VALUABLE AI FOR HEALTH

EARLY HTA ON THE VALUE OF AN AI-BASED DECISION SUPPORT SYSTEM IN MULTIPLE SCLEROSIS

Externally translated edition

institute for Medical Technology Assessment
Erasmus University Rotterdam

Valérie Wester, MSc

Simone Huygens, PhD

Matthijs Versteegh, PhD

Study Sponsor:	Ministry of Health, Welfare and Sport
Department:	Information Policy Directorate
Therapeutic Area:	Artificial Intelligence, Multiple Sclerosis case study
Key Sponsor Contact:	Ministry of Health, Welfare and Sport Rosalie Oostrom er.v.oostrom@minvws.nl
Key iMTA contact	institute for Medical Technology Assessment (iMTA B.V.) Erasmus University Rotterdam P.O. Box 1738 3000 DR Rotterdam The Netherlands
Date:	17 May 2021

CONTENTS

INHOUD

SUMMARY

- Artificial intelligence (AI) is a promising technique for using data in healthcare to improve health by optimising treatment options, diagnoses or logistics.
- To substantiate which AI investments are meaningful and which healthcare benefits an AI application can have, and to mobilise parties around the use of AI applications in the healthcare sector, economic evaluations can be used that make the value of AI applications tangible.
- Health Technology Assessment (HTA) offers a framework for carrying out economic evaluations of AI applications. It expresses health benefits in quality-adjusted life years (QALYs). Costs and savings in the healthcare sector and elsewhere are taken into account in this assessment. While HTA offers an objective analysis of the added value of a treatment, a positive HTA is not a guarantee for the uptake of an AI application in practice.
- This method has been applied to a case study of a promising AI application in healthcare: MS sherpa. MS sherpa is a medical device (CE-certified) consisting of a smartphone application and an integrated healthcare provider portal/dashboard. MS sherpa is intended for the monitoring of patients with multiple sclerosis (MS) with the aim of providing patients and their practitioners with personalised insight into the presence and progression of MS-related symptoms.
- The assumed added value of MS sherpa is that this insight into the presence and progression of symptoms will allow a proportion of patients to switch to a more effective drug sooner because the disease activity, in terms of MS relapses or disease progression, is detected early.
- The added value of the MS sherpa application has not yet been proven; hence, this report works with the application's *potential added value* on the basis of assumptions: a so-called early HTA.
- The results of the early HTA show that the MS sherpa application can be a cost-effective addition to standard of care (more health for an acceptable level of higher costs). The application can be cost-effective (better health and fewer costs) if favourable assumptions about effectiveness are made and if benefits outside of healthcare, such as labour productivity, are taken into consideration.
- The early HTA of the use of MS sherpa shows that the application leads to better health and larger healthcare costs. However, when weighing the benefits of better health that are incurred outside of the healthcare sector, such as labour productivity or informal care, MS sherpa can save costs for society: the additional healthcare costs are smaller than the benefits outside of healthcare.
- MS sherpa may have broader effects on patients, positive (e.g., stimulating self-efficacy and shared decision-making) as well as negative (e.g., being confronted with one's illness and being reluctant to share data). Broader effects of this kind are only weighed in an HTA if they concern patients' health-related quality of life.
- Whether the smart use of MS sherpa can contribute to the early detection of disease activity in MS patients and can thus influence treatment decisions has yet to be studied in clinical practice. In the short term, a clinical study will start at the MS Center of Amsterdam UMC to study this and other potential benefits of MS sherpa.
- Those who want to carry out an HTA will need to make decisions regarding, among other things, the perspective applied, the analysis technique and the structure of the data collection.
- Investing in carrying out an HTA has the most added value if it needs to be determined whether the intervention is to be reimbursed by the basic health insurance package.

1. INTRODUCTION

The availability and analysis of data has burgeoned in the computer age. In healthcare, the disciplines of statistics and epidemiology have played an important role in identifying determinants of health. New data analysis techniques closely related to statistics make it possible to answer pressing societal questions and to contribute to public health. Artificial intelligence (AI) is one of the promising examples of these new techniques. In this report, we reflect on the way the Health Technology Assessment (HTA) can be used to assess the added value of AI applications in healthcare. An AI-based decision support system for patients with multiple sclerosis (MS) is elaborated on as a case study of a value assessment of an AI application.

ARTIFICIAL INTELLIGENCE

AI can be an important source of innovation in many societal sectors, including healthcare. More data than ever are being generated and stored and we have become increasingly capable of analysing and using those data, partly due to AI. AI applications have the potential to analyse the effect of previous actions and can work fully or semi-autonomous. As AI applications continually improve themselves on the basis of new data, new insights arise. The central government uses the definition of the European Commission's High-level expert group (see the boxed text).

Definition artificial intelligence (AI)

"Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. [1]*"*

One application of AI in healthcare is to analyse large amounts of public or other health data and to find patterns in order to offer predictions, personalised advice or personalised actions. In this way, AI applications contribute to prevention, self-management, diagnosis, treatment and logistics. There are AI applications, for example, that can read X-rays or do a detailed analysis of cancer cells, speeding up the diagnosis of medical conditions. Healthcare professionals can also use AI as a decision support aid by translating data on patient characteristics, clinical observation and laboratory results such as blood levels into personalised advice. In addition, AI applications can support patients increasingly well to self-manage their health thanks to the numerous possibilities of, for example, portable technology and health applications. These applications of AI demonstrate that AI has the potential to influence the quality of our healthcare and health.

VALUABLE AI FOR HEALTH

The 'Valuable AI for health' programme of the Ministry of Health, Welfare and Sport aims to help convert the potential of AI applications into tangible value and value creation for patients, healthcare providers and citizens. The development of scalable and broadly deployable applications requires investment in many areas: setting up large data sets, the technology, the research itself, legal and policy necessities and cooperation between various parties. These are considerably (structural) investments, in both public and private context. In order to substantiate which investments for the benefit of AI are meaningful, and to mobilise parties around the use of AI in the healthcare sector, it is essential to make the value of AI applications tangible. To this end, the Ministry's 'Valuable AI for health' programme has asked the institute for Medical Technology Assessment (iMTA) of Erasmus University Rotterdam to evaluate the value of an AI application by way applying the HTA methodology to a case study. The value of an AI application in MS is quantified in terms of cost-effectiveness by looking at both costs and benefits in terms of health gains.

This report discusses the possibilities and limitations of the HTA methodology for determining the value of AI, with an emphasis on the economic evaluation. The early HTA of a decision support system for MS patients, MS sherpa, is elaborated here as a case study of the value assessment of an AI application. The report starts out with

an explanation of HTA and early HTA and the methodology for economic evaluations in the Netherlands. Next, more background information is provided about the case study and the methods and results of the early HTA for MS sherpa are described. This is followed by a discussion of the quantitative and qualitative benefits of an AI application such as MS sherpa that are not revealed by an economic evaluation. The report includes two appendices: the 'Road map for HTA research' (Appendix 1) and the 'Step-by-step plan for carrying out an economic evaluation' (Appendix 2).

