VALUE IN HEALTH 16 (2013) S20–S26

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Current Methodological Issues in the Economic Assessment of Personalized Medicine

Lieven Annemans, PhD, MSc[1,*], *Ken Redekop, PhD, MPH*[2], *Katherine Payne, PhD, MSc*[3]

[1]*Department of Public Health, Ghent University, Ghent, Belgium;* [2]*Institute of Health Policy & Management, Erasmus University, Rotterdam, The Netherlands;* [3]*Centre for Health Economics, The University of Manchester, Manchester, UK*

ABSTRACT

There is a need for methodological scrutiny in the economic assessment of personalized medicine. In this article, we present a list of 10 specific issues that we argue pose specific methodological challenges that require careful consideration when designing and conducting robust model-based economic evaluations in the context of personalized medicine. Key issues are related to the correct framing of the research question, interpretation of test results, data collection of medical management options after obtaining test results, and expressing the value of tests. The need to formulate the research question clearly and be explicit and specific about the technology being evaluated is essential because various test kits can have the same purpose and yet differ in predictive value, costs, and relevance to practice and patient populations. The correct reporting of sensitivity/specificity, and especially the false negatives and false positives (which are population dependent), of the investigated tests is also considered as a key element. This requires additional structural complexity to establish the relationship between the test result and the consecutive treatment changes and outcomes. This process involves translating the test characteristics into clinical utility, and therefore outlining the clinical and economic consequences of true and false positives and true and false negatives. Information on treatment patterns and on their costs and outcomes, however, is often lacking, especially for false-positive and false-negative test results. The analysis can even become very complex if different tests are combined or sequentially used. This potential complexity can be handled by explicitly showing how these tests are going to be used in practice and then working with the combined sensitivities and specificities of the tests. Each of these issues leads to a higher degree of uncertainty in economic models designed to assess the added value of personalized medicine compared with their simple pharmaceutical counterparts. To some extent, these problems can be overcome by performing early population-level simulations, which can lead to the identification and collection of data on critical input parameters. Finally, it is important to understand that a test strategy does not necessarily lead to more quality-adjusted life-years (QALYs). It is possible that the test will lead to not only fewer QALYs but also fewer costs, which can be defined as "decremental" cost per QALYs. Different decision criteria are needed to interpret such results.

Keywords: guidelines, health economics, modeling, personalized medicine.

Introduction

Personalized medicine is promising from an economic perspective because in principle only those patients who are most likely to benefit from a treatment will receive that treatment. It is well recognized, however, that there is a need for well-structured economic assessment to provide robust data on the potential added value of the technology providing the personalized approach to medicine. This need for methodological scrutiny in the economic assessment of personalized medicine is consistent with any evaluation of a health care technology, and there are up to now very few specific guidelines available for the economic assessment of personalized medicine (for an example of a work in progress, see National Institute for Health and Care Excellence's [NICE's] Diagnostics Assessment Programme) [1]. This lack of specific guidance may be viewed to be appropriate given the same evaluative framework is likely to be generally applicable to identify and quantify the incremental costs and benefits of a technology that personalizes medicine. In this article, however, we present a list of 10 specific issues that we argue pose specific methodological challenges that require careful consideration when designing and conducting robust model-based economic evaluations in the context of personalized medicine. The goal of this article was to discuss these issues with reference to the standard components of guidelines on the design and conduct of model-based economic evaluations.

Specific Issues in Model-Based Economic Evaluations of Personalized Medicine

In the sections that follow, we describe and suggest solutions to 10 methodological issues that we believe are specific to the challenge of conducting robust model-based cost-effectiveness analysis of technologies designed to personalize medicine. The ultimate goal

Conflicts of interest: The authors have indicated that they have no conflicts of interest with regard to the content of this article.

* *Address correspondence to:* Lieven Annemans, Ghent University - Brussels University, De Pintelaan 185 Blok A, 2nd Floor, Ghent 9000, Belgium.

E-mail: lieven.annemans@ugent.be.

1098-3015/$36.00 – see front matter © 2013 Published by Elsevier Inc. on behalf of International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

http://dx.doi.org/10.1016/j.jval.2013.06.008

is to suggest methodological areas that need close scrutiny if the resulting economic evaluation is to provide robust and useful information for informed resource allocation decisions.

