

Use in Surrogate Outcomes in HTA: Survey of UK reports

Rod Taylor

Associate Professor in Health Services Research

Peninsula College of Medicine & Dentistry
Universities of Exeter & Plymouth, UK

HTAi 2009 Annual Conference – Parallel Panel Session
HTAi Policy Forum WGSO



PENINSULA
— MEDICAL SCHOOL —
UNIVERSITIES OF EXETER & PLYMOUTH

Disclosure of Interests

- Peninsula Technology Assessment Group (PenTAG) is one of the independent review groups for the NICE Technology Appraisal process
- Member of NICE Technology Appraisal & Interventional Procedures Advisory Committees
- Consultant for the healthcare industry

UK Guidance on Surrogates

- **Evidence submitted to NICE**

4.4.3 The written submissions provide a unique contribution outlining the professional view of the place of the technology in current clinical practice. This includes evidence that relates to some or all of the following:

.....

- the identification of appropriate outcome measures and the appropriate use of surrogate outcome measures.

Aim

- To explore the use of surrogate outcomes in the health technology assessment (HTA) and by doing so provide a basis for guidance for their future use, validation and reporting.
-focus on role of surrogate outcomes in cost-effectiveness models (CEMs) within UK HTA Programme reports.

Objectives

1. Assess prevalence of use of surrogate outcomes in the CEMs in UK HTA programme reports
2. Review current practice around the use of surrogate outcomes in CEMs in UK HTA programme reports
3. Provide recommendations on the use of surrogate outcomes in the CEM within future HTA reports.

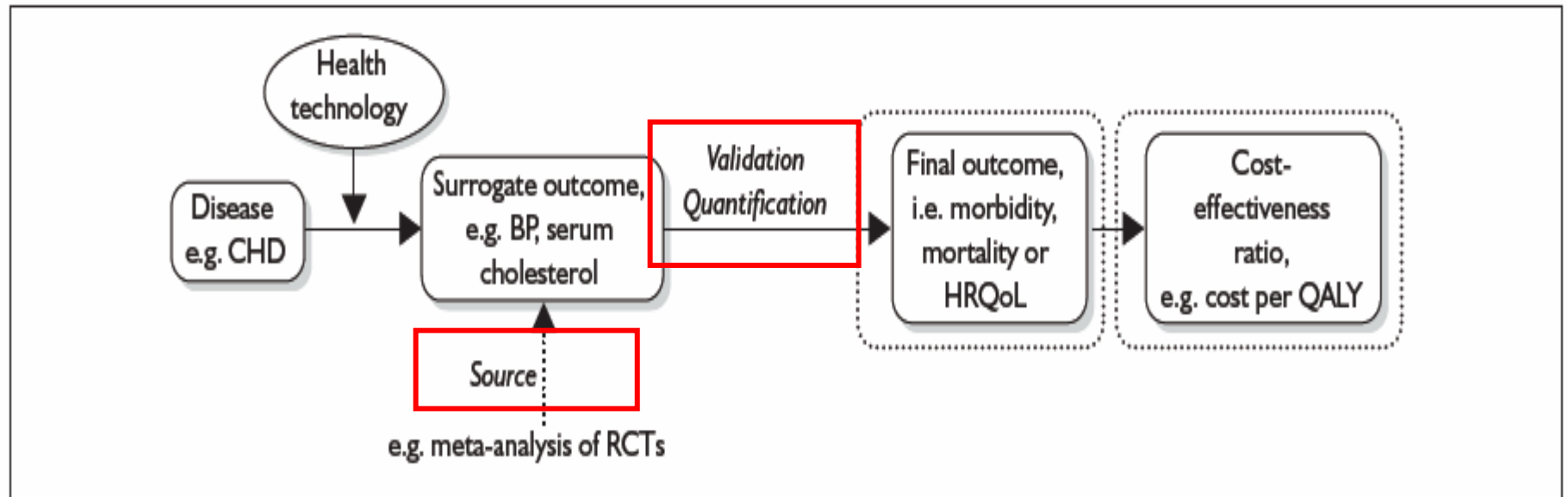


FIGURE 2 Schematic representation of the use of a surrogate outcome in a health technology assessment cost-effectiveness model. 'Source' refers to the source of the surrogate outcome data (usually a systematic review/meta-analysis of clinical effectiveness literature in an HTA); 'validation' refers to the evidence supporting the relationship between the surrogate outcome and the final outcome; and 'quantification' refers to how this relationship has been quantified. The two dotted boxes show that quantification of the surrogate outcome to final outcome may take place either within or outside the cost-effectiveness model per se. BP, blood pressure; CHD, coronary heart disease; HRQoL, health-related quality of life; QALY, quality-adjusted life-year; RCT, randomised controlled trial.

