International Journal of Technology Assessment in Health Care, **25:3** (2009), 315–322.

doi:10.1017/S0266462309990213

Surrogate outcomes in health technology assessment: An international comparison

Marcial Velasco Garrido
Technische Universität Berlin

Sandra Mangiapane
Central Research Institute of Ambulatory Health Care in Germany

Objectives: Our aim was to review the recommendations given by health technology assessment (HTA) institutions in their methodological guidelines concerning the use of surrogate outcomes in their assessments. In a second step, we aimed at quantifying the role surrogate parameters take in assessment reports.
Methods: We analyzed methodological papers and guidelines from HTA agencies with International Network of Agencies for Health Technology Assessment membership as well as from institutions related to pharmaceutical regulation (i.e., reimbursement, pricing). We analyzed the use of surrogate outcomes in a sample of HTA reports randomly drawn from the HTA database. We checked methods, results (including evidence tables), and conclusions sections and extracted the outcomes reported. We report descriptive statistics on the presence of surrogate outcomes in the reports.
Results: We identified thirty-four methodological guidelines, twenty of them addressing the issue of outcome parameter choice and the problematic of surrogate outcomes. Overall HTA agencies call on caution regarding the reliance on surrogate outcomes. None of the agencies has provided a list or catalog of acceptable and validated surrogate outcomes. We extracted the outcome parameter of 140 HTA reports. Only around half of the reports determined the outcomes for the assessment prospectively. Surrogate outcomes had been used in 62 percent of the reports. However, only 3.6 percent were based upon surrogate outcomes exclusively. All of them assessed diagnostic or screening technologies and the surrogate outcomes were predominantly test characteristics.
Conclusions: HTA institutions seem to agree on a cautious approach to the use of surrogate outcomes in technology assessment. Thorough assessment of health technologies should not rely exclusively on surrogate outcomes.

Keywords: Surrogate outcomes, Biological Markers, Intermediate outcomes, Outcome assessment

Surrogate end points represent, in the best case, preliminary steps in the causal chain leading to relevant clinical and patient outcomes such as mortality or morbidity. They are measured in lieu of the actual outcome of interest and are typically biochemical markers, physiological parameters, or subclinical end points that are not generally perceived directly by patients, but are nevertheless correlated with clinically relevant end points (e.g., high blood pressure is associated with a higher risk of stroke, high LDL cholesterol is a risk factor for myocardial infarction, the CD4 cell count is associated with AIDS-related morbidity and mortality). Surrogate outcomes are used not only in trials of pharmaceuticals, but also in studies of other clinical technologies. In addition, parameters with an intermediary character

This study originated in a project commissioned and funded by the German Agency for Health Technology Assessment (DAHTA@DIMDI) (HTA Report Grant 53–06). The Agency is part of the German Institute for Medical Documentation and Information, which is a federal authority subordinated to the Federal Ministry of Health. The full report of the project will be published in German during 2009 at http://www.dimdi.de/static/en/hta/index.htm.

are also used in the field of community and public health interventions.

Their use in assessing the benefits of health technologies, however, is problematic. In the past, reliance on surrogate outcomes has led to false conclusions concerning the effects of technologies on the relevant health outcomes (4;11). In many situations, relying on the strong correlation observed between surrogate and relevant end points to spread an intervention has had fatal consequences (i.e., positive effects on the surrogate, but increased mortality with the intervention in question). These issues have been already raised 30 years ago, as the World Health Organization study on the prevention of myocardial infarction with clofibrate rebuted the expectations put in this strategy—which were founded in its cholesterol lowering effect—by showing that mortality was higher in the clofibrate group than in the placebo group (7).

A classic example of the potential for fatal consequences when relying on surrogates is the case of class I antiarrhythmic drugs: these drugs increase mortality although they suppress ventricular arrhythmias, which on their own are associated with higher mortality after myocardial infarction (4). Some drugs have been removed from the market after increased mortality or morbidity was observed in association with their use, contrary to the expectations that had been raised by the observation of positive effects on a surrogate end point (4;29). However, in other occasions reliance on surrogates has led to the withholding of effective therapies. For example, beta blockers—due to their bradycardic effect—were considered for many years to be contraindicated in patients with heart failure, because—following pathophysiological reasoning—a reduction in heart rate was thought to have deleterious effects in these patients. These agents, however, turned to have a beneficial effect by reducing all-cause mortality in patients with chronic heart failure (15;22).