Importance of defining the scope/research question of the economic evaluation

The first step in any guideline for the good design and conduct of a cost-effectiveness analysis stresses the need to be clear about the scope of the proposed evaluation (e.g., see Drummond and Jefferson [2]). In practice, this means defining the research question and being clear about the technology to be evaluated. For pharmaceuticals, this step is relatively simple and tends to involve defining the technology in line with how it will be used as described in the product license in terms of mode of administration, dosing, and frequency. The scope of the evaluation is then completed by being clear about the relevant comparators, which should be current practice and may sometimes involve a "do-nothing" scenario. The eligible study population is then selected, which drives the subsequent scope of the evaluation. For companion diagnostic medicines, the process of defining the technology for evaluation can be more complex and adequate time and thought should be paid to ensure that the subsequent evaluation provides information relevant to how the technology will be used in clinical practice. Understanding the precise nature will often require the input of expert opinion from laboratory scientists and other clinicians who may be involved in delivering the diagnostic component of the combination product. This is because in most jurisdictions and particularly across Europe, the diagnostic component is not defined as a particular "product." Even in situations in which a biomarker, for example, has been mentioned within a product license for a particular medicine to target the eligible population, the technological approach to use will generally remain unspecified. Yet, this information is key, because it will determine the prevalence of the underlying (genetic) characteristic to be tested and hence the test results. This also means that it is left to a health service provider in the implementation stage to select which technology will be used to identify which biomarker to use in clinical practice. Clearly, this has implications for defining the technology to be evaluated within the context of an economic evaluation, in terms of both the precise nature of the diagnostic and being able to identify a relevant unit price. Furthermore, the unspecified nature of the diagnostic test will also affect the number of potential comparators in the evaluation because it may be necessary to consider whether the analysis should compare multiple ways of defining the biomarker as relevant alternatives. This challenge of defining the technology is exemplified by two health technology assessments relating to CYP2D6 testing for patients taking antipsychotics and tamoxifen. Both health technology assessments were commissioned on the basis of the availability of a marketed test, the AmpliChip CYP450 test, which tests for mutations in genes that encode the drug-metabolizing enzymes CYP2D6 and CYP2C19 [3,4]. The initial scoping phase and systematic review of the clinical literature identified numerous potential technologies, including bespoke laboratory tests, that could be used to test for a mutation in the *CYP2D6* gene as well as multiple alleles that may be relevant depending on the ethnicity of the study population. Understanding the actual companion diagnostic to be evaluated is not an insurmountable challenge but does add another layer of complexity and analysts should dedicate some time and resources to ensure that the cost-effectiveness analysis provides information for decision makers relevant to how the technology can be used in clinical practice.

Sensitivity, specificity, false negatives and false positives, and predictive value need to be considered explicitly

The test is the central component of all clinical applications of the concept of personalized medicine. The majority of companion diagnostic tests are much like humans and can and do make mistakes. Therefore, we need to examine and quantify how these diagnostic mistakes can impact the health status and costs in the design and conduct of economic evaluations of personalized medicine. *Sensitivity* and *specificity* are terms used to describe the performance (or diagnostic accuracy) of a medical test. When these terms are used to describe the ability of a medical test to detect the presence of a specific disease in a person, *sensitivity* refers to the probability that a person with the disease will have a positive (or abnormal) test result and *specificity* refers to the probability that a person without the disease will have a negative (or normal) test result. Tests can therefore make two types of mistakes: one when a patient with the disease has a negative test result (a so-called *false-negative* result) and the other when a patient without the disease has a positive test result (a *false-positive* result). But how should these terms be interpreted in the context of personalized medicine? This can be explained by using the HER2 test and trastuzumab. Studies have shown that adding trastuzumab to adjuvant chemotherapy of patients with HER2-positive breast cancer improves patient survival and reduces the risk of distant metastases [5]. Figure 1 illustrates the three general treatment options available: 1) to treat all patients with chemotherapy and trastuzumab, 2) to perform a HER2 test first and give chemotherapy and trastuzumab to all patients with a HER2-positive test result and chemotherapy only to all patients with a HER2-negative test result, or 3) to treat all patients with chemotherapy only. There are different ways, however, to assess the HER2 status of a tumor, including immunohistochemistry and fluorescence in situ hybridization, and these methods vary in their ability to establish whether a tumor is HER2-positive or HER2-negative. Moreover, there is also variation in interpretation methods. In other words, the different HER2 tests and methods vary in sensitivity and specificity; some of the tumors labeled as HER-positive will actually be HER2-negative (false positive), while some of the tumors labeled as HER-negative will actually be HER2-positive (false negative). What does this mean in clinical practice? The false-positive patients are HER2-negative patients who will receive, but not benefit from, trastuzumab treatment;

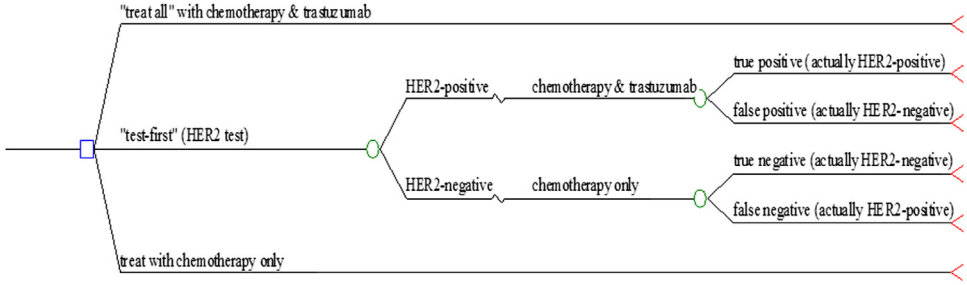

Fig. 1 – Example showing three treatment strategies (treat all, test first, and treat with chemotherapy only).