International Journal of Technology Assessment in Health Care, 32:1/2 (2016), 46–53.

doi:10.1017/S0266462316000064

THE EARLY BIRD CATCHES THE WORM: EARLY COST-EFFECTIVENESS ANALYSIS OF NEW MEDICAL TESTS

Leander R. Buisman
Institute of Health Policy and Management, Erasmus University Rotterdam; Institute for Medical Technology Assessment, Erasmus University Rotterdam
buisman@bmg.eur.nl

Maureen P.M.H. Rutten-van Mölken
Institute of Health Policy and Management, Erasmus University Rotterdam; Institute for Medical Technology Assessment, Erasmus University Rotterdam

Douwe Postmus
Department of Epidemiology, University of Groningen, University Medical Center Groningen

Jolanda J. Luime
Department of Rheumatology, Erasmus MC, University Medical Center Rotterdam

Carin A. Uyl-de Groot, William K. Redekop
Institute of Health Policy and Management, Erasmus University Rotterdam; Institute for Medical Technology Assessment, Erasmus University Rotterdam

Objectives: There is little specific guidance on performing an early cost-effectiveness analysis (CEA) of medical tests. We developed a framework with general steps and applied it to two cases.
Methods: Step 1 is to narrow down the scope of analysis by defining the test's application, target population, outcome measures, and investigating current test strategies and test strategies if the new test were available. Step 2 is to collect evidence on the current test strategy. Step 3 is to develop a conceptual model of the current and new test strategies. Step 4 is to conduct the early-CEA by evaluating the potential (cost-)effectiveness of the new test in clinical practice. Step 5 involves a decision about the further development of the test.
Results: The first case illustrated the impact of varying the test performance on the headroom (maximum possible price) of an add-on test for patients with an intermediate-risk of having rheumatoid arthritis. Analyses showed that the headroom is particularly dependent on test performance. The second case estimated the minimum performance of a confirmatory imaging test to predict individual stroke risk. Different combinations of sensitivity and specificity were found to be cost-effective; if these combinations are attainable, the medical test developer can feel more confident about the value of further development of the test.
Conclusions: A well-designed early-CEA methodology can improve the ability to develop (cost-)effective medical tests in an efficient manner. Early-CEAs should continuously integrate insights and evidence that arise through feedback, which may convince developers to return to earlier steps.

Keywords: Early cost-effectiveness analysis, Medical test, Decision support, Research and development, Manufacturer, Test developer

In healthcare systems worldwide, comparative cost-effectiveness research is mostly done in the late stages of medical test[1] development. Such evaluation is mainly used by medical test manufacturers/developers (referred to as test developers) to demonstrate to payers (i.e., healthcare insurers, governments, and managed care organizations) that a test is good value for money (2–6). To test developers, adequate reimbursement of new tests is important for wide implementation in clinical practice, which improves return on investment. Medical tests are most often reimbursed as part of a Diagnosis Related Group (DRG) payment or fee-for-service. Hence, those who actually decide about reimbursement include opinion leaders among the clinicians, hospital managers, government authorities, and healthcare insurers. Given the serious consequences of negative reimbursement decisions, test developers take a considerable risk when conducting cost-effectiveness analyses only in the late stages of medical test development, when large research and development investments have already been made. Therefore, evaluation of new tests in the early stages of development has attracted increasing attention.

Early cost-effectiveness analysis (CEA) helps test developers to decide about further development of medical tests, set realistic performance-price goals, and design and manage reimbursement strategies (5). Early-CEAs may guide the resources invested in the development process. However, they require close cooperation between test developers and the researchers performing early-CEAs. Over the past 2 decades, most early-CEAs focused on new drug therapy (7). The increasing use of medical tests in various phases of disease prevention and treatment, including companion diagnostics to "personalize" medicine, calls for early-CEAs to assess how much these tests could really improve health outcomes and healthcare efficiency. However, there is little specific guidance on performing early-CEAs of medical tests. Some general steps of early-CEAs were

This research was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), projects PARISk (grant 01C-202) and TRACER (grant 04I-202), and supported by the Dutch Heart Foundation and the Dutch Arthritis Foundation.

[1] Medical tests are used "to determine the presence or absence of a definite disease or of some substance in any of the fluids, tissues, or excretions of the body, or to determine the presence or degree of a psychological or behavioural trait" (1). An early-CEA can be important for any type of device, including the medical tests addressed in this study.