Although these—and other—previous experiences have shown the risks of basing technology assessment on surrogate outcomes, claims for the potential benefits—for example, faster access to effective innovations—of relying more on these type of parameters are recurrent. For example, the "accelerated approval procedure" of the US Food and Drug Administration introduced in 1993 the possibility to base approval assessments of drugs for life-threatening conditions currently lacking of treatment options on surrogate end points (12). In a recent international symposium organized by the German Federal Ministry of Health to debate the assessment of cost-effectiveness—with gathered among others participants from academia and industry—the issue of the potential role of surrogate outcome parameters in the assessment of health technologies was raised (1). More recently, a Working Group on Surrogate Outcomes has been established under the auspices of the international society for Health Technology Assessment (HTAi) with the aim of encouraging the debate about the issue of surrogate outcomes use in HTA (16).

In light of these debates, it seems relevant to get an overview of the actual handling of the issue of surrogate outcomes in the international HTA community.

The aim of our study was thus to assess the role currently assigned to surrogate outcomes in the field of HTA by reviewing institutional methodological guidance and by estimating the actual use of surrogate end points in international HTA reports.

METHODS

We followed two different methodological approaches for addressing the above-mentioned aim: a document review and the analysis of a random sample of HTA reports.

Document Review

We searched for documents describing the general methods recommended and followed by HTA institutions in the conduction of their assessments (i.e., in the elaboration of their HTA reports). We included two types of institutions for this review, both of them publicly funded:

- HTA agencies which are members of the International Network of Agencies for HTA (International Network of Agencies for Health Technology Assessment, INAHTA) and
- Institutions undertaking assessments in connection with pharmaceuticals' regulation (i.e., reimbursement and pricing), without INAHTA membership.

Documents were included if they had been made available to the public and were written in English, French, German, Portuguese, or Spanish. General methodological guidance published by these institutions was retrieved through the institutional Web sites. To identify relevant documents we scanned the publication lists and performed an additional search using the institutional Web sites' search engine targeting at agencies' methodological work on the issue of surrogate outcomes.

We analyzed the relevant documents and extracted the passages dealing with the selection of outcome parameters in general and with the use of surrogate outcomes in particular. The statements were compared and summarized in a narrative way.

Random Sample of HTA Reports

For the purpose of this part of the project, we considered the HTA reports gathered in the HTA database (www.crd.york.ac.uk/crdweb/Home.aspx?DB = HTA). To be eligible for our survey, reports had to be finished—i.e., documents described as "(ongoing) project" were excluded—and written in English, French, German, Portuguese, or Spanish.

We drew a representative random sample stratified by country of origin and by publication year (before 2000, between 2000 and 2004, after 2004). The sample size was set

Table 1. Definitions of Surrogate Outcomes Proposed in HTA Methodological Guidelines

Agency (country)	Definition
Medical Devices Advisory Committee (Australia) (21)	The term 'surrogate outcome' has a variety of definitions, and other terms are often used in its place (e.g., intermediate outcome or biological marker). The common features of the definitions are: • that the surrogate outcome is commonly a physiological variable (e.g., serum cholesterol concentration, blood pressure); • there is a statistical association between the surrogate outcome and the clinical outcome of interest (e.g., bone mineral density and fracture, CD4 cell concentrations and progression of HIV); or • there is a biological and pathophysiological basis for believing that the surrogate outcome is a major determinant of the clinical outcome in the disease being studied (e.g., glycosylated hemoglobin measurements and diabetic complications). A surrogate outcome should possess all of these features, but experience suggests that few do.
Canadian Agency for Drugs and Technologies in Health (Canada) (5)	A surrogate outcome is "a laboratory measurement or a physical sign used as a substitute for a clinically meaningful end point that measures directly how a patient feels, functions, or survives." [Bucher et al. (1999) (4)]
Pharmaceutical Management Agency (New Zealand) (28)	'Surrogate/intermediate' outcomes are essentially biological markers. Commonly a physiological variable (e.g., serum LDL-cholesterol concentration, blood pressure), a surrogate/intermediate outcome has a statistical association with clinical outcome of interest (e.g., bone mineral density with fracture, CD4 cell concentrations with progression of HIV). There will also be a biological and pathophysiological basis for believing that the surrogate/intermediate outcome is a major determinant of the clinical outcome in the disease being studied (e.g., glycosylated hemoglobin (HbA1c) and diabetes complications). A surrogate/intermediate outcome should possess all of the above features, but few do.

at 5 percent of the eligible reports for convenience. Random numbers were generated with statistical software (SAS®).