Definitions and Validation Criteria for Biomarkers and Surrogate Endpoints: Development and Testing of a Quantitative Hierarchical Levels of Evidence Schema

MARISSA N. LASSERE, KENT R. JOHNSON, MAARTEN BOERS, PETER TUGWELL, PETER BROOKS, LEE SIMON, VIBEKE STRAND, PHILIP G. CONAGHAN, MIKKEL ØSTERGAARD, WALTER P. MAKSYMOWYCH, ROBERT LANDEWÉ, BARRY BRESNIHAN, PAUL-PETER TAK, RICHARD WAKEFIELD, PHILIP MEASE, CLIFTON O. BINGHAM III, MICHAEL HUGHES, DOUG ALTMAN, MARC BUYSE, SALLY GALBRAITH, and GEORGE WELLS

ABSTRACT. *Objective.* There are clear advantages to using biomarkers and surrogate endpoints, but concerns about clinical and statistical validity and systematic methods to evaluate these aspects hinder their efficient application. Our objective was to review the literature on biomarkers and surrogates to develop a hierarchical schema that systematically evaluates and ranks the surrogacy status of biomarkers and surrogates; and to obtain feedback from stakeholders.

Methods. After a systematic search of Medline and Embase on biomarkers, surrogate (outcomes, endpoints, markers, indicators), intermediate endpoints, and leading indicators, a quantitative surrogate validation schema was developed and subsequently evaluated at a stakeholder workshop.

Results. The search identified several classification schema and definitions. Components of these were incorporated into a new quantitative surrogate validation level of evidence schema that evaluates biomarkers along 4 domains: Target, Study Design, Statistical Strength, and Penalties. Scores derived from 3 domains — the Target that the marker is being substituted for, the Design of the (best) evidence, and the Statistical strength — are additive. Penalties are then applied if there is serious counterevidence. A total score (0 to 15) determines the level of evidence, with Level 1 the strongest and Level 5 the weakest. It was proposed that the term "surrogate" be restricted to markers attaining Levels 1 or 2 only. Most stakeholders agreed that this operationalization of the National Institutes of Health definitions of biomarker, surrogate endpoint, and clinical endpoint was useful.

Conclusion. Further development and application of this schema provides incentives and guidance for effective biomarker and surrogate endpoint research, and more efficient drug discovery, development, and approval. (*J Rheumatol* 2007;34:607-15)

Key Indexing Terms:

SURROGATE
LEVELS OF EVIDENCE

TRIAL ENDPOINT

BIOMARKER
PREDICTIVE FACTORS

From the Department of Rheumatology, St. George Hospital, University of New South Wales, Sydney, Australia.

M.N. Lassere, MB, BS, Grad Dip Epi, PhD, Associate Professor in Medicine, Department of Rheumatology, St. George Hospital, University of NSW, Sydney, Australia; K.R. Johnson, MD, Senior Lecturer in Medicine, University of Newcastle, Newcastle, University of New South

Wales; P.P. Tak, MD, PhD, Professor, Division of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; R. Wakefield, BM, Senior Lecturer in Rheumatology, Academic Department of

OMERACT Scoring

Domain	Definition	Scoring
Target	The final outcome* that the surrogate substitutes for	0 to 5
Study design	The level of evidence for the relationship between the surrogate and final outcome	0 to 5
Statistical strength	The strength of the association & its statistical significance between the surrogate and final outcome	0 to 5
Penalties	Lack of, opposing or inconsistent evidence from biology, clinical epidemiology or therapeutic trials	-1 to -3
		-3 to 15 [cut off ≥ 10]

