) and methods of valuation – and so are not on the same scale.



Comparison of 3 generic measures mean (SD) scores by visual impairment

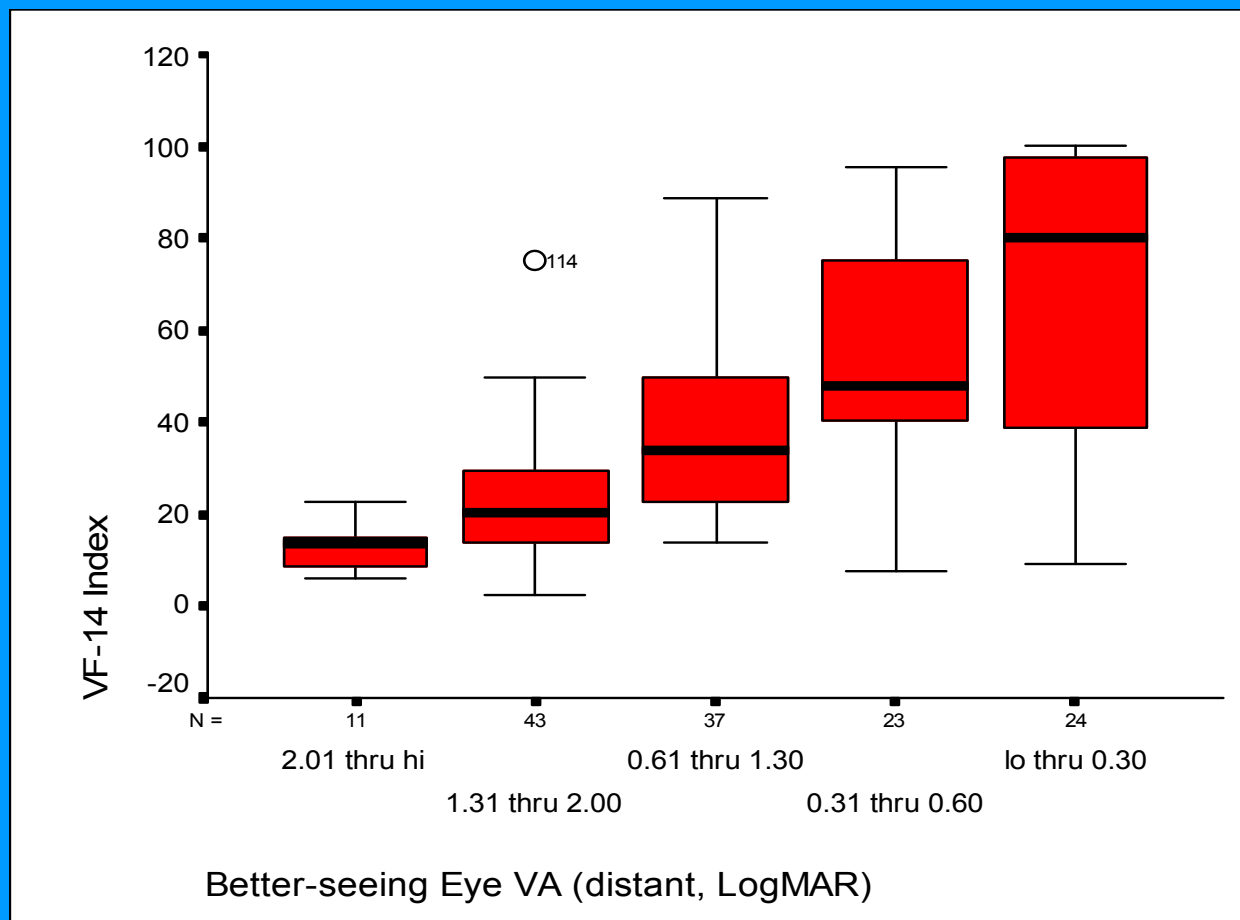
Contrast sensitivity (binocular, log units)	N	TTO	HUI3	SF-6D	EQ-5D
<0.30	67	0.58 (0.32)	0.25 (0.25)	0.65 (0.11)	0.70 (0.20)
0.30 thru 0.90	67	0.56 (0.32)	0.30 (0.26)	0.64 (0.14)	0.70 (0.24)
0.91 thru 1.3	48	0.70 (0.28)	0.42 (0.24)	0.68 (0.14)	0.78 (0.16)
>1.30	26	0.83 (0.25)	0.53 (0.31)	0.73 (0.16)	0.70 (0.28)
R-squared		0.09*#	0.14*#	0.05*#	0.03