Documents were excluded from the final sample of HTA reports if they had to be purchased for a fee or if they did not report a technology assessment (e.g., guidelines, methods reports).

To assess what kind of outcomes had been used in the assessment, we analyzed research questions, methods and results sections as well as—if available—the evidence tables displaying information extracted from studies included in the HTA reports. We classified the types of outcome parameter used in the HTA reports according to the following definitions:

- *Clinical Relevant End Point*: "A characteristic or variable that reflects how a patient feels, functions, or survives" (3). It includes end points measurable or verifiable by thirds such as death or morbidity (i.e., hard end points) as well as subjective end points such as symptoms, quality of life, performance, etc. (i.e., so-called patient reported outcomes).
- *Surrogate End Point*: Measurable parameters which are based on epidemiologic, pathophysiologic, therapeutic, or other scientific evidence and which are expected to predict a clinical relevant end point (3). A surrogate end point represents an intermediary stage in the causal chain leading to the clinical relevant outcome. According to the framework proposed by Fryback und Thornbury (13) for the assessment of diagnostic technologies, we consider test characteristics (sensitivity, specificity, etc.) to be also surrogate parameters.

In addition, we evaluated whether a prospective definition of the outcome parameter to be considered in the assessment had been reported in the documents. We also analyzed whether only effectiveness outcomes had been considered or if the report also indicated safety assessment as an additional end point.

We provide descriptive statistics of the frequencies of the different outcomes.

RESULTS

Recommendations from Assessment Institutions

A total of fifty-six institutions were scanned (see Supplementary Table 1, which can be viewed online at www.journals.cambridge.org/thc, for the complete list), including both INAHTA members and nonmembers.

At the time we performed our document search (June 2007), INAHTA counted 45 member institutions. We identified a total of twenty-four relevant documents from fifteen agencies. The documents had been published between 1995 and 2007. One document was excluded, because it was only available in Norwegian. The German Institute for Quality and Efficiency in Healthcare published an updated version of its general methods paper in 2008—during the term of our project—which we included instead of the one identified trough our search and dating from 2006.

We scanned eleven further nonmember institutions of INAHTA. This search line yielded twelve further relevant methodological guidelines published by eight institutions between 1999 and 2006. One of the documents was, however, excluded because it was only available in Finnish. Figure 1 depicts the document selection process.

Overall, the problematic of surrogate end points and outcome selection was addressed in twenty of the analyzed

https://doi.org/10.1017/S0266462309990213 Published online by Cambridge University Press