Lassere et al. *J Rheumatol* 2007;34:607–15

OMERACT criteria – scoring

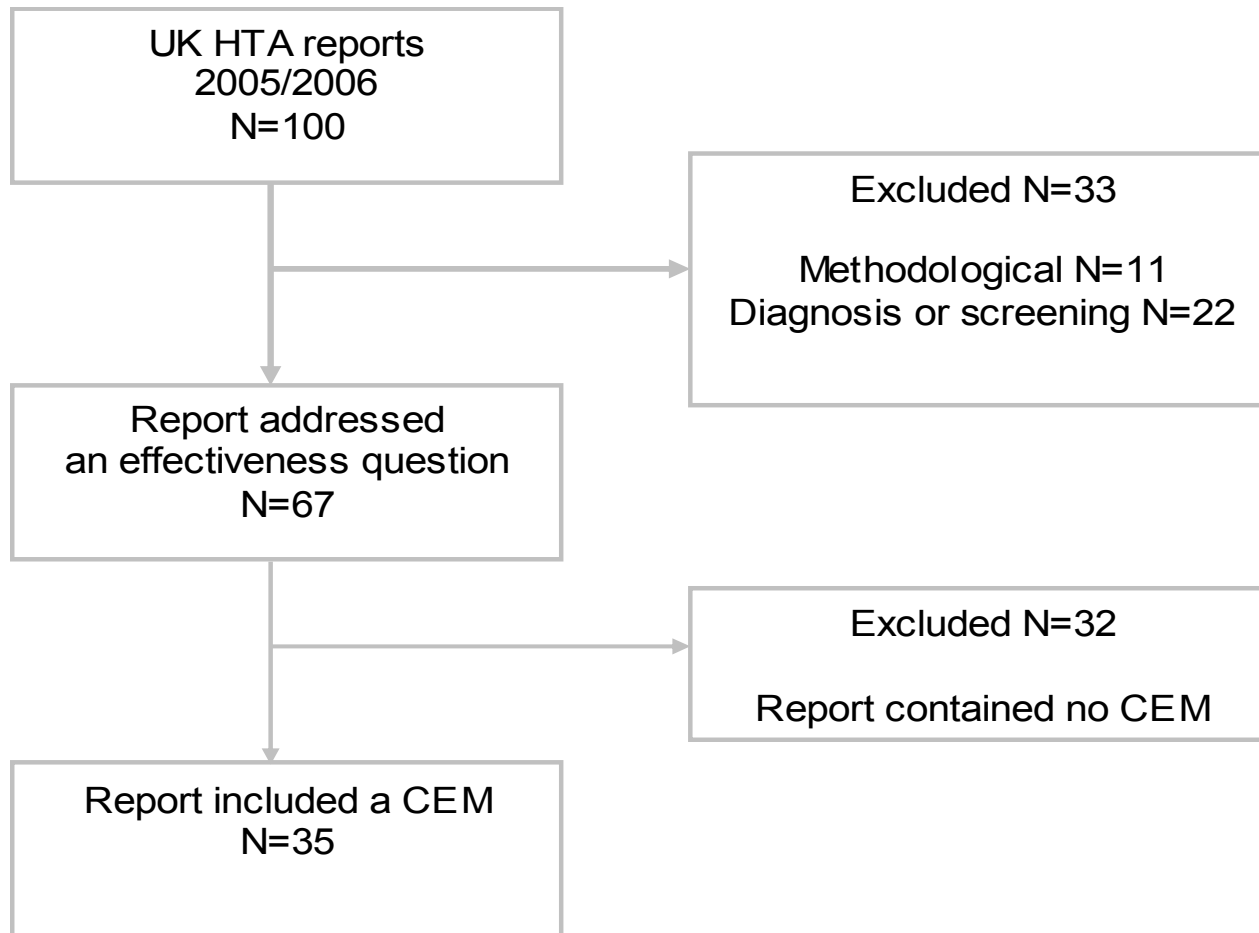
Circle appropriate rank for domains A–D and provide total score.