* $p < 0.05$ between groups, # $p < 0.05$ linear trend

Source: Espallargues et al, 2006

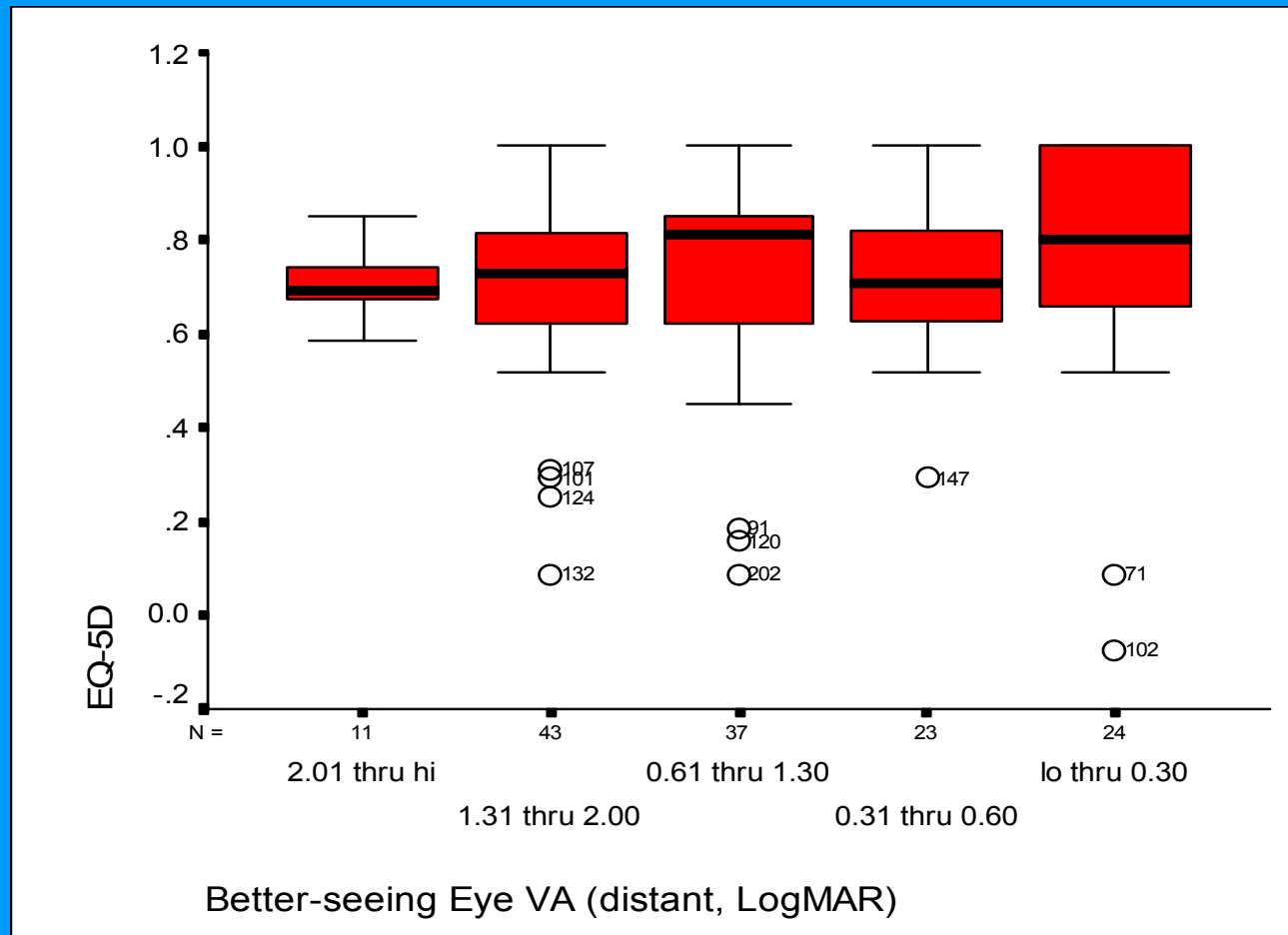