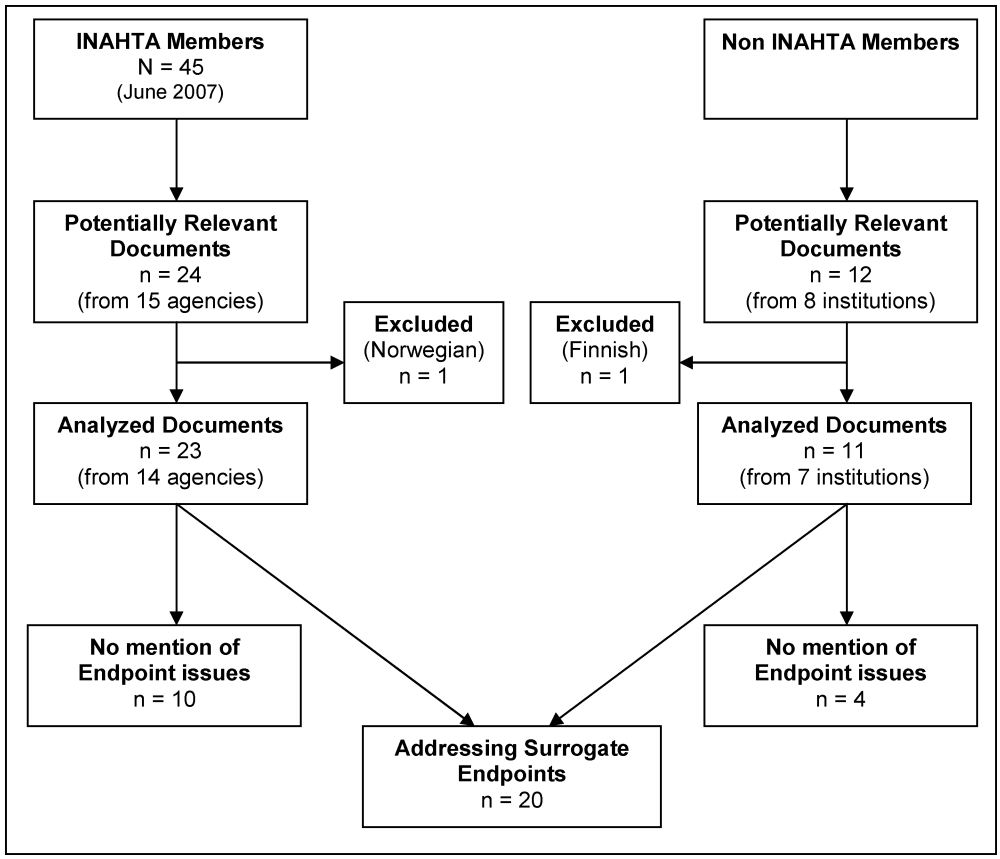

Figure 1. Pool of methodological guidelines from HTA institutions.

methodological documents at least to some extent, although the depth and breadth with which the issue was addressed varied considerably across documents. The complete pool of documents and references included in this part of the study as well as the extracted passages related to the issue of surrogates is available in Supplementary Table 2, which can be viewed online at www.journals.cambridge.org/thc.

The documents analyzed witness a general consensus among HTA institutions regarding the type of outcome parameter which are to be taken into account when assessing health technologies, namely those considered to be directly relevant for patients. Mortality, morbidity, symptomatology, quality of life, as well as adverse and undesirable effects are consistently mentioned as the outcomes needed to perform a reliable assessment of a health technology. In many of the documents, these were termed as being "definitive" or "final" outcomes, enhancing their conceptual opposition to the intermediary, indirect and short-termed character associated with surrogate outcomes. There seems to be a broad consensus that information on these kinds of outcomes (i.e., final) is the one required for an appropriate weighting of the good and harm associated with the application of an intervention.

Only a few institutions provided an elaborated definition of the concept of "surrogate" outcome or end point—these being very similar and partly literally the same (see Table 1). Common to all these definitions is that surrogates outcomes are physiologic parameters, biomarkers, or physical signs measured instead of a relevant end point.

Whereas all the agencies underline that clinical or patient relevant outcomes are to be given priority in HTA, many of them declare in their methodological guidance that a surrogate outcome parameter could be considered acceptable in exceptional situations, provided that the validity of the surrogate outcome has been proven (2;5;6;14;17;21;24;26–28). In the view of some agencies, in some situations, evidence of effects on surrogate outcomes could be combined with other evidence in pharmacoeconomic modeling to simulate final outcomes (5;6;14;27;28). The validity of a surrogate is considered to be given if a set of conditions has been fulfilled. These conditions can be summarized as follows (5;17;21;24;28): There is a statistical association between the surrogate and the relevant end point; The association between the surrogate and the relevant end point is plausible from a biological and pathophysiological point of view; The association is strong and consistent across studies; and There is evidence from randomized controlled trials that improvements in the surrogate has led to improvements in the final outcome.

https://doi.org/10.1017/S0266462309990213 Published online by Cambridge University Press

Table 2. Results of HTA Reports Survey

Outcome type[a]	No. of reports (%)[b]				
	All reports ($n = 140$)	Reports with nondiagnostic technologies ($n = 121$)	Reports with diagnostic/ screening technologies ($n = 19$)	Reports with prospective definition of assessment parameters ($n = 67$)	Reports with *ad hoc* extraction of assessment parameters ($n = 73$)
Effectiveness assessment					
Clinical relevant outcomes	134 (96%)	120 (99%)	14 (74%)	65 (97%)	69 (94%)
Hard end points	131 (93%)	118 (97%)	13 (68%)	62 (92%)	69 (94%)
Patient reported outcomes	78 (56%)	72 (60%)	6 (32%)	41 (61%)	37 (51%)
Surrogate outcomes	87 (62%)	70 (58%)	17 (89%)	42 (63%)	45 (62%)
used only surrogate end points	6 (4%)	1 (1%)	5 (26%)	2 (3%)	4 (5%)
Safety assessment	122 (87%)	111 (92%)	11 (58%)	56 (84%)	66 (90%)
Full Assessment used only surrogate outcomes	*5 (4%)*	***0 (0%)***	***5 (26%)***	***2 (3%)***	***3 (4%)***

[a]Definitions according to methods section.
[b]Multiple mentions possible,% rounded.
HTA, health technology assessment.