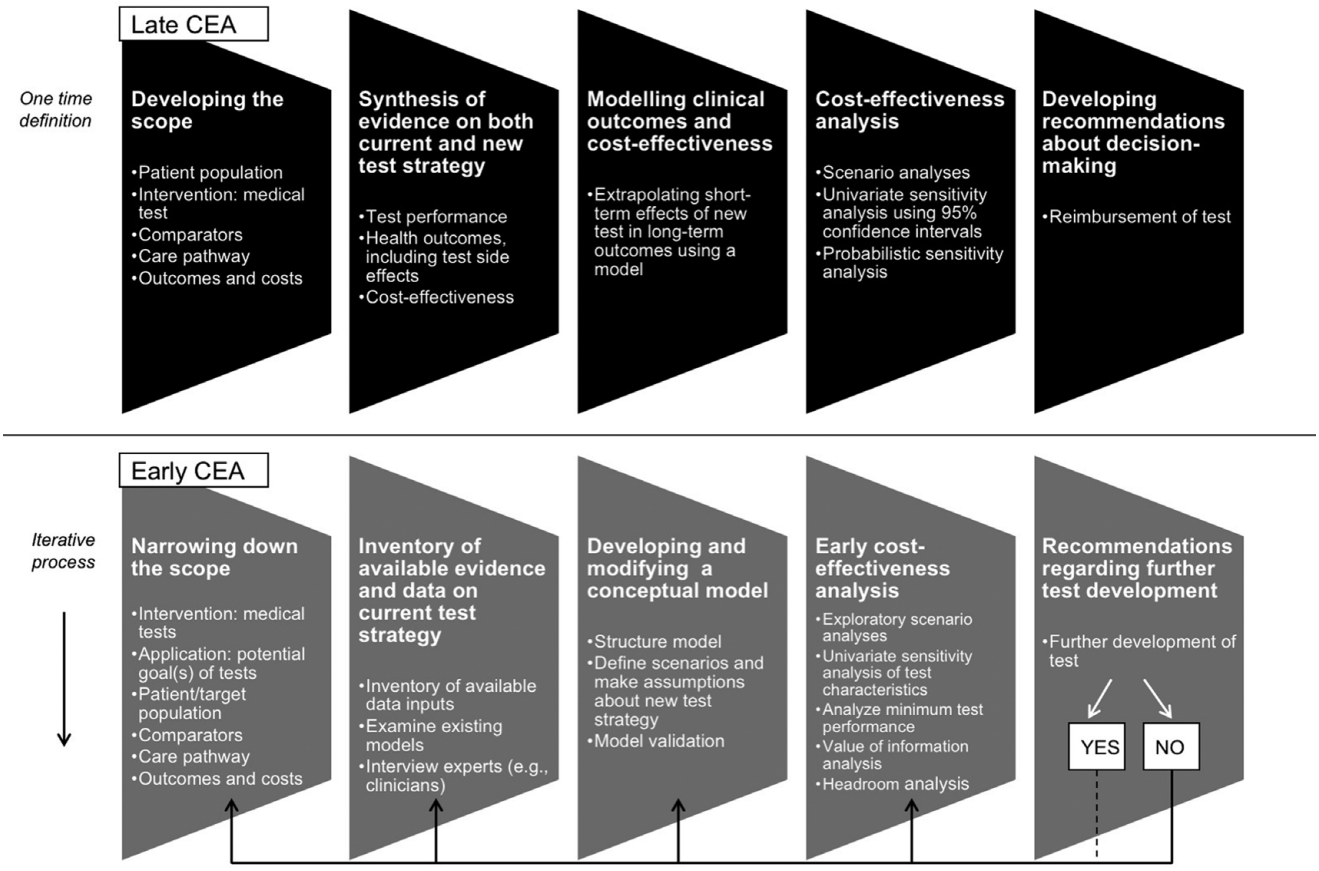

Figure 1. Differences and similarities of late and early-CEAs of medical tests. Steps of late-CEA based on *NICE Diagnostic Assessment Programme Manual* (9); Steps of early-CEA were defined by the authors.

described by O'Prinsen et al. (8). In this study, we developed a framework with the general steps of early-CEAs of new medical tests and applied it to two cases. The development of the framework was an iterative process because in the cases we applied the methods in our framework but used the experience and insights gained from the cases to refine the framework.

GENERAL STEPS OF EARLY-CEAS

Before presenting the general steps of early-CEAs, we describe the differences between early and late-CEAs of tests. Figure 1 shows the five general steps of late and early-CEAs of medical tests and the main differences between them.

The general steps of *late-CEAs* of medical tests are based on the Diagnostic Assessment Programme (DAP) Manual from NICE (9). The DAP manual promotes the consistent and rapid adoption of clinically innovative and cost-effective medical tests in the United Kingdom. There are at least four main differences between early and late-CEAs of medical tests. First, the first step of a late-CEA of medical tests is to *define* the scope, while the first step of an early-CEA is *to narrow down* the scope. Second, early-CEAs are much more iterative than late-CEAs. Throughout the development process of a test, new data and ideas may emerge, which may convince developers to return to earlier steps. Third, much less data are available for early-CEAs than for late-CEAs, making sensitivity analyses of early-CEAs much more exploratory. Fourth, late-CEAs are mainly used by payers in reimbursement decisions, while early-CEAs are used by test developers for internal decision making about further development of a test and setting realistic performance-price goals. The following section describes the five steps of an early-CEA as shown in Figure 1.

Step 1: Narrowing Down the Scope

The test developer needs to consider where in the healthcare system the new test will be used. Figure 2 shows a variety of applications of medical tests following the sequence of screening/case finding, diagnosis, disease progression and treatment.

One way to start exploring where a test could be used in patient care is to apply the Patient Population, Intervention, Comparator, and Outcomes (PICO) method (11). This way, one can systematically compare different potential areas of application in the healthcare system. If the application of the test is clear, the target population may also be clear. However, if the application is not clear, literature research and discussions with test developers and clinicians on areas of highest unmet need and greatest added benefit would help to determine the application in the healthcare system and define the target population. Given the context of an early-CEA, we extend the PICO

https://doi.org/10.1017/S0266462316000064 Published online by Cambridge University Press