VF-14 by VA





EQ-5D by VA





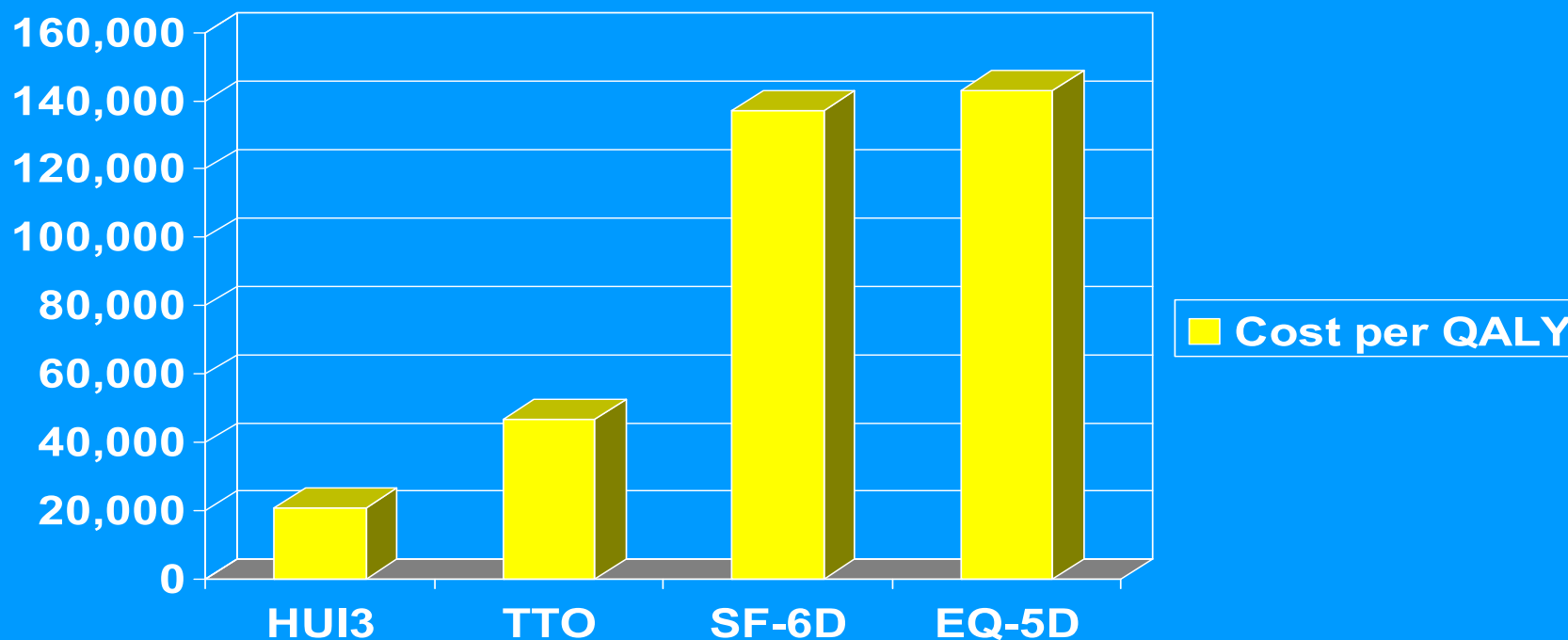
The
University
Of
Sheffield.