In addition, one agency—the German Institute for Quality and Efficiency in Health Care (IQWiG)—describes situations in which a surrogate can not be considered to be valid, namely when an intervention: shows an effect on the surrogate but not on the patient relevant end point, shows an effect on the patient relevant end point but not on the surrogate, and shows inconsistent effects on surrogate and patient-relevant end points (17).

Another institution—the Australian Pharmaceutical Benefits Committee (PBAC), which assesses evidence submitted by manufacturers, requires the submission of systematic review(s) showing the biological and epidemiological relationship between surrogate and final outcome, systematic review(s) of RCTs of other interventions showing that a positive effect on the proposed surrogate leads to a positive effect on the final outcome as well as a reasoning on why the relationship between surrogate and final outcome observed with other interventions would apply to the proposed intervention (2).

None of the documents includes a list or catalog of surrogate parameters considered to be generally acceptable or well established for different conditions. However, the PBAC guidelines suggest that left ventricular ejection fraction and viral load can be considered to be established for survival after myocardial infarction and cure for viral hepatitis, respectively (2).

Use of Surrogate Outcomes in HTA Reports

A total of 118 reports had to be excluded from our original 5 percent sampling for different reasons. Main reasons for exclusion were that the report was not available or because it was addressing a methodological issue and not a health technology (see Figure 2). Thus, we analyzed a final effective sample of 151 reports, which represents 2.8 percent of the eligible HTA reports. Although the reports covered different types of technologies, the majority dealt with the assessment of medical and surgical interventions (52 percent), followed by drugs (32.5 percent).

Overall, an explicit prospective description of the outcome parameters upon which the assessment would be based was present in less than half of the HTAs (i.e., no mention of relevant outcomes in the research or methods sections was found). Thus, in 55 percent of the assessments, we extracted the outcome parameters used from the results section or from evidence tables. Eleven of the reports did not contain any information on the outcomes considered in the assessment of the technology, and were subsequently not further taken into account. These HTAs were both lacking a prospective definition of the outcome parameters upon which the assessment would have been based and in addition reported that no evidence was found. Because no studies were included in the assessment and subsequently no results were reported (i.e., no evidence tables were shown in the report), we could not obtain any information on the assessment outcomes.

Surrogate outcomes had been considered and/or extracted and reported in 62 percent (87 of 140) of the reports. As shown in Table 2, the role of clinically relevant outcomes was, however, much more prominent. Almost all reports assessed effectiveness upon patient relevant parameters, either hard outcomes (e.g., mortality, incidence of morbid events) or outcomes reported by patients. In addition, the majority of HTAs reported to also address the safety of the technology taking into account any kind of potential adverse effects. In general, surrogate outcome parameters were more frequently used in the assessment of diagnostic and screening technologies than in other types of technologies.

Only one of the documents addressing therapeutical interventions reported to assess effectiveness of the technology

https://doi.org/10.1017/S0266462309990213 Published online by Cambridge University Press

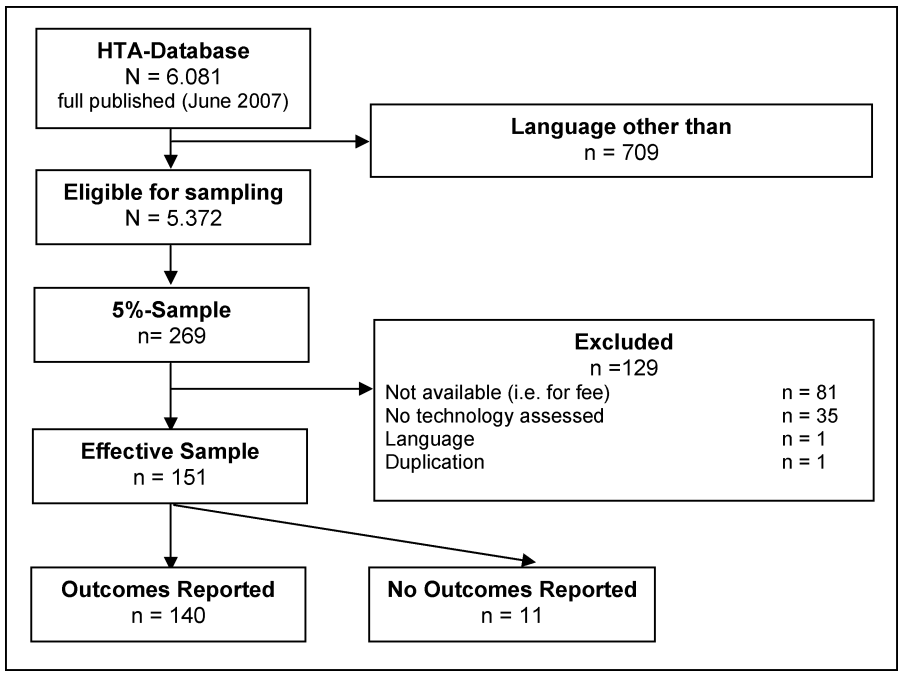

Figure 2. HTA reports sampling from the HTA database (available at www.crd.york.ac.uk/crdweb/Home.aspx?DB = HTA).