2. HEALTH TECHNOLOGY ASSESSMENT

In order to ensure that the limited budget we as a society are prepared to spend on healthcare delivers the most possible health to those who need it, we have to gain insight into the existing and potential cost-effectiveness of healthcare interventions. An intervention could be any number of things, from preventive measures, medication, and devices to palliative care. A healthcare intervention's cost-effectiveness can be studied by comparing the costs and benefits of alternative interventions in an economic evaluation. An economic evaluation is an important aspect of an HTA, as it makes an intervention's value tangible and comparable. Other aspects within the HTA framework include systematic evaluations of social, organisational and ethical questions around an intervention. The results of an HTA can inform policy decisions, for example about whether or not to include an intervention in the basic health insurance package. An HTA is usually carried out when a new intervention is ready for use in clinical practice. If an HTA is applied earlier in an intervention's development process and there are no data about its effectiveness yet, or the data are insufficient, we call this an 'early' HTA. An important benefit of an early HTA is that it can provide insight into the intervention's *potential* value, even if this value has not yet been demonstrated in clinical studies.

2.1 THE VALUE OF HTA FOR PROVIDERS OF AI APPLICATIONS

An economic evaluation can provide insight into where the benefits of an intervention lie (in healthcare or elsewhere), what the value of the health gained is and at which points in time costs are incurred and benefits are enjoyed. While these insights are useful, they are also expensive to obtain. Providers of AI applications would therefore be well adviced to ask themselves: is an economic evaluation as part of an HTA useful for me?

Providers of AI applications will at some point be faced with the question of how best to employ their new instrument as an integrated part of healthcare in the Netherlands and possibly abroad. Even if the organisation that has developed the instrument is not commercial in nature, it is very likely that compensation will have to be applied for to cover development costs. If compensation is required, the question arises "how much, and paid by whom?".

If the intervention is to be paid for by the basic health insurance premiums, an economic evaluation is of great importance, as the market mechanisms that normally determine prices are absent from the insured healthcare market. This is why economic evaluations are used to evaluate prices. After all, the evaluation can determine whether the expenses are proportionate to the health benefits gained – an insight that is indispensable for the intervention to be covered by the basic health insurance package. Earlier studies showed, for example, that economic evaluations using decision models can yield valuable information about the impact of the use of diagnostics and prediction models in clinical practice on health outcomes and costs [2, 3, 4]. However, there is no standard admission procedure for AI applications yet, which is why it cannot be stated with certainty that an economic evaluation is a necessary prerequisite for inclusion in the basic health insurance package. Similarly, a positive HTA is not a guarantee that an AI application will be used in practice.

If a company wants to offer an AI application to healthcare providers directly regardless of whether it will be included in the basic health insurance package, an economic evaluation is less self-evident. If healthcare providers pay for the application, a business case will probably suffice. A business case is a more narrow economic evaluation that only considers costs and does not consider (health) benefits. Moreover, it only considers the

costs for the parties involved, and does not incorporate the costs for other actors or sectors. A relevant question in such cases is whether the costs the healthcare institution will have to incur will lead to savings for that same institution, irrespective of the improved health of patients. However, the business case may not properly showcase the value of an AI application if the financial gains fall largely outside of the purchasing healthcare institution, which is often the case when the application is to expect to lead to new patterns of referral by primary healthcare providers, the postponement of admission to a different healthcare institution, a reduction in informal care required or an increase in labour productivity. In such cases – benefits outside of the purchasing healthcare institution – an economic evaluation will remain important to demonstrate the added value of an AI application, for instance towards funders such as health insurers.

In some cases, an AI application is not yet ready to be brought to market, but providers of AI applications are interested in internally identifying the most promising candidate application for further development. In those cases, too, an economic evaluation will provide important insights, such as the total health benefit to be achieved (theoretically) per application and the maximum price that would be proportionate with that health benefit. There are also standard methods for identifying which characteristics of an AI application have a decisive influence on the economic evaluation. Systematically changing all input parameters in the model (such as price or effectiveness) creates a systematic analysis of outcome 'drivers'. The results are often represented visually in a so-called tornado diagram. For the case study, we elaborated an example of this, which will be found in Chapter 3.

In some situations, an HTA can be difficult to conduct, such as in the case of AI applications that affect a large number of healthcare processes at once. The HTA is a *vertical* research methodology that charts the costs and benefits of a specific intervention. Large process changes that have consequences for a very large group of interventions are *horizontal* interventions that are not easily evaluated with an HTA, unless all intended interventions are evaluated separately. This is an expensive, time-consuming undertaking for which there may not be enough data at the time of analysis. If, on the basis of the above, it is decided that an HTA is desirable, the 'Road map for HTA research' in Appendix 1 may be consulted on the steps to follow in designing an HTA.

2.2 ECONOMIC EVALUATION ACCORDING TO THE GUIDELINES OF THE NATIONAL HEALTH CARE INSTITUTE

This paragraph elaborates on the main aspects and steps of an economic evaluation. For a detailed explanation of how to carry out an economic evaluation, which is a core component of an HTA, we refer to Appendix 2. This appendix provides a visual summary of the steps described in the National Health Care Institute's 'Guideline for conducting economic evaluations in healthcare' [5].

An economic evaluation is often carried out using a decision model that combines various sources of information, such as clinical studies and cost data, to allow for predictions about the future costs and benefits of interventions. A model takes a cohort of patients and simulates their entire remaining life span from the moment the intervention is provided, thus predicting the number of life years, the quality of these life years and the associated costs. With this information, it is possible to determine the intervention's additional costs and benefits as compared to an alternative intervention, which makes it possible to determine cost-effectiveness (more on this under the heading 'cost-effectiveness'). A model is underpinned by various sources and assumptions, therefore healthcare providers are consulted when a model is being developed to ensure the model is aligned with the clinical practice in the best way possible. In addition, uncertainty analyses can be carried out to clarify the impact of the assumptions on the results. By following standardised methodologies, a model provides a comparable, predictable assessment of an intervention's added value. One drawback is that a model is always a simplified version of reality; researchers depend on data that are already available and their comparability.

SOCIETAL PERSPECTIVE

An economic evaluation can adopt various perspectives: that of a healthcare institution, that of the insured package or a 'societal perspective'. The choice of perspective depends on the question an economic evaluation is intended to answer, and for whom. The National Health Care Institute uses a societal perspective for its economic evaluations of healthcare interventions. This means that all relevant costs and benefits of the interventions are considered, irrespective of who shoulders the costs or receives the benefits. The costs of an intervention are therefore not limited to costs within the healthcare sector, costs outside the healthcare sector are also considered in an economic evaluation. Examples of the latter are the societal costs of providing informal care and reduced productivity in paid or unpaid jobs due to impaired health. In addition, healthcare costs include not only those costs that result directly from the intervention but also all healthcare costs for the remainder of the patient's life, regardless of whether these are related to the health problem the intervention is aimed at. An analysis from the societal perspective can reveal where the benefits of an intervention are enjoyed. If this is outside of the purchasing healthcare institution, this may explain why a positive societal added value does not result in the intervention's uptake in clinical practice.

QALY

A comparison of a healthcare intervention's costs is easier than a comparison of its benefits, as the costs share the same currency (the euro). To enable a comparison of interventions with different health outcomes, the benefits are expressed in quality-adjusted life years, or QALYs. The QALY combines two components: length of life and quality of life. It is calculated by multiplying life years by a weight for the quality of those life years. This weight can take on several values, with 0 representing a health state as bad as being dead and 1 representing perfect health. One QALY equals one life year in perfect health. According to the guidelines of the National Health Care Institute, the quality-of-life component of the QALY must be measured using the EQ-5D-5L[1] (or, in case of earlier studies, the EQ-5D-3L). The EQ-5D-5L is a generic questionnaire containing five questions about the extent to which a person experiences problems related to mobility, self-care, usual activities, pain and discomfort, and anxiety and depression (Figure 1). As the questionnaire can be administered in all disease areas, the health benefits and cost-effectiveness of interventions may be compared for different conditions. Although the EQ-5D is the most widely accepted standard for health benefits in economic evaluations, it is acknowledged it that the EQ-5D – and consequently the QALY which is calculated with it – is not a perfect measure. The central question is whether the quality-of-life component of a QALY captures all relevant effects of a treatment, as discussed in Chapter 2.2. Other, broader questionnaires are being developed that might mitigate the shortcomings of EQ-5D and the resulting QALYs[6].

EQ-5D-3L	EQ-5D-5L
Level 1: No problems	Level 1: No problems
Level 2: Some problems (MO, SC, UA) or Moderate problems (AD, PD)	Level 2: A few problems
	Level 3: Moderate problems
Level 3: Severe or very severe problems (AD, PD)/unable to … (SC, UA)/bedridden (MO)	Level 4: Severe problems
	Level 5: Extreme problems (MO, PD, AD)/unable to (SC, UA)

MO = mobility, SC = self-care, UA = usual activities, PD = pain/discomfort, AD = anxiety/depression

Figure 1. Dimensions of the EQ-5D (source: EuroQol [7])

[1]The EQ-5D-5L can be licenced via www.euroqol.org

COST-EFFECTIVENESS

The outcome of an economic evaluation is presented as the amount of money spent to deliver one additional QALY with the new intervention when compared to the alternative intervention. This outcome is called the incremental cost-effectiveness ratio (ICER). Incremental refers to the comparison of the new intervention with an alternative treatment, e.g. standard of care. The ICER is calculated by taking the difference in cost between the two alternatives and dividing it by the difference in health benefits (see the formula below).

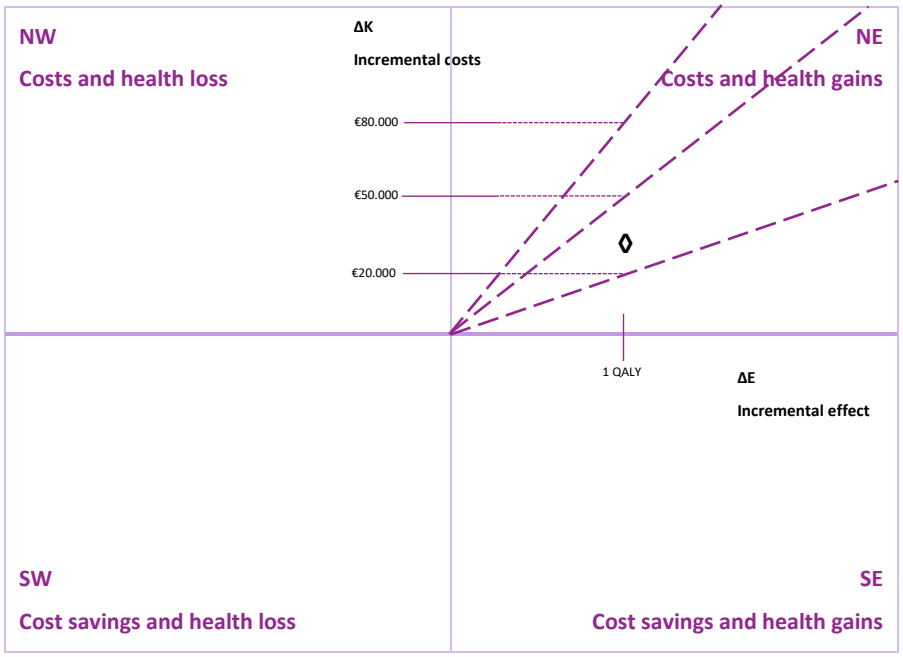
Incremental cost effectiveness ratio

$$= \frac{costs\ of\ new\ intervention - costs\ of\ standard\ care}{health\ benefit\ of\ new\ intervention - health\ benefit\ of\ standard\ care} = \text{€ per QALY}$$

The outcome of an economic evaluation is compared with the reference value (threshold) for the maximum cost we are prepared to pay for an additional QALY: if the ICER is below the reference value, the new intervention is considered to be cost-effective. This reference value is set by the National Health Care Institute and depends on the disease burden caused by the health problem. The disease burden is expressed as a number between 0 and 1 and represents the difference between the life expectancy of patients with the health problem and the life expectancy of the general population that does not suffer this health problem. The reference value is larger for health problems with a higher disease burden: the reference value for a disease burden from 0.10 to 0.40 is €20,000 per QALY (e.g., migraine); the value for a burden from 0.41 to 0.70 is €50,000 per QALY (e.g., MS); and the value for a burden from 0.71 to 1.00 is €80,000 per QALY (e.g., an aggressive oncological disorder). In general, the highest category of disease burden mainly includes disorders that significantly reduce life expectancy. The report 'Disease burden in practice' by the National Health Care Institute uses a number of fictional case studies to explain how the disease burden is calculated [8]. The disease burden is normally calculated using the same model as the one used to carry out the cost-effectiveness analyses. Chapter 3.3 details the calculation of the disease burden of MS.

INTERPRETING THE OUTCOMES OF AN ECONOMIC EVALUATION

Figure 2 shows a cost-effectiveness plane that plots the incremental costs ΔK on the y-axis against the incremental effect ΔE on the x-axis. The various scenarios for outcomes of cost-effectiveness analyses are represented in the Figure. If a new intervention both saves cost and results in health benefits, the ICER is situated in the south-eastern (SE) quadrant. This is the most favourable scenario: in this case, the intervention is always cost-effective, and the intervention is referred to as being 'dominant'. However, in many cases, a new intervention results in better health against higher costs, causing the ICER to fall in the north-western (NW) quadrant. In these scenario's, the reference value will be consulted, i.e. how much we are prepared to pay for one QALY. In the north-eastern (NE) quadrant in the Figure, lines have been drawn for the reference values €20,000, €50,000 and €80,000 per QALY. Depending on the asking price for the intervention in case of an observed effect, the ICER will be situated higher or lower in the quadrant. If the ICER is below the disease burden's applicable reference value, the treatment is cost-effective. If the ICER is above the reference value, the treatment is not cost-effective. The reference values play no role in the other quadrants: in those cases, the intervention is always cost-effective (SE), the intervention is never cost-effective (NW) or one would have to sacrifice health in return for cost savings (SW).



◊: fictional treatment that yields one additional QALY against €30,000 in additional costs.

Figure 2. Possible outcomes of a cost-effectiveness analysis

The '◊' in the Figure above shows the cost-effectiveness of a fictional treatment: the treatment yields one additional QALY against an additional cost of €30,000. This treatment is cost-effective if the reference value for the health problem is €50,000 or €80,000, but is not cost-effective if the reference value of €20,000 is applied.

FEASIBILITY OF THE HTA METHODOLOGY IN AI

The HTA methodology has been applied in a wide array of healthcare interventions, including genetic tests, implants, pharmaceutical products, medical devices and e-health. In the field of AI, too, various HTA reports have been described [9]. Although the existing HTA methodologies are often well-suited, there are situations in which the standard analyses are less appropriate for demonstrating an intervention's value. For example, the traditional methodologies are less suitable for studies in very small patient groups, such as in the case of rare conditions or small sub-populations (personalised medicine) and in interventions with a fast continuous development process, such as AI applications with a self-learning element that remains active after being introduced to the market. Such situations may require small or larger adjustments to the analysis techniques. To offer more tools to researchers for demonstrating the value of healthcare interventions, the National Health Care Institute and ZonMw will in the coming years be working on a learning guide for HTA methodologies [10]. The guide, which can be refined on an ongoing basis, will offer an overview of HTA methodologies and explain why and in which situations specific research designs, analytical methods and outcome measures are suitable. One consideration in the programme of the National Health Care Institute and ZonMw is the need for additional outcome measures when the traditional ones, such as the EQ-5D, appear unsuited to the evaluation of interventions that seek to improve well-being. In the next section, we will go into more detail on the discussion about including broader benefits in the HTA framework.

2.3 BROADER EFFECTS OF HEALTHCARE INTERVENTIONS

The main goal of an economic evaluation within an HTA framework is to identify, measure, value and compare costs and benefits of the healthcare interventions being assessed [11]. Based on the rationale that all healthcare interventions funded out of the healthcare budget should improve health, economic evaluations express benefits in terms of health benefits, or QALYs, as described in the previous chapter. Quality of life, as operationalised in a QALY, is also called 'health-related' quality of life.

However, health interventions often have more effects than the improvement of health-related quality of life alone. Recent years have seen increasing attention for the valuation of the 'full' added value of interventions, insofar as this can be known. The rationale behind this alternative approach is that a broader outcome measure may be a better match to the expectations that patients and citizens have of healthcare interventions. For example, research shows that the needs and expectations around healthcare interventions go beyond health benefits alone: people also attach value to aspects such as satisfaction with the healthcare process, effects outside of health dimensions and the opportunity to participate in society on an equal footing [12]. The QALY only captures such aspects if they influence one or more of the five dimensions of the EQ-5D (mobility, self-care, usual activities, pain/discomfort and anxiety/depression, see Figure 1) – in other words, if the effects of those aspects are sufficiently strong to influence health-related quality of life. A study of the effectiveness of self-management interventions for MS patients, for example, found that four out of five interventions had an effect on depression, anxiety or general health-related quality of life [13].

The growing interest in the broader benefits of healthcare interventions has led to the exploration of a broader outcome measure in terms of general quality of life, also called 'well-being'. For instance, Statistics Netherlands (CBS) describes well-being as the degree to which people are satisfied with their life. CBS looks at general satisfaction with life, satisfaction with various aspects of life and the degree of personal control people experience with regard to their own life [14]. This kind of broad outcome measure is referred to as the well-being-adjusted life year (WALY). An example of a questionnaire for measuring well-being is the ICECAP, which measures the degree to which a person experiences stability, attachment, autonomy, progress and pleasure [15]. A challenge in implementing well-being measures in economic evaluations is that well-being can develop independently from health as people adapt to deteriorating circumstances (coping), which can make it seem like improving their health is not worthwhile [16, 8].

There is currently no broad consensus about the extent to which broader benefits should influence HTA decision-making. Bringing dimensions other than health into the equation can displace health under a constrained healthcare budget. After all, including broader benefits, for instance through the WALY, implies a willingness to trade health in terms of health-related quality of life or longevity for other aspects of value, as the example below illustrates. In addition, a recalibration is required to determine whether there is a willingness to use the healthcare budget to fund interventions that mostly have benefits other than health.

Including broader benefits in an economic evaluation

Table 1 illustrates the consequences of including broader benefits in decision-making. The health benefit in terms of QALYs is greater for treatment 2. However, treatment 1 offers the additional value of 'enhancing self-reliance', whereas treatment 2 does not improve this. If these broader benefits are given a large weight and a decision is made to reimburse treatment 1, with the amount of €80,000 two life years in perfect health, QALYs, can be generated. Treatment 2, however, could have yielded two-and-a-half QALYs for the same amount, albeit without the additional benefits beyond health-related quality of life.

Table 1. Illustrative example of the consequences of weighing broader benefits as part of an economic evaluation

Treatment	Difference in costs*	Difference in QALYs*	Broader benefits	Costs per QALY
Treatment 1	€80,000	2	Enhancing self-reliance	€40,000
Treatment 2	€80,000	2.5	None	€32,000

** Compared with standard of care*

To be able to compare AI applications with other healthcare interventions, the standard analysis for economic evaluations should be followed to the extent possible, in accordance with the guidelines of the National Health Care Institute. It is important, however, to determine whether there are any benefits in dimensions other than health dimensions. This enables consideration of these benefits as contextual factors in the HTA decision-making process. An important addition to the standard analysis thus is the AI application's objective: often, the objective does not primarily comprise health benefits in terms of quality of life as measured with the EQ-5D-5L or longevity. Rather, targets are set that are aimed primarily at general well-being, such as self-reliance, learning to cope better with a chronic condition, or autonomy. In such cases, the National Health Care Institute recommends using a validated, weighted well-being questionnaire such as the ICECAP in addition to the EQ-5D-5L [8, 15]. Beside the value the intervention can have for the patient, an analysis of broader benefits can also include values for society as a whole. For this, CBS uses values such as the environment, security, housing and the functioning of society [14]. Other values that can be considered include equal opportunities, creating future opportunities (option value) and value to science (scientific spillover) [17].

As this is an early HTA, no data are available yet about the effects of MS sherpa, either within health dimensions or beyond. Nonetheless, to offer as complete a picture as possible of the value of this AI application, Chapter 3.4 will address the potential broader effects that MS sherpa may have based on assumptions and exploratory studies.

3. CASE STUDY

A Health Technology Assessment (HTA) is a useful framework for quantifying the value of AI applications, in terms of both health benefits and possible cost savings in the healthcare sector and elsewhere. As such, an HTA can contribute to the justification in the political and public debate regarding which structural investments in AI are meaningful. It can also help mobilise parties such as health insurers and healthcare institutions regarding the use of AI in the healthcare sector. At the recommendation of specialists, the AI programme has decided to focus the intended HTA on a single promising project where societal impact is concerned and to elaborate this in great detail. To illustrate what an economic evaluation of an AI application entails, this chapter describes the early HTA of an AI application for MS patients: **MS sherpa**. This case study offers a good opportunity to translate the potential value of an AI application in healthcare into what it can deliver in terms of cost-effectiveness.

Early HTA

We speak of a 'potential value' and 'early HTA' because the value of MS sherpa has not yet been demonstrated, so at this point, we can only speak of the product of the product. A quantitative assessment will need to demonstrate the effectiveness of MS sherpa in clinical practice.

3.1 BACKGROUND INFORMATION ON MULTIPLE SCLEROSIS

MULTIPLE SCLEROSIS

MS is a chronic disease, characterised by damage to the protective and insulating layer around the cranial nerves, spinal cord and optic nerves. People with MS may experience, among others, walking difficulty, vision problems, and sensory issues. Patients experience unpredictable exacerbation of their illness, often resulting in hospital admissions and significantly reduced functionality. About 17,000 people in the Netherlands have MS; worldwide between 2 and 2.5 million people are affected by the disease[18]. Unfortunately, there are no treatments that can cure MS.

CURRENT TREATMENT OF MULTIPLE SCLEROSIS

Current treatment of MS follows the guideline of the Netherlands Society of Neurology. There are 12 drugs that can slow down disease progression or counter relapses in which patients suffer from severe MS symptoms. Annual checks at the neurologist usually involve an MRI scan, and often cognitive tests and walking tests are also completed. The therapy and medication can be adjusted on the basis of this annual check and the experiences shared by the patient.

MS drugs are divided into three lines (Table 2). Generally, patients start with a drug from the first line, after which they may switch to a drug from the second or even the third line. This treatment sequence has been established because neurologists have the most experience with first-line drugs and because their side effects are less serious than those of some second-line or third-line drugs. The effectiveness and side effects of drugs, however, differ per patient. As a result, it can take a long time before the optimal treatment is found. Over the course of their lifetime, patient therefore generally use several drugs to treat MS. Patients can switch to a different drug if their current drug does not work well, causes side effects, or if patients prefer a different mode of administration. Some patients stop taking medication, especially if they have not had any relapses or relapses in a long time. The 2019 MS Drug Monitor (*Monitor MS-Geneesmiddelen 2019*) of the National Health Care Institute showed that 30 to 40% of users of MS drugs switched to a different drug within five years [19].

Table 2. Available drugs for multiple sclerosis in the Netherlands.

Line 1	Line 2	Line 3
Dimethyl fumarate 240 mg PO (DIF)	Cladribine 3.5 mg PO (CLA3.5)	Alemtuzumab 12 mg IV (ALE)
Glatiramer 20 mg SC (GLA20)	Fingolimod 0.5 mg PO (FIN)	
Interferon ß-1a 30 mcg IM (IFNa30)	Natalizumab 300 mg IV (NAT)	
Interferon ß-1a 44 mcg SC (IFNa44)	Ocrelizumab 600 mg IV (OCR)	
Interferon ß-1b 250 mcg SC (IFNb250)		
PEG Interferon ß-1a 125 mcg SC (PEG)		
Teriflunomide 14 mg PO (TER14)		

SELF-MONITORING OF MULTIPLE SCLEROSIS

Worldwide, there are about ten smartphone applications for MS patients that are dedicated to providing information about MS, setting up 'communities', registering health and symptoms on a daily basis, recording the use of medication or sharing data with healthcare providers [20]. People with MS generally seem to respond positively to the use of such health applications [21, 22]. An online survey in the Netherlands of 143 people with MS found that >90% of patients have a smartphone and use it regularly [23]. Over 50% of people with MS are already monitoring themselves (e.g., by keeping a diary), and over 60% of people with MS would like to monitor themselves with a smartphone application in order to gain greater insight into their disease. Patients especially noted the added value of gaining insight into the disease progression and the cause of symptoms (43%). A smaller percentage sees gaining insight to take action as the main goal of self-monitoring using a smartphone application (11%), followed by gaining an insight for the physician (6%). Almost half of the patients noted added value in the application with regard to personal disease progression. If an application can measure MS progression, 46% of people with MS are willing to pay for the tool (most of whom want to spend a maximum of five euros) without having further details about the application. Patients in a lower age category appear to use applications more often, have a greater desire to self-monitor and see more added value in a smartphone application that monitors disease progression. The older patients are, the more problematic they seem to find disclosing personal information.

Studies have shown that the use of a health application by people with MS may help them feel more independent, may make it easier to access healthcare and may save time [21]. Other possible benefits that health applications may achieve include easier access to information about MS, communication with healthcare providers, value for monitoring MS, more efficient management of administrative MS tasks, and the possibility of communicating with other people who have MS. Complaints heard from people with MS about health applications often revolve around problems with the application's performance and the limited possibilities it offers [24]. For example, while existing applications often offer the possibility to register fatigue symptoms, this is in many cases not linked to actions or advice to help patients cope better with fatigue [25]. A reminder to take medication can be another valuable function of an MS application and can improve compliance with therapy [26]. In spite of the large and growing availability of MS applications, the existing applications do not yet appear to be able to assist people with all dimensions of MS self-management programmes [20].

3.2 MS SHERPA

The unpredictable nature and progression of MS makes life with the disease challenging. In view of this, software company Orikami B.V. partnered with the Nationaal MS Fonds, Stichting MS Research, Amsterdam UMC, Radboud University and the MS4 Research Institute to develop an application that monitors disease activity in MS patients in order to develop greater insight into the progression of the disease. This application is called MS sherpa. To develop and research the application, Orikami obtained funding from such sources as the Nationaal MS Fonds, the Dutch Research Council, investment fund Healthy.capital, the RedMedTech fund, the European Regional Development Fund, healthholland and qredits. Orikami has also won several prizes, including the regional Healthcare Innovation Prize (*Zorginnovatieprijs*) in 2018.

The Figure below illustrates the potential added value of more frequent monitoring of MS patients. The dots stand for all of the days in a patient's year; the colour of each dot indicates how severe the symptoms were on that day. On average, MS patients see their neurologist once a year. The tests that are completed on that day provide a snapshot of the severity of their symptoms. Also, in their talk with the neurologist, patients will only recall and discuss a few days' worth of symptoms. With the aid of the MS sherpa application, more frequent measurements take place, which can provide a more complete picture of the disease progression over time in the talk with the neurologist.

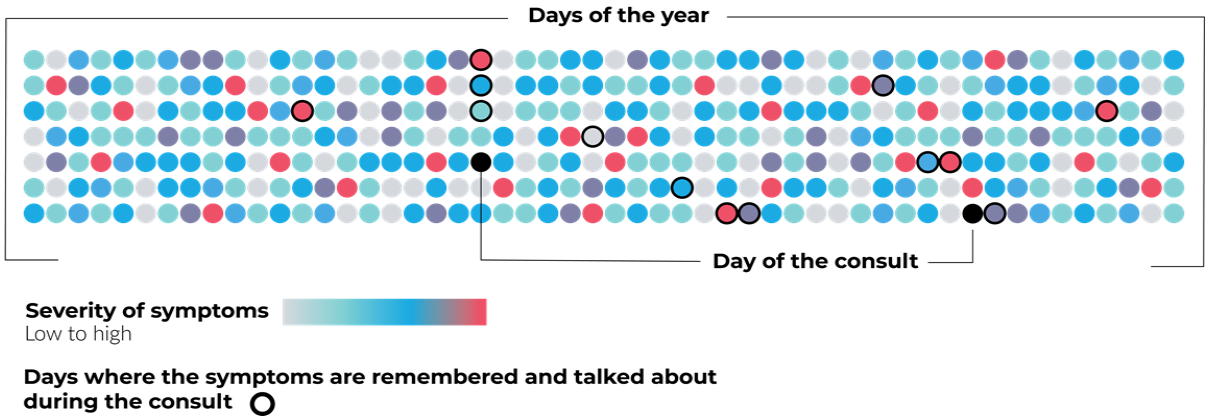


Figure 3. Representation of the progression of symptoms over a year [source: Orikami, based on roche.com/about/priorities/personalised_healthcare/digital-biomarkers.htm]

MS sherpa is a medical device (CE certified) consisting of a smartphone application and an integrated healthcare provider portal/dashboard. MS sherpa aims to help monitor MS patients to provide them and their therapists with a personalised insight into the presence and progression of MS-related symptoms. Two clinically validated tests were developed for the application that make use of digital biomarkers: the Symbol Digit Modalities Test (SDMT) of cognitive deterioration and the 2-Minute Walk Test (2MWT) of reduced walking speed. These tests, which were developed using self-learning software, have been digitalised and validated for home use by patients

[27, 28]. The walking test and the cognitive test are to be completed at least once a month. The data registered by patients at home are analysed and plotted over time, showing how the disease behaves and develops. In addition, patients can use the application to complete questionnaires about quality of life (e.g., regarding mood and energy) and to keep a diary of observations, providing insight into gradual changes in the disease manifestation and to identify possible links with lifestyle choices. Currently, MS sherpa is only used in research context; it will be released on the market in the short term. In the early HTA, an estimated product price of €480 per patient per year will be considered.

The expectation is that the insights MS sherpa provides for a proportion of patients may result in an earlier switch to an MS dug from the next line, as the more frequent monitoring allows for earlier detection of signs of disease activity (i.e., the disease progression of MS relapses). In general, these drugs are more effective than drugs from previous lines, increasing the chance of preventing disease progression or MS relapses. The prevention of disease progression or MS relapses will improve quality of life and may lead to costs savings in terms of MS-related healthcare costs, informal care, or the lost productivity from paid or unpaid jobs.

Various studies are being carried out before the application will be implemented more broadly in clinical practice. In the course of 2021, the Multiple Sclerosis self-monitoring & self-management (MSSM) study will begin. About 250 MS patients will participate in this scientific study from the MS Center of Amsterdam UMC, half of whom will use the MS sherpa application [29]. In this study the application's effect on self-efficacy will be evaluated. Self-efficacy can be described as the confidence people have in their own ability to complete tasks successfully and to influence matters they find important. The MSSM study will also look at the effects of MS sherpa on the healthcare process (such as treatment decisions, contact moments with the physician, the use of clinical tests, referrals to specialists) and at its effects on MS relapses and disease progression. The study will furthermore look at quality of life, cost-effectiveness, patient preferences and the broader usability of the data. Another study, DOT-MS, investigates whether it is safe to stop taking medicaiton when MS has been stable for many years; MS sherpa will be used to monitor whether there is any deterioration [30, 31]. The DOT-MS study will also investigate whether biomarkers can be used to predict for which patients it is most beneficial to be taken off medication.

The added value of an AI application depends on its appropriate use (by patients as well as healthcare providers) to useful purposes. As the added value of MS sherpa has not yet been proven in clinical studies, assumptions were formulated for this early HTA about how the application is used and for which purpose. During an initial analysis of the benefits of MS sherpa, it was decided that the main focus in the early HTA will be on the possibility of detecting MS disease activity early with MS sherpa, allowing for an early switch to a different MS drug – even before the disease activity has actually occurred. Chapter 3.4 will discuss the broader effects of the use of MS sherpa.

3.3 EARLY HTA OF MS SHERPA

The early HTA of MS sherpa is carried out using a decision model in which information from the literature and from talks with MS neurologists are integrated. With this model, the potential cost-effectiveness of MS sherpa was calculated and several scenarios with varying effectiveness of the AI application were explored.

In this chapter, we first describe the NICER MS model, followed by a description of how the potential effect of MS sherpa is included in this model. Next, we elaborate on the information in the literature about the effectiveness of MS drugs, costs and quality of life. Lastly, the results of the early HTA are reported and we discuss how these can be interpreted.

MS MODEL

The analyses were carried out through additional programming of the NICER MS model, a complex decision model for MS costs and effects written in programming language 'R', developed by Huygens and Versteegh of iMTA. This model simulates the entire lifetime of MS patients on the basis of MS relapses and disease

progression. Disease progression was expressed using the Extended Disability Status Scale (EDSS). The EDSS is a scale that measures the level of disability of a person with MS. It runs from 0 to 10, where 0 represents the absence of problems and 10 represents death caused by MS. At EDSS score 7 and above, an MS patient is wheelchair dependent.

The speed of MS progression and the number of relapses are influenced by the choice of MS drugs. The model also considers the side effects of MS drugs, the possibility of no longer taking drugs, patients' quality of life, and risk of death. In addition, it considers all costs stemming from the treatment of MS: the cost of MS drugs, other healthcare, informal care and loss of productivity. Disease progression, a relapse or a severe side effect is assigned extra costs and a reduction in quality of life. This allows for the calculation of the costs and effects of various sequences of MS drugs.

Figure 4 is a graphical representation of the model. It illustrates the possible treatment pathways an MS patient can go through in the NICER MS model. The potential effect of using MS sherpa (an earlier switch to a different drug) has been added to this. A detailed explanation of these treatment pathways is provided below the Figure.

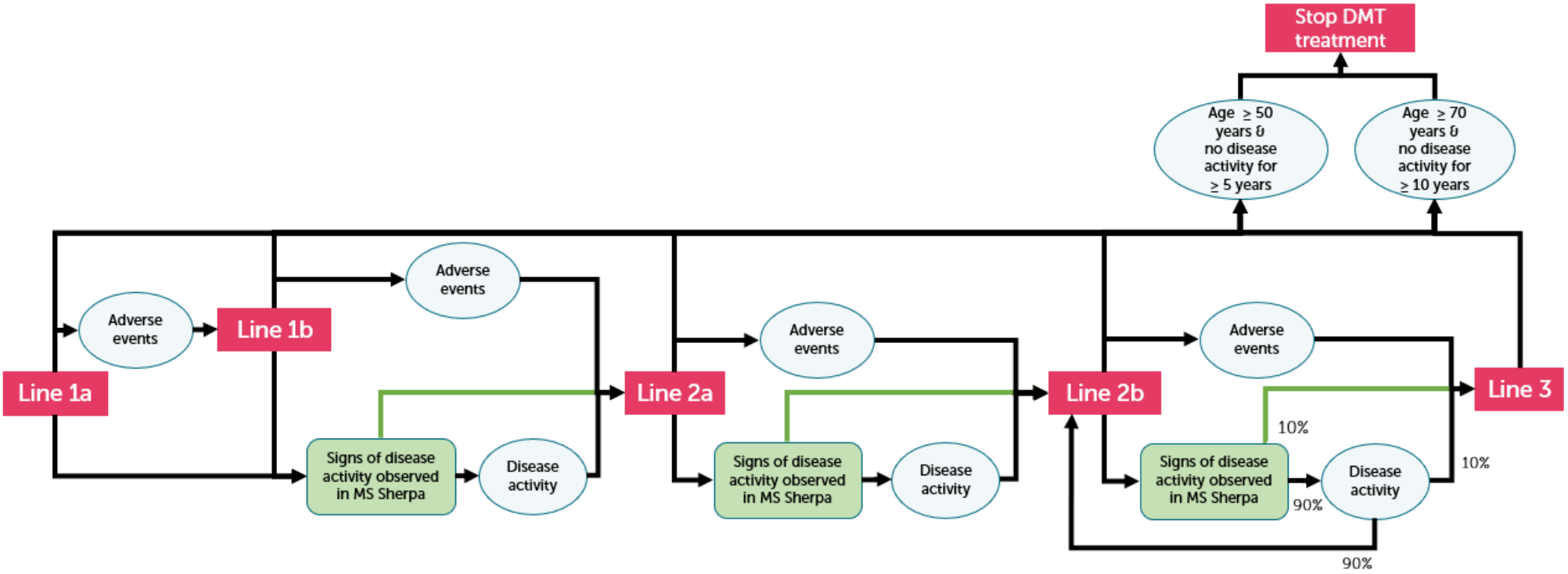


Figure 4. Schematic representation of the MS model.

Patients start out with an MS drug from the first line (Line 1a). If they experience adverse effects from this drug, they switch to a different first-line MS medicine (Line 1b). In case of disease activity during the use of a first-line MS drug (Line 1a or Line 1b), patients switch to a second-line MS drug (Line 2a). If patients experience side effects of the second first-line MS drug (Line 1b), they also switch to a second-line MS drug (Line 2a). Next, patients switch to an MS drug from the second line (Line 2b) in case of disease activity. Lastly, if patients experience side effects or disease activity, they can switch from a second MS drug from the second line (Line 2b) to an MS drug from the third line. However, in view of the reluctance of neurologists to prescribe the only third-line MS drug currently available (due to the risk of side effects), only 10% of patients switch from the second MS drug from the second line to the third line in case of disease activity. In each line, patients can also stop taking MS medication (Stop DMT treatment) if they are aged 50 or over and have not had any disease activity for at least 5 years, or if they are aged 70 or over and have not had any disease activity for at least 10 years. The green blocks in the Figure (Signs of disease activity observed in MS sherpa) reflect the possible impact of MS sherpa on the treatment pathways: based on the insights from the MS sherpa data on the progression of disease activity, it may be decided in consultation with the physician to switch to a drug from the next treatment line at an earlier stage.

**In the NICER MS model, disease activity is defined as: a severe relapse (18.7% of all relapses); a relapse in two consecutive years during the use of first-line MS medication; a relapse in two consecutive years during the use of second-line MS medication and lesions visible on MRI scans; or disease progression (defined as a reduction in the EDSS score) and a relapse.*

INPUT VALUES FOR THE MODEL

CLINICAL EFFECTS

If an AI application is to support decisions about follow-up treatments, a solid scientific basis for the effectiveness of these treatments is required [32]. That is, the actual impact of MS sherpa on clinical outcomes is driven by the accuracy of MS sherpa in detecting deterioration *in combination with* the effectiveness of the treatment the patient can switch to. In the NICER model, the effectiveness of the pharmacological treatment of MS for the calculation of cost-effectiveness is based on clinical studies. Clinical studies usually compare the effect of a new drug with a placebo or a single existing drug, but for most drugs, there are no clinical studies that compare the various available drugs with each other. In such cases, a network meta-analysis can be carried out, using existing one-on-one comparisons to determine the differences in effectiveness between various MS drugs. The results are displayed in Figure 5, which shows an overlap in the effectiveness of MS drugs, while second-line and third-line MS drugs seem more effective on average relative to a placebo in preventing disease progression and relapses than first-line MS drugs relative to a placebo.

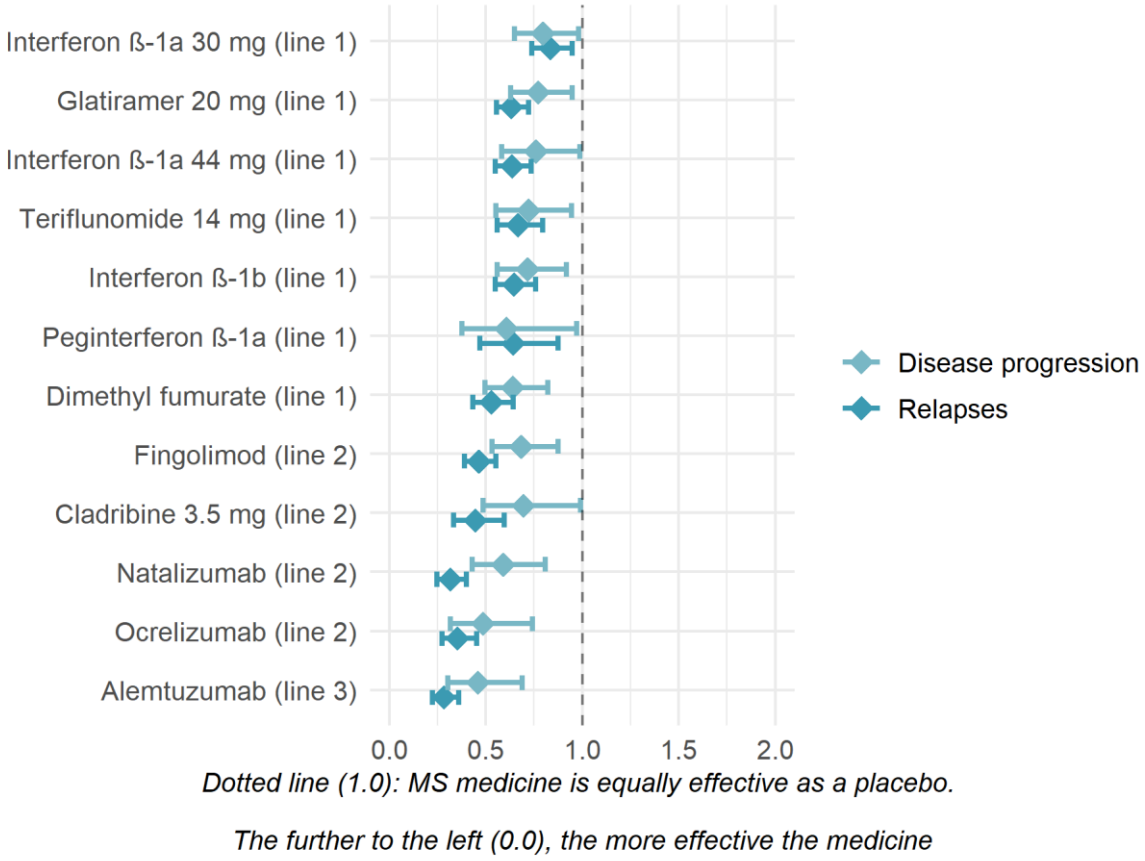


Figure 5. Results of the network meta-analysis of the effectiveness of MS drugs.

COSTS OF STANDARD OF CARE FOR MULTIPLE SCLEROSIS

MS-related costs have increased in recent years, as new, more expensive drugs have been brought onto the market. In 2019, about 170 million euros were spent in the Netherlands on MS drugs [33]. The costs of MS drugs range from 8,000 to 38,000 euros per patient per year. Table 3 shows the costs of the first five years of treatment with MS drugs. These costs are based on publicly available list prices and dosages from the Pharmacotherapeutic Compass (*Farmacotherapeutisch Kompas*) of the National Health Care Institute, combined with the costs of hospital admissions for the administration of these drugs on the basis of information from neurologists [34].

Table 3. Annual costs of MS drugs* during the first five years of use

MS drug	Year 1	Year 2	Year 3	Year 4	Year 5
ALE	€ 40,625	€ 24,375			
CLA3.5	€ 30,667	€ 30,667			
DIF	€ 15,105	€ 15,242	€ 15,242	€ 15,242	€ 15,242
FIN	€ 22,840	€ 22,345	€ 22,345	€ 22,345	€ 22,345
GLA20	€ 8,749	€ 8,749	€ 8,749	€ 8,749	€ 8,749
NAT	€ 26,292	€ 26,292	€ 26,292	€ 26,292	€ 26,292
OCR	€ 25,631	€ 25,136	€ 25,136	€ 25,135	€ 25,136
PEG	€ 12,686	€ 12,806	€ 12,806	€ 12,806	€ 12,806
TER14	€ 12,045	€ 12,045	€ 12,045	€ 12,045	€ 12,045
IFNa30	€ 9,150	€9,421	€ 9,421	€ 9,421	€ 9,421
IFNa44	€ 10,509	€ 11,062	€ 11,062	€ 11,062	€ 11,062
IFNb250	€8,224	€8,432	€ 8,432	€ 8,432	€ 8,432

** Assumptions of hospital admissions for the administration of drugs based on estimates by MS neurologists: Alemtuzumab (ALE) involves a five-day hospitalisation in the first year and a three-day hospitalisation in the second year. Fingolimod (FIN) involves a one-day hospitalisation in the first year. Natalizumab (NAT) involves a one-day hospitalisation every four weeks. Ocrelizumab (OCR) involves two induction days and one maintenance day in the first year, as well as two annual hospitalisation days after the first year.*

The costs of healthcare, informal care and loss of productivity in the NICER MS model are based on information from a survey study. Uitdehaag et al. administered this survey to 382 Dutch MS patients, asking questions about the degree of disability (measured with the EDSS), quality of life, symptoms, use of healthcare and the impact of MS on productivity [35].

Based on the information about resource use, Uitdehaag et al. calculated the average annual healthcare costs of three groups of MS patients: patients with mild MS (EDSS 0-3), patients with moderate MS (EDSS 4-6.5) and patients with severe MS (EDSS 7-9) [35]. The costs are higher for patients with a more severe form of the disease. In the NICER MS model, we added together the healthcare costs excluding the costs of the MS drug given in Table 3. This sum was corrected for inflation to reflect the price level of 2019. Table 4 gives the healthcare costs per EDSS cluster, consisting of the costs of hospital admissions, outpatient treatments, diagnostic examinations and other drugs. These are the costs of the total healthcare consumption by MS patients, irrespective of whether it is (directly) related to MS.

The survey by Uitdehaag et al. also asked about the use of informal care by MS patients [35]. Three quarters (75%) of patients with severe MS use informal care, compared to 58% of patients with moderate MS and 22% of patients with mild MS. The amount of informal care receive is larger for patients with severe MS (an average of 20 days per month and 3.9 hours per day), compared to an average of 16 days per month and 2.8 hours per day for patients with moderate MS and an average of 8.9 days per month and 2.9 hours per day for patients with mild MS. The proportion of MS patients who access informal care was multiplied by the average number of days per month, the average number of hours of informal care per day and the unit price of informal care in 2019, which stood at €14.74 euros per hour, as per the Cost Manual (*Kostenhandleiding*) of the National Health Care Institute [36]. Table 4 shows the average annual costs per EDSS cluster.

The NICER MS model includes the cost of lost productivity due to MS in cases where patients with a paid job had an MS relapse or lost their job due to disease progression. According to Uitdehaag et al., the share of MS patients with a paid job was 51.5% among patients with mild MS and 10.7% among patients with severe MS [35]. Based on linear interpolation between these two numbers, assuming an average difference of 5 EDSS classes between these two groups, the proportion of patients who lose their job because of MS increases by 8.2% per EDSS class.

The productivity costs of an MS relapse were calculated by taking the average hourly wage as reported by Statistics Netherlands corrected for the gender distribution of MS patients (€33.22), and multiplying this by the average number of hours worked per week (26.5 hours) by MS patients as reported by Egmond et al. [37]. Based on an estimate by MS neurologists, MS