LAUREL AND HARDY



Cost effectiveness of drug for MD by preference-based measure





Causes and policy implications

Descriptive System

- Dimensions – e.g. vision not covered by EQ-5D
- Severity – range (e.g. floor effect in SF-6D)
- Sensitivity – number of levels (e.g. EQ-5D is very crude at upper end)

Valuation

- Time trade-off, standard gamble and visual analogue scales produce different values

→ Implications for policy

- Policy makers insist on one measure to achieve comparability across patient groups
- Problem: no single measure covers all patient groups (e.g. what about children?) or all medical conditions (e.g. visual impairment in AMD, hearing loss, OAB, leg ulcers etc.)



The
University
Of
Sheffield.

One solution: mapping





Mapping empirically onto a generic measure

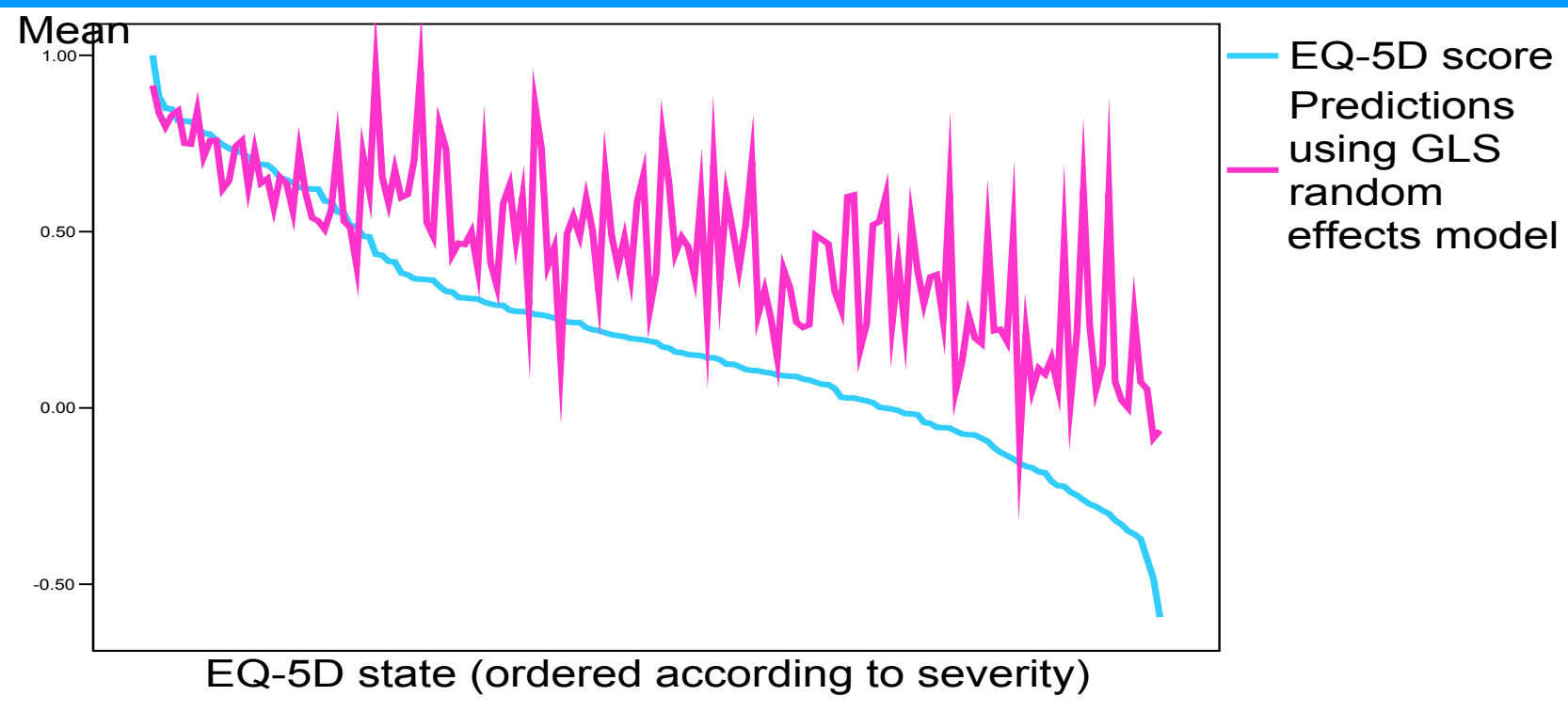
- Uses regression to 'map' between measures (i.e. estimate a relationship between them in order to use one (e.g. SF-36) to predict a value for the other (e.g. EQ-5D))
- Requires both measures to be used in the same patient sample

Advantages:

- It can be quick and in some circumstances it may be adequate
- How well does it do?



Using SF-36 to predict the EQ-5D scores





Mapping: where is it going?

Current practice:

- Review found variable performance across conditions
- Published studies usually do not explore the impact of size and pattern of errors (related to severity) on cost effectiveness
- Recent application to mapping function from SF-36 to EQ-5D found average MAE of incremental differences between trial arms of around 0.05. Impact on CE depends on context.

Implications:

- Always second best to directly using a generic in a trial
- Important dimensions may not appear in the generic measure – so may be better to value preferred measure in the first place



Condition specific measures: asthma

Feel concerned about having asthma

- [1]None of the time [2]A little or hardly any of the time [3]Some of the time
[4]Most of the time [5] All of the time

Feel short of breath as a result of asthma

- [1]None of the time [2]A little or hardly any of the time [3]Some of the time
[4]Most of the time [5] All of the time

Experience asthma as a result of air pollution

- [1]None of the time [2]A little or hardly any of the time [3]Some of the time
[4]Most of the time [5] All of the time

Asthma interferes with getting a good night's sleep

- [1]None of the time [2]A little or hardly any of the time [3]Some of the time
[4]Most of the time [5] All of the time

Overall, the activities I have done have been limited

- [1] Not at all [2] A little [3] Moderate or some
[4] Extremely or very [5] Totally

Valuation

- Asthma (AQL-5D) states valued using TTO (MVH protocol) by a representative sample of UK general population (n=308)
- Statistical modelling to estimate an algorithm for valuing all states defined by AQL-5D (and hence AQLQ) (Yang et al, 2007)

Similar studies have been undertaken with SF-36 (SF-6D), King Health Questionnaire, OABq, SQoL, ADQoL and others are on their way



Generic vs. condition specific measures

For Condition specific measures:

- more sensitive and relevant to condition

Against:

- excludes side-effects of treatment (unless built into measure)
- excludes co-morbidities and these may alter the impact of the main medical condition (i.e. preference dependence)
- Problem of achieving comparability across measures due to framing effects and dependence between preferences for included and excluded dimensions (i.e. the impact of role limitation may depend on a persons mental well-being)



The
University
Of
Sheffield.





Why use QALY?

- Combines different dimensions of health (e.g. mobility, pain, social functioning) into a single measure of effectiveness
 - Combines the two main benefits of health care: improved health related QoL with survival
 - Can be used to assess all types of interventions: those that impact on length of life, quality of life and both
 - Can be used to make comparisons across interventions in the same patient group and between patient groups
- can be used to inform resource allocation in health care



Problems with QALYs?

- Ignores equity concerns (e.g. QALYs to those in poor health may be more valuable to society) – can adjust QALYs to take such concerns into account
- Ignores non-health issues – can be included but limited by size of descriptive system
- Results are dependent on methods used methods – so need a reference case
- Existing measures may be too crude – so need research into better descriptive systems



Conclusions about QALYs

- A health care system without QALYs lacks a systematic basis for comparing the cost effectiveness of health care interventions
- They are contentious, but have become accepted in a number of countries e.g. UK, Australia, Netherlands, Sweden, Canada
- Further research is required into developing QALYs based on more sensitive and relevant descriptive systems and to ensuring comparability is retained