in question based on evidence on a surrogate parameter, although the full assessment also considered safety and thus it was not solely based in surrogate outcomes. This was a report from an appraisal of the National Institute for Health and Clinical Excellence (NICE) on antivirals for the treatment of hepatitis C that assessed effectiveness upon viral load. However, the document acknowledged the need for further research on relevant outcomes such as survival and health-related quality of life (25). The appraisal was based on a more comprehensive report in whose methods section other outcomes such as morality, morbidity, or health-related quality of life had been defined as the relevant parameters for the assessment of this technology (31).

Overall, only five reports used exclusively surrogate parameters for the full assessment of the target technology. None of these reports assessed a therapeutical intervention. One of the reports assessed the effectiveness of strategies to increase women participation in breast cancer screening. This report analyzed exclusively screening uptake rates (30). The other four HTAs assessed diagnostic technologies and extracted and/or reported to have considered only detection rates and/or test characteristics such as specificity or sensitivity (9;19;23;32). Three of them addressed the potential for replacing a diagnostic method by another one with different degrees of similarity. One of them aimed at assessing the replacement of the current laboratory method for neonatal screening by tandem-mass spectrometry (19). Another report assessed the potential for replacement of magnetic resonance imaging (MRI) in the stroke diagnostic chain by ultrafast MRI (23). Finally, another HTA addressed optical coherence tomography as potential substitute for fluorescein angiography in the diagnostic work-up of age-related macular degeneration (32).

DISCUSSION

Our results indicate that HTA institutions are cautious regarding the use of surrogate outcome parameters in their assessments. According to their methodological guidelines, non-surrogate parameters are considered the first choice in HTA. These parameters should allow a more reliable and comprehensive assessment of the tandem effectiveness-safety. At the same time, there seems to be a common place among these agencies that, at least theoretically, surrogate outcomes could be used as the main source of evidence for the assessment in well-grounded situations. Our results show that in practice these situations can be considered exceptional. Only 1 percent of reports assessing a therapeutic technology based their assessment of effectiveness exclusively on surrogate parameters and none of them based the full assessment on such parameters. This is not surprising because, according to the agencies' method papers, surrogate outcomes would only be accepted when there is ample evidence of its validity to reflect relevant effects of the technology. The evidence required to accept the validity of surrogate outcomes includes results from randomized controlled trials and from systematic reviews which show a consistent relationship between the surrogate and the final outcome which holds for the technology in question (20). Current works regarding the use of surrogate parameters in HTA also underline the preference

https://doi.org/10.1017/S0266462309990213 Published online by Cambridge University Press

for patient-relevant outcomes and the need to thoroughly assess the evidence and the uncertainties regarding the relationship between surrogate and final outcome (10;33).

The use of surrogate outcomes was much more common in the field of assessment of diagnostic technologies. Eighty-nine percent of HTAs on diagnostic technologies reported surrogates as compared to 58 percent of the reports assessing therapeutical technologies. Around 25 percent of reports on diagnostic technologies based the assessment solely on these kind of outcomes—for example, on tests characteristics. The reason for this discrepancy between diagnostic and therapeutic technologies is probably the fact, that restricting the assessment to tests characteristics is more often considered acceptable. Often the replacement of a diagnostic technology by another one has no other consequences for the patient than the ones derived from different accuracy. In the case of a well established diagnostic workup and therapeutic management chain, the assessment of a diagnostic technology which aims at replacing an existing one can be limited to the accuracy parameters if the new technology raises no safety concerns. That can be the case when two different laboratory procedures for analyzing a blood sample are compared or when different image processing procedures for the same imaging technique are under comparison.

To our knowledge this is the first comprehensive review on how HTA agencies and related institutions are handling the issue of surrogate outcomes in their assessments. Recently, a working group established by the Australian Pharmaceutical Benefits Advisory Committee has published a framework for evaluating the use of surrogate outcomes in submissions for coverage of pharmaceuticals (33). In addition to a review of the methodological scientific literature, the Australian report provided a comparison of the guidance on the surrogate issue of only two institutions (the English NICE and the Canadian Agency for Drugs and Technologies in Health). Our review of agencies' guidance—although it excluded two documents because of language reasons—is broader, thus allowing for a more complete international comparison (see Supplementary Table 2).

In the winter issue of this journal, a survey on the use of surrogate outcomes in HTA reports appeared, which, however, only included reports published within the United Kingdom HTA Program and was limited to cost-effectiveness models (10). Our survey was planned to be representative of the universe of HTA reports contained in the HTA database. Because of time and financial constraints in the project, we had to exclude reports that were not available through the Internet or that required to be purchased from the publishing institution. This resulted in a reduction of our original sample, which on the one side results in loss of statistical precision and on the other side also affects the representativeness of the sample. Reports from the United States were underrepresented in the effective sample. Only 9 percent of the reports in our definitive sample are from the United States, whereas these account for 24 percent of the reports contained in the HTA database. Similarly, older HTA reports are underepresented in our sample (i.e., reports published before the year 2000 account for 17 percent of the HTA database but only for 9 percent of our sample). Nevertheless, unlike previous surveys of HTA reports (8;10;18), our sampling was not restricted to a specific group of countries or agencies and is, thus, representative to a considerable extent.

Another limitation of our survey is that it did not include reports from most of the institutions involved in decision-making processes for coverage and pricing of pharmaceuticals. At the time of our sampling, the HTA database included reports from agencies with a membership by the INAHTA as well as from other ones. However, only the reports of one of the institutions involved in assessments in connection with pharmaceuticals' regulation are included in this database (i.e., of the English NICE). Assessment reports from other such agencies, for which we analyzed their methodology guidances, were not included in our survey. To our knowledge, these assessments are usually not public and are not gathered in any common database and, thus, are difficult to retrieve. In the light of the methodological recommendations of these agencies regarding the use of surrogate outcomes, we believe that the exclusion of their assessment reports from our survey, however, does not have major consequences for our results.

All in all, our study suggests that the role of surrogate outcomes in HTA is very limited, both from a theoretical point of view and from the observed practice of HTA institutions.

SUPPLEMENTARY MATERIALS

Supplementary Tables 1 and 2 (www.journals.cambridge.org/thc)

CONTACT INFORMATION

Marcial Velasco-Garrido, MPH (marcial.velasco@tu-berlin.de), Research Assistant, Health Care Management (Dir. Prof. Reinhard Busse) H80, Technische Universität Berlin, Strasse des 17. Juni, 135, Berlin, 10623, Germany
Sandra Mangiapane, PhD, MSc (smangiapane@kbv.de), Research Associate, Central Research Institute of Ambulatory Health Care in Germany, Herbert-Lewin-Platz 3, Berlin, 10623, Germany

REFERENCES

1. Antes G, Jöckel KH, Kohlmann T, Raspe H, Wasem J. *Kommentierende Synopse der Fachpositionen zur Kosten-Nutzenbewertung für Arzneimittel. Erstellt im Auftrag des Bundesministeriums für Gesundheit*. Freiburg, Essen: Greifswald, Kiel; 2007.
2. Australian Government – Department of Health and Aging. *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.1)*. Canberra: Commonwealth of Australia; 2006.

3. Biomarkers Definition Working Group. Biomarkers and surrogate end points: Preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001;69:89-95.
4. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA.* 1999;282:771-778.
5. Canadian Agency for Drugs and Technologies in Health. *Guidelines for the economic evaluation of health technologies: Canada*. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
6. Cleemput I, Crott R, Vrijens F, Huybrechts M, Van Wilder P, Ramaekers D. *Recommandations provisoires pour les évaluations pharmacoéconomiques en Belgique*. Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2006.
7. Committee of Principal Investigators. A co-operative trial in the primary prevention of ischaemic heart disease using clofibrate: Report from the Committee of Principal Investigators. *Br Heart J.* 1978;40:1069-1118.
8. Draborg E, Gyrd-Hansen D, Poulsen PB, Horder M. International comparison of the definition and the practical application of health technology assessment. *Int J Technol Assess Health Care*. 2005;21:89-95.
9. Dretzke J, Cummins C, Sandercock J, Fry-Smith A, Barrett T, Burls A. Autoantibody testing in children with newly diagnosed type 1 diabetes mellitus. *Health Technol Assess*. 2004;8:1-204.
10. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: A review of United Kingdom health technology assessment reports. *Int J Technol Assess Health Care*. 2009;25:6-13.
11. Fleming TR, de Mets DL. Surrogate end points in clinical trials: Are we being misled? *Ann Intern Med*. 1996;125:605-613.
12. Food and Drug Administration. *New drug, antibiotic and biological drug product regulations: Accelerated approval*. Proposal Rule. 57 Federal Register 13234-13242. Washington; 1992.
13. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making*. 1991;11:88-94.
14. Health Care Insurance Board. *Guidelines for pharmacoeconomic research, updated version*. Diemen: College voor zorgverzekeringen; 2006.
15. Heidenreich PA, Lee TT, Massie BM. Effect of beta-blockade on mortality in patients with heart failure: A meta-analysis of randomized clinical trials. *JAMA.* 1997;30:27-34.
16. HTAi Policy Forum Working Group on Surrogate Outcomes (2008). *Call for expressions of interest and comments*. September 2008. http://www.htai.org/index.php?id = 265.
17. Institut für Qualität und Wirtschaflichkeit im Gesundheitswesen. *Allgemeine Methoden. Version 3.0 vom 27.05.2008*. Cologne: Institut für Qualität und Wirtschaflichkeit im Gesundheitswesen; 2008.
18. Lavis JN, Wilson MG, Grimshaw J, et al. *Towards optimally packaged and relevance assessed health technology assessments. Report submitted to the Canadian Agency for Drugs and Therapeutics in Healthcare*. Hamilton: McMaster University; 2007.
19. Makni H, St-Hilaire C, Robb L, Larouche K, Blancquaert I. *Tandem mass spectrometry and neonatal blood screening in Quebec*. Montreal: AETMIS; 2007.
20. Mangiapane S, Velasco Garrido M. *Surrogate endpoints for the assessment of the benefit of health technologies*. Cologne: DIMDI; 2009 (in press).
21. Medical Services Advisory Committee. *Funding for new medical technologies and procedures: Application and assessment guidelines*. Canberra: Commonwealth of Australia; 2005.
22. MERIT-HF Study Group. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL randomized intervention trial in congestive heart failure. (MERIT-HF). *Lancet*. 1999;353:2001-2007.
23. Mundy L, Merlin T. *Ultrafast magnetic resonance imaging (MRI) for patients undergoing conventional MRI. Horizon scanning prioritising summary – Volume 2*. Canberra: Commonwealth of Australia; 2003.
24. National Institute of Health and Clinical Excellence. *Guidance for manufacturers and sponsors*. London: National Institute for Clinical Excellence; 2001.
25. National Institute for Health and Clinical Excellence. *Peginterferon alfa and ribavirin for the treatment of mild chronic hepatitis C*. London: National Institute for Clinical Excellence; 2006.
26. Norwegian Medicines Agency. *Norwegian guidelines for pharmacoeconomic analysis in connection with applications for reimbursement*. Oslo: Statens legemiddelverk; 2005.
27. Pharmaceutical Benefits Board. *General guidelines for economic evaluations from the pharmaceutical benefits board*. Stockholm: TLV; 2003.
28. Pharmaceutical Management Agency. *Recommended methods to derive clinical inputs for proposals to PHARMAC*. Christchurch; Pharmaceutical Management Agency Ltd; 2005.
29. Psaty BM, Weiss NS, Furberg CD, et al. Surrogate end points, health outcomes and the drug-approval process for the treatment of risk factors for cardiovascular disease. *JAMA*. 1999;282:786-790.
30. Sarria-Santamera A, Timoner-Aguilera J. *Assessment of the effectiveness of health promotion and disease prevention technologies*. Madrid: Agencia de Evaluación de Tecnologías Sanitarias; 2002.
31. Shepherd J, Jones J, Hartwell D, et al. Interferon alfa (pegylated and nonpegylated) and ribavirin for the treatment of mild chronic hepatitis C – a systematic review and economic evaluation. *Health Technol Assess.* 2007;11:1-224.
32. Stürzlinger H, Fröschl B, Genser D. *Evaluation of optical coherence tomography in the diagnosis of age related macular degeneration compared with fluorescence angiography*. Cologne: DIMDI; 2007.
33. Surrogate to Final Outcome Working Group. *Report of the Surrogate to Final Outcome Working Group to the Pharmaceutical Benefits Advisory Committee: A framework for evaluating proposed surrogate measures and their use in submissions to PBAC*. Canberra: 2008. www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/$File/STFOWG%20paper%20FINAL.pdf