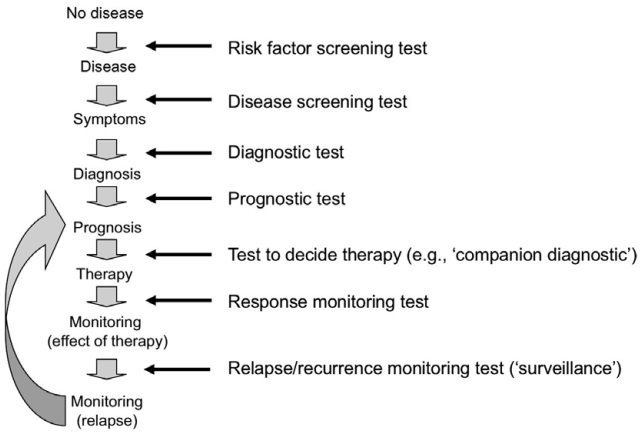

Figure 2. Various applications of medical tests in the healthcare system. The vertical arrows represent the different phases of disease and its treatment (based on Redekop and Uyl-de Groot (10)).

method and re-order the different elements into an "APCOI" (Application, Patient Population, Comparator, Outcomes, Intervention) method, resulting in the following questions: (i) Application (in healthcare): What is the anticipated application of the test in the healthcare system? (ii) Patient Population (participants): What is the target population? (iii) Comparator: How are current test strategies (i.e., current tests and treatment options based on the test result), anticipated future comparators, and resulting clinical care organized? (iv) Outcomes: How will effectiveness be defined and which costs need to be taken into account? (v) Intervention: What will clinical care look like with the new medical test?

Discussions with different stakeholders are essential during test development. For example, a test can be cost-effective, but if clinicians are not convinced that the test improves patient care, the new test will not be used in clinical practice and will not yield a sufficient return on investment.

Step 2: Inventory of Available Evidence and Data on Current Test Strategy

A proper examination of the strengths and weaknesses of current care is invaluable in estimating the added value of the new test. For example, if current care is already highly effective, it may be difficult, if not impossible, to improve upon it (although the new test can still be considered cost-effective if it reduces costs). Researchers should examine how current test strategies, anticipated future comparators, and clinical care are organized and which evidence and data are available on the costs and health outcomes. In addition, potentially relevant existing models of the disease and target population should be reviewed.

Step 3: Developing and Modifying a Conceptual Model

A conceptual model of both the current and new test strategies should be developed in Step 3 and informed by Step 2. Because little is known about the impact of the new test strategy, various scenarios should be defined, which could include varying subpopulations as part of the target population (e.g., specific age-sex groups), the prevalence of the disease of interest, applications of the test, costs, test performance, and health improvements. Due to the stronger iterative nature of early-CEAs as compared to late-CEAs, conceptual models should be revised as new insights and evidence arise during test development. The model's validity should be scrutinized on its validity similar to late-stage modeling (face, cross-model, external, predictive, and verification validity), to the extent that is possible in the early stages of test development (12).

Step 4: Early Cost-effectiveness Analysis

An early-CEA can be conducted to evaluate the potential impact of the new test in clinical practice when all parameters and their values have been determined. Normally, initial estimates of the model parameters may have to be derived from expert opinion, observational studies or small clinical trials. Therefore, exploratory scenario analyses can help to set a benchmark for the minimum performance that is required for a test to become cost-effective compared with current practice. For example, one could derive the minimum sensitivity and specificity at which the test becomes an attractive alternative from an effectiveness standpoint. These scenario analyses can be reapplied and modified throughout the entire development process. Moreover, an early-CEA should at least include a univariate sensitivity analysis in which a range of parameters are varied to identify which of them have the most impact on incremental costs, effectiveness and cost-effectiveness. When uncertainty can be quantified, a value of information analysis is recommended in the early stages of test development to decide whether additional research is needed to decide which test scenario should be chosen (13–15).

Although early-CEAs provide valuable insights into the clinical and economic value of new medical tests, they do not directly provide an answer to the question of whether the test developer should continue developing the test. To address this question, the results must be translated into an estimate of the test's maximum sales price. This price can be derived from the cost-effectiveness model, for any given combination of parameters. This price can be further substantiated by applying the principle of value-based pricing, in which the price is largely driven by the innovative nature of the test and the extent to which it addresses unmet needs, and by making assumptions about how the volume of sales is affected by the maximum sales price (16). The maximum sales price can then be fed into an appropriate product-investment evaluation method, such as the headroom method (16). The headroom (or potential profit) method assesses the maximum additional cost at which the medical test is still likely to be cost-effective at a given willingness-to-pay threshold (17;18).

Step 5: Developing Recommendations Regarding Further Test Development

The results of early-CEAs can help test developers of medical tests to decide about the further development of the test (go/no-go decision). If an early-CEA shows that the test is unfeasible

https://doi.org/10.1017/S0266462316000064 Published online by Cambridge University Press

or unlikely to be cost-effective, it is unlikely that the test will be reimbursed. Therefore, test developers may decide to return to earlier steps of the early-CEA. Even if an expensive test is found to be cost-effective, it may be difficult for clinicians to use the test until the higher test costs have been incorporated into a higher DRG payment.

CASES OF EARLY-CEAS OF MEDICAL TESTS

This section applies the steps of early-CEAs of medical tests to two cases: a diagnostic test to detect early rheumatoid arthritis (RA) and a prognostic test to assess the risk of a recurrent ischemic stroke.

Case 1: Diagnostic Test for RA

RA is a chronic inflammatory disease characterized by structural irreversible joint damage, leading to severe disability and premature death (19–24). Early detection is important because early treatment with synthetic and biologic disease modifying antirheumatic drugs (DMARDs) has been shown to slow down disease progression (25–29).

When applying the general steps of an early-CEA, we started to narrow down the scope of the analysis to tests that help to diagnose RA in an early stage (Step 1). The current diagnostic standard for RA, and thus the comparator in the CEA, is the ACR/EULAR 2010 RA classification criteria (referred to as RA-2010 criteria) (24;30;31). Use of the RA-2010 criteria at baseline as a risk prediction tool usually results in a considerable proportion of early arthritis patients being classified as having an intermediate-risk to develop RA in the near future (3–5 points). What is needed is an additional test to reclassify these intermediate-risk patients into high-risk and low-risk ones. A B-cell test is a candidate for this purpose (32).

As part of Step 2 of our early-CEA, we examined how current clinical care is organized using the RA-2010 criteria. Moreover, we investigated which tests are used to diagnose RA by conducting interviews with rheumatologists and analyzing resource use data. We found that a variable number of diagnostic tests (average: thirty-two diagnostic tests per patient) is requested by rheumatologists during the initial outpatient visit to exclude differential diagnoses of RA (Benner et al., 2015, unpublished data). The mean diagnostic costs per patient were €422 (SD: €168). In addition, interviews were conducted with the test developer of the B-cell test about the performance and costs of the test to diagnose RA in a population of early arthritis patients. The B-cell test has been studied in a cohort of seropositive arthralgia patients (i.e., subsample of early arthritis patients) and a sensitivity of 60 percent and specificity of 90 percent were found (32). We assumed that this test performance was applicable for all early arthritis patients.

As Step 3, a conceptual decision model with a 5-year time horizon was developed for the diagnosis and treatment of patients with early arthritis based on evidence obtained in Step 2 (see Supplementary Figure 1 for the model). Interviews with rheumatologists and test developer were conducted about the potential use of B-cell test in the early diagnosis of RA to formulate test scenarios. In the new test strategy, a B-cell test was used as add-on for patients with an intermediate-risk according to the RA-2010 criteria. In this strategy, intermediate-risk patients can be reclassified as high-risk or low-risk. At a sensitivity of 60 percent and specificity of 90 percent, 29 percent (75/263) would be reclassified as high-risk and 71 percent (188/263) as low-risk. Patients were classified as true positive (TP) if they had a positive test result (scored ≥6 points on the RA-2010 criteria or were B-cell positive) at baseline and used methotrexate (MTX) at 12 months. Patients were considered as true negative (TN) if they had a negative test result (scored <6 points on the RA-2010 criteria or were B-cell negative) at baseline and did not use MTX at 12 months. False positive (FP) patients were patients who scored ≥6 points on the RA-2010 criteria or were B-cell positive at baseline but did not use MTX at 12 months. Patients were classified as false negative (FN) if they scored <6 points on the RA-2010 criteria or were B-cell negative at baseline but used MTX at 12 months.

Patients classified as TP or FN at 12 months were defined as RA patients and entered a patient-level state transition model at 12 months in which the disease activity (DAS28) and treatment course were simulated based on data from two cohorts (REACH (30) and tREACH (31,33)) and published data (34). Patients could switch from MTX to more expensive biologic DMARDs if they had a DAS28 ≥ 3.2 with a swollen joint count > 0 and no comorbidity. Treatment with a biologic DMARD incurs additional treatment costs but also improves health-related quality of life (EQ-5D utilities). Patients classified as TN or FP at 12 months entered a background model in which they stayed for the remaining four years, assuming no change in utilities, biologic DMARD costs for 10% of FPs in the first year after diagnosis, and otherwise no RA-related costs.

The decision model was populated with data of patients in the REACH cohort (30). The prevalence of RA among these patients is 54 percent at 12 months using the RA-2010 criteria. The quality of life (EQ-5D) at time of diagnosis and follow-up was obtained from the literature and the REACH cohort. Direct medical costs were based on the Dutch Manual of Costing (35). Costs of blood tests were based on tariffs provided by the Dutch Healthcare Authority (36), and costs of MTX and biologic DMARDs were obtained from the National Health Care Institute (37). All costs were adjusted to 2014 Euros using the general price index from the Dutch Central Bureau of Statistics (38), and a healthcare perspective was used.

In Step 4, exploratory scenario analyses were performed using different potential test scenarios of the B-cell test. In this study, we present the results of the B-cell test as *add-on test for patients with an intermediate-risk* based on the RA-2010 criteria. The sensitivity and specificity of these criteria in the REACH data were estimated to be 62 percent and 77 percent,

https://doi.org/10.1017/S0266462316000064 Published online by Cambridge University Press

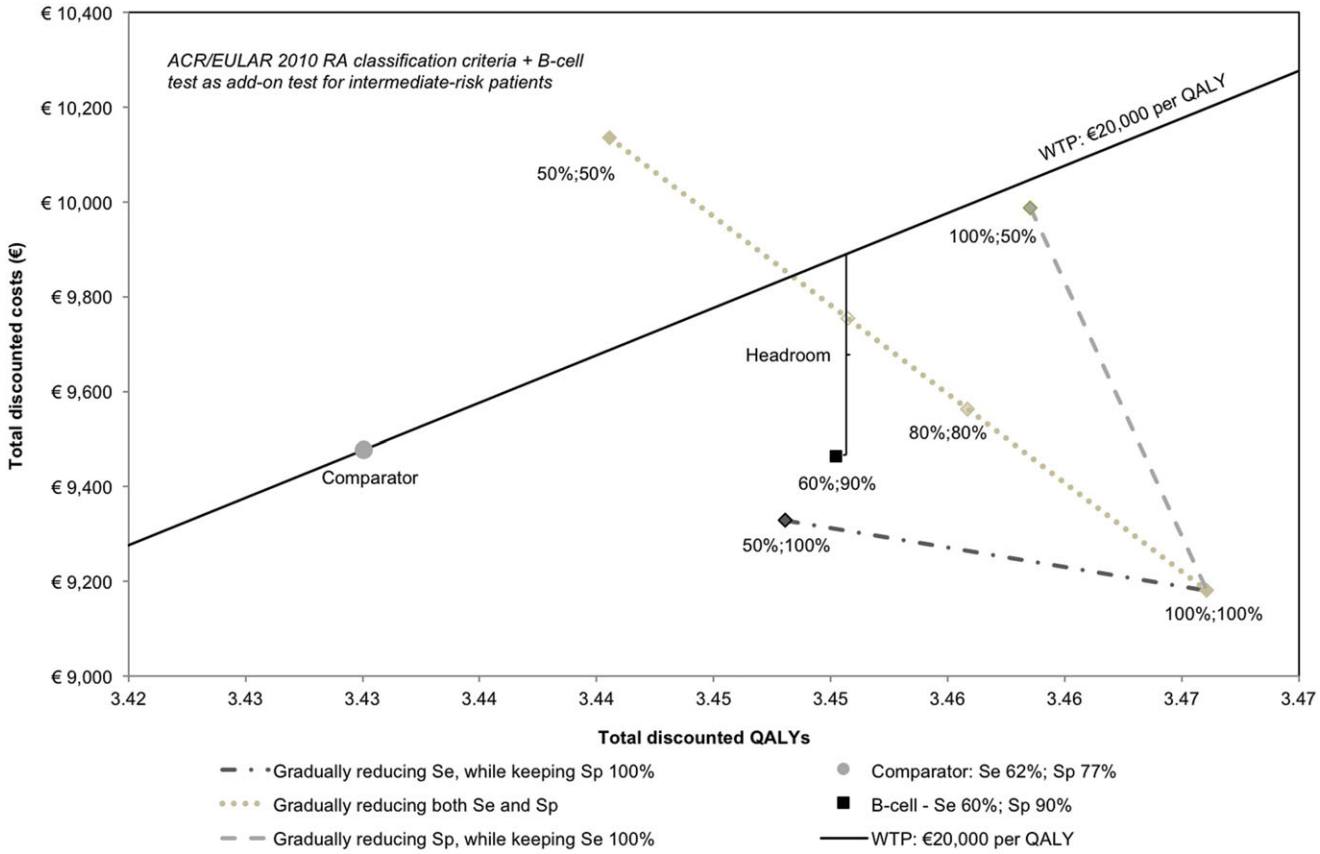

Figure 3. Cost-effectiveness plane for a B-cell test as add-on for patients with an intermediate-risk based on the ACR/EULAR 2010 RA classification criteria. The figure presents the sensitivity and specificity of an add-on test for intermediate-risk patients in addition to the ACR/EULAR 2010 RA classification criteria; Se = sensitivity; Sp = specificity; WTP = willingness-to-pay; QALY = quality-adjusted life year.

respectively. Probabilistic sensitivity analysis was performed and the maximum cost of a B-cell test required for the incremental cost-effectiveness ratio (ICER) of the new test strategy to stay below €20,000 per quality-adjusted life year (QALY) (i.e., the headroom) was calculated (i.e., headroom). Figure 3 shows the impact of varying the sensitivity and specificity between 50 percent and 100 percent on the headroom of a B-cell test.

An add-on B-cell test with 100 percent sensitivity and specificity dominates the RA-2010 criteria because it reduces costs (by €296) and increases health (by 0.036 QALYs). Similarly, a B-cell test with 60 percent sensitivity and 90 percent specificity dominates the comparator because it also reduces costs (by €14) and increases health (by 0.020 QALYs). The headroom of the B-cell test (€417) is shown in Figure 3 by the difference in costs between the willingness-to-pay threshold line of €20,000 per QALY gained and the cost-effectiveness of the new test strategy (where the test is assumed to be free).

Case 2: Prognostic Test for Recurrent Stroke

Patients with a recent transient ischemic attack (TIA) or minor ischemic stroke are at risk of a recurrent ischemic stroke, which may be caused by a plaque rupture in the carotid artery. Surgery (carotid endarterectomy) can reduce this risk, but the procedure can lead to death and morbidity. Better stroke risk prediction would therefore help to determine which patients should undergo surgery. Noninvasive molecular imaging technologies, such as contrast enhanced magnetic resonance imaging, computed tomography angiography, and biomechanical analysis are technologies that can be used to improve risk prediction.

When applying the general steps of an early-CEA, we narrowed down the scope of the analysis using the APCOI method (Step 1). The application (and target population) of the new medical test was a *prognostic test* to predict the risk of a recurrent stroke caused by plaque rupture in *patients with recent TIA or minor ischemic stroke and 30–69 percent carotid stenosis*. We then examined how current test strategies and clinical care are organized and discovered a substantial amount of variation. For example, the national stroke guidelines were followed by 60 percent of Dutch hospitals while other hospitals used various other test combinations (39). The comparators representing current clinical practice were therefore defined as "guideline-based care" (40) and other current test strategies. For the new test strategy, interviews were conducted with test developers and clinicians about the optimal combination of tests and potential test performance, because many combinations are possible. The most likely application for a new test would be as a confirmatory imaging test for patients with a 30–69 percent carotid stenosis according to an initial duplex ultrasonography (sensitivity: 89

https://doi.org/10.1017/S0266462316000064 Published online by Cambridge University Press

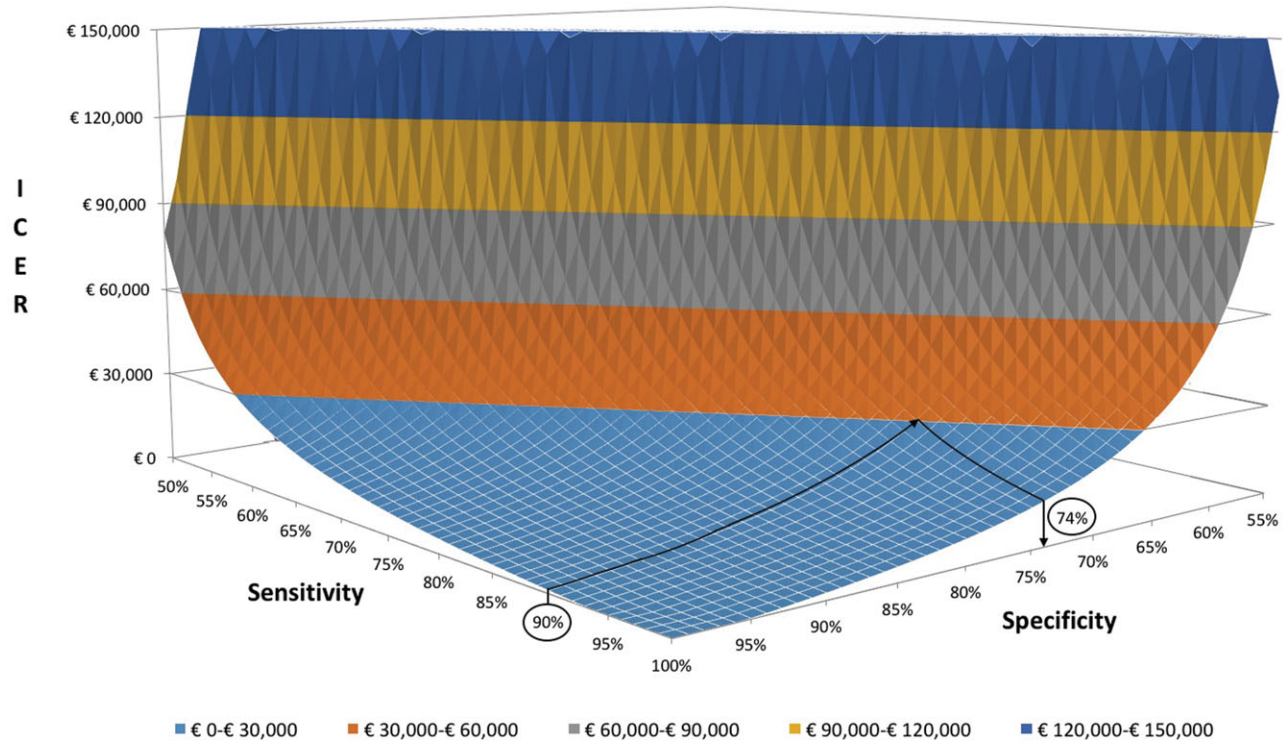

Figure 4. Minimum sensitivity and specificity of a new confirmatory imaging test (test costs: €362). The minimum values of sensitivity and specificity can be found by starting from a value of sensitivity at the sensitivity axis, moving vertically up to the corresponding gridline, following the gridline to a predefined threshold regarding the willingness-to-pay to gain one quality-adjusted life year (QALY), following the gridline for specificity to the right-hand side, and then moving vertically down to the specificity axis; ICER = incremental cost-effectiveness ratio.

percent and specificity: 84 percent [41]; costs: €125 [36]). We assumed that if the confirmatory imaging test identified patients as being at high-risk of a recurrent stroke, patients underwent surgery, while patients with a low-risk of recurrent stroke received medicines alone.

As Step 2, we examined the costs and health outcomes of current care using clinical stroke guidelines and literature. These findings were discussed with vascular neurologists and radiologists to assess the quality and relevance. Also, any existing ischemic stroke CEA models were reviewed including those for diagnostic tests (e.g., Tholen et al. [41]).

The results from Step 2 were used to develop a conceptual model in Step 3. The model consisted of 3 parts: prognostic testing, treatment, and health outcomes. The initial version focused on the use of a new confirmatory imaging test for patients with 30–69 percent carotid stenosis tested with an initial duplex ultrasonography.

At present, little is known about the performance of a test to predict the risk of plaque rupture. Therefore, decision modeling was used to estimate the minimum test performance (i.e., sensitivity and specificity) that a new confirmatory imaging test must have to be cost-effective compared with current care (Step 4). Exploratory sensitivity analyses were performed to identify which combinations of test costs and performance resulted in acceptable cost-effectiveness ratios. All costs were adjusted to 2014 Euros using the general price index from the Dutch Central Bureau of Statistics (38), and a societal perspective was used.

A perfect confirmatory imaging test (100 percent sensitivity and specificity) with a cost of €362 for 60-year-old men appears to be cost-effective compared with "guideline-based care" using a life-time horizon. A perfect confirmatory imaging test dominates "guideline-based care" because it reduces costs (by €110) and increases health (by 0.066 QALYs), because the test identifies all patients correctly and ensures that they all receive the appropriate treatment. Figure 4 shows the minimum values of sensitivity and specificity needed for a new confirmatory imaging test at different thresholds regarding the willingness-to-pay to gain one QALY. For example, if a test is 90 percent sensitive, it must have a specificity of at least 74 percent to be cost-effective given a willingness-to-pay threshold of €30,000 per QALY gained. This combination can be found in Figure 4 by starting from the sensitivity axis at 90 percent, following the gridline to the boundary between the blue and orange parts (i.e., willingness-to-pay threshold of €30,000 per QALY gained), following the gridline for specificity to the right-hand side, and then moving vertically down to the specificity axis.

DISCUSSION

CEAs in the early stages of medical test development have important benefits. For test developers, they are useful in guiding further development of tests, for example, by estimating the maximum cost of a new test and the minimum test performance required for the test to be cost-effectiveness. For

https://doi.org/10.1017/S0266462316000064 Published online by Cambridge University Press

clinicians, early-CEAs provide valuable information about the patient (sub)populations in which the test is potentially cost-effective. An early-CEA can convince clinicians of the potential improvements in patient care and health outcomes which may result in faster take-up of tests in clinical practice. Published early-CEA results can also help payers to identify promising new tests early, resulting in more timely decisions about reimbursement. A major advantage of early-CEA is that it generates optimal product development and pricing, although it might seem resource-intensive at first. This was illustrated by the two cases. We showed how the inevitable uncertainties in the early stage of the development can be addressed by applying different scenario and sensitivity analyses.

New tests can lead to significant improvements in health outcomes and efficiency only if effective treatments are available. In the RA case, effective treatment was available to slow down disease progression in patients diagnosed with a high-risk of having RA, while in the stroke case, patients with a high-risk of recurrent ischemic stroke underwent surgery.

We developed a framework with general steps of an early-CEA of new medical tests. Our framework is developed to evaluate the potential cost-effectiveness of new medical tests and is not completely applicable to new drugs (see Drummond et al. [42] for important differences in CEAs for tests and drugs). Some studies have previously been performed, which can be seen as examples of the steps in our framework. Postmus et al. (17) describe in more formal terms how Steps 2–5 can be performed for a risk factor screening test. Cao et al. (43) describe how to handle the use of expert opinion in an early-CEA model of medical tests in a probabilistic way. We recommend further research that applies our proposed framework of early-CEAs of medical tests to the assessment of specific tests.

Applying the framework to our example cases, we saw that varying the sensitivity and specificity influenced the headroom of an add-on test for intermediate-risk RA patients. To calculate the headroom, we used a fixed willingness-to-pay threshold per QALY gained, but different thresholds led to different estimates of the maximum sales price. If this maximum cost offers sufficient degree of headroom from a commercial standpoint, test developers may opt to continue developing the test as planned. In the second case, to predict individual stroke risk, different combinations of sensitivity and specificity of a new confirmatory imaging test were cost-effective at a given willingness-to-pay threshold. Decisions about further test development may therefore be dependent on the threshold used.

The framework has also some potential limitations. First of all, the general steps might be too general to make the framework applicable to all early-CEAs of different types of tests in all diseases; the necessary test-specific details will have to be developed and documented. Furthermore, the success of a new medical test depends on more factors than cost-effectiveness. Factors such as total revenue, future market with potential competitors, future investments during test development, costs of scaling-up the production, stakeholder preferences, and marketing among professionals should be considered by test developers in the early stages of development. A business case developed and refined during the early stages should incorporate these factors (4;5). At the later stages, payers will have important considerations, such as the safety of the test and budget impact.

CONCLUSION

A well-designed early-CEA methodology as presented in this study can improve the ability to develop effective and cost-effective medical tests in an efficient manner. The continuous integration of new insights and evidence that arise through feedback during the test development may convince developers to return to earlier development steps and will result in more informed decisions by test developers about its potential application in the healthcare system and target population.

SUPPLEMENTARY MATERIAL

Supplementary Figure 1
http://dx.doi.org/10.1017/S0266462316000064

CONFLICTS OF INTEREST

The authors report no conflicts of interest relevant to the manuscript.

REFERENCES

1. Dirckx JH. *Stedman's concise: Medical dictionary for the health professions*. 4th ed. Baltimore: Lippincott Williams & Wilkins; 2001.
2. Drummond MF. Economic evaluation of pharmaceuticals. Science or marketing? *Pharmacoeconomics*. 1992;1:8-13.
3. Johannesson M. Economic evaluation of drugs and its potential uses in policy-making. *Pharmacoeconomics*. 1995;8:190-198.
4. Annemans L, Genesté B, Jolain B. Early modelling for assessing health and economic outcomes of drug therapy. *Value Health*. 2000;3:427-434.
5. Pietzsch JB, Paté-Cornell ME. Early technology assessment of new medical devices. *Int J Technol Assess Health Care*. 2008;24:36-44.
6. IJzerman MJ, Steuten LM. Early assessment of medical technologies to inform product development and market access: A review of methods and applications. *Appl Health Econ Health Policy*. 2011;9:331-347.
7. Hartz S, John J. Contribution of economic evaluation to decision making in early phases of product development: A methodological and empirical review. *Int J Technol Assess Health Care*. 2008;24:465-472.
8. O'Prinsen AC, Gaultney J, Redekop WK. Universal steps for conducting early-stage medical technology assessment. *ISPOR Connect*. 2009;November/December:4-6.
9. National Institute for Health and Care Excellence (NICE). *Diagnostic assessment programme manual*. London: National Institute for Health and Care Excellence; 2011.
10. Redekop K, Uyl-de Groot C. Diagnostiek en economische evaluatie. In: Rutten-van Molken M, Uyl-de Groot C, Rutten F, eds. *Van kosten tot effecten: Een handleiding voor economische evaluatiestudies in de gezondheidszorg*. Amsterdam: Elsevier Gezondheidszorg; 2010. p 157-176.
11. O'Connor D, Green S, Higgins JPT. Defining the review question and developing criteria for including studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews

of Interventions Version 5.1.0. The Cochrane Collaboration; 2011. http://www.cochrane-handbook.org (accessed May 2, 2014).
12. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making*. 2012;32:733-743.
13. Claxton K. The irrelevance of inference: A decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ*. 1999;18:341-364.
14. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: When is there sufficient evidence? *Value Health*. 2005;8:433-446.
15. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: Some lessons from recent UK experience. *Pharmacoeconomics*. 2006;24:1055-1068.
16. Girling A, Young T, Brown C, Lilford R. Early-stage valuation of medical devices: The role of developmental uncertainty. *Value Health*. 2010;13:585-591.
17. Postmus D, De Graaf G, Hillege HL, Steyerberg EW, Buskens E. A method for the early health technology assessment of novel biomarker measurement in primary prevention programs. *Stat Med*. 2012;31:2733-2744.
18. Cosh E, Girling A, Lilford R, McAteer H, Young T. Investing in new medical technologies: A decision framework. *J Commerc Biotechnol*. 2007;13:263-271.
19. Pincus T, Callahan LF, Sale WG, et al. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum*. 1984;27:864-872.
20. Mitchell DM, Spitz PW, Young DY, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum*. 1986;29:706-714.
21. Scott DL, Symmons DP, Coulton BL, Popert AJ. Long-term outcome of treating rheumatoid arthritis: Results after 20 years. *Lancet*. 1987;1:1108-1111.
22. Isomäki H. Long-term outcome of rheumatoid arthritis. *Scand J Rheumatol Suppl*. 1992;95:3-8.
23. Wolfe F. The natural history of rheumatoid arthritis. *J Rheumatol Suppl*. 1996;44:13-22.
24. Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis*. 2010;69:1580-1588.
25. Bijlsma JW, Weinblatt ME. Optimal use of methotrexate: The advantages of tight control. *Ann Rheum Dis*. 2007;66:1409-1410.
26. Jacobs JW. Optimal use of non-biologic therapy in the treatment of rheumatoid arthritis. *Rheumatology (Oxford)*. 2012;51:iv3-8.
27. Chen YF, Jobanputra P, Barton P, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. *Health Technol Assess*. 2006;10:iii–iv, xi–xiii, 1-229.
28. Donahue KE, Garthehner G, Jonas DE, et al. *Comparative effectiveness of drug therapy for rheumatoid arthritis and psoriatic arthritis in adults*. Rockville, MD: AHRQ Comparative Effectiveness Reviews; 2007.
29. Singh JA, Christensen R, Wells GA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: A Cochrane overview. *CMAJ*. 2009;181:787-796.
30. Alves C, Luime JJ, Van Zeben D, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis*. 2011;70:1645-1647.
31. Claessen SJ, Hazes JM, Huisman MA, et al. Use of risk stratification to target therapies in patients with recent onset arthritis; Design of a prospective randomized multicenter controlled trial. *BMC Musculoskelet Disord*. 2009;10:71.
32. Van Baarsen LG, Bos WH, Rustenburg F, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum*. 2010;62:694-704.
33. De Jong PH, Hazes JM, Han HK, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis*. 2014;73:1331-1339.
34. Vermeer M, Kievit W, Kuper HH, et al. Treating to the target of remission in early rheumatoid arthritis is cost-effective: Results of the DREAM registry. *BMC Musculoskelet Disord*. 2013;14:350.
35. Tan SS, Bouwmans CA, Rutten FF, Hakkaart-van Roijen L. Update of the Dutch Manual for Costing in Economic Evaluations. *Int J Technol Assess Health Care*. 2012;28:152-158.
36. Dutch Healthcare Authority. Tariffs application [Internet]. Utrecht: Dutch Healthcare Authority. http://dbc-zorgproducten-tarieven.nza.nl/nzaZpTarief/ZoekfunctieDot.aspx (accessed February 4, 2015).
37. National Health Care Institute [Internet]. Diemen: National Health Care Institute. http://www.medicijnkosten.nl/default.asp (accessed February 4, 2015)
38. Statistics Netherlands [Internet]. The Hague: Statistics Netherlands. http://statline.cbs.nl (accessed February 4, 2015)
39. Buisman LR, Rijnsburger AJ, Den Hertog HM, Van der Lugt A, Redekop WK. Clinical practice variation needs to be considered in cost-effectiveness analyses: A case study of patients with a recent transient ischemic attack or minor ischemic stroke. *Appl Health Econ Health Policy*. 2016;14:67-75.
40. Dutch Institute for Healthcare Improvement CBO. *Guideline: Diagnostics, treatment and healthcare for patients with a stroke* (In Dutch: Richtlijn: Diagnostiek, behandeling en zorg voor patiënten met een beroerte). Utrecht: Dutch Institute for Healthcare Improvement CBO; 2008.
41. Tholen ATR, De Monyé C, Genders TSS, et al. Suspected carotid artery stenosis: Cost-effectiveness of CT angiography in work-up of patients with recent TIA or minor ischemic stroke. *Radiology*. 2010;256:585-597.
42. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs–Same or different? *Value Health*. 2009;12:402-404.
43. Cao Q, Postmus D, Hillege HL, Buskens E. Probability elicitation to inform early health economic evaluations of new medical technologies: A case study in heart failure disease management. *Value Health*. 2013;16:529-535.

https://doi.org/10.1017/S0266462316000064 Published online by Cambridge University Press