Domain	Rank	Criteria
A. Target (for all studies ranked in domain B)	0	All targets studied are disease centred and reversible
	1	At least one target studied that is disease centred is irreversible
	2	At least one patient-centred target that is reversible
	3	At least one patient-centred target of irreversible minor organ morbidity or minor irreversible clinical burden of disease
	4	At least one patient-centred target of irreversible major organ morbidity or major irreversible clinical burden of disease
	5	Death

Methods

- Sampling frame: UK HTA Programme monograph reports published in 2005 and 2006
 - Included reports with (1) economic model, (2) economic model based on a surrogate outcome & (3) addressing treatment/therapy question
 - Excluded reports addressing diagnostic/aetiological/prognostic/methodological question

Results – Included studies



Elston & Taylor (2009) *Int J Technol Assess Healthcare* 2009;25:6-13.

Results – Surrogate CEMs

	Disease/ technology	Surrogate endpoint	Final outcome
Woodroffe (2005)	Adults undergoing renal transplantation	Biopsy confirmed acute rejection (BPAR)	Graft survival
Loveman (2005)	Alzheimer's disease/new drugs	Cognitive function score (ADAS-COG)	Need for full time residential care
Shephard (2006)	Chronic hepatitis/new drugs	Biochemical (ALT) and viral (HBV) outcomes	Cirrhosis, liver cancer, liver transplant
Yao (2006)	Children undergoing renal Transplantation	Biopsy confirmed acute rejection (BPAR)	Graft survival

Results – Validation & Quantification

	Surrogate Validation	Surrogate quantification
Woodroffe (2005)	SR of observational studies	Regression-based model
Loveman (2006)	Single observational study	Regression-based model
Shephard (2006)	Natural history model	Transition probabilities
Yao (2006)	SR review of observational studies Comparison of change in surrogate outcome in one RCT	Regression-based model

Surrogate Quantification

Woodroffe et al (2005)

“A systematic review by the authors of this report found evidence that biopsy confirmed acute rejection is predictive of future 5-year or longer graft survival in adults....pooled hazard ratio was 1.96 [95% CI: ...].....”

TABLE 57 Sensitivity analysis – varying hazard ratio

	ICER (mean) (£/QALY)
CAS vs TAS	
HR 1.41	145,540
HR 1.96	58,801
CAS vs CMS	
HR 1.41	194,559
HR 1.96	76,958
CAS vs BCAS	
HR 1.41	Dominant
HR 1.96	Dominant
CAS vs DCAS	
HR 1.41	Dominant
HR 1.96	Dominant
TAS vs BTAS	
HR 1.41	Dominant
HR 1.96	Dominant

OMERACT Scoring

	Woodroffe (2005)	Shephard (2005)	Loveman (2006)	Yao (2006)
Target	4	4	4	4
Study design	2	0	2	2
Statistical design	3	0	3	3
Penalties	0	0	0	0
Total score	9	4	9	9

Limitations

- Small sample size...generalisability?
- Operational definition of surrogate
- Documentary analysis

Conclusions

- 4/35 (11%) HTA reports with CEM based on surrogate outcome
- All reports sourced surrogate outcome from systematic review but only two undertook SR of relationship between surrogate and final outcome
- Only one report provided level 1 validation evidence (trial-based) of relationship between surrogate and final outcome

Conclusions cont.

- None of reports achieved OMERACT score to indicate acceptable evidence of a surrogate outcome
- All included surrogate reports undertaken on behalf of NICE
- Number of CEM reports (n=7) used 'intermediate outcome' (e.g. HRSD) and extrapolated to utility

Recommendations

- Where possible, HTAs should be based on final patient-related outcome
- If need to use surrogate outcome
 - Undertake SR of evidence for relationship of surrogate/final outcome
 - Evidence should be categorised according to validation hierarchy (level 1, 2 or 3)
 - (Strong) consideration to only basing CEM analysis if at least level 2 evidence

Recommendations cont

- Where CEM analysis based on surrogate outcome is undertaken
 - Transparent explanation of relationship between surrogate and final outcome
 - Explicit exploration of uncertainty of relationship through sensitivity analysis
 - Specific research recommendations considering relationship
 - Include term ‘surrogate outcome’ in report executive summary

Acknowledgements

- Julian Elston
- Ken Stein
- Colin Green
- Paul Tappenden, ScHARR
- NIHR HTA Programme TAR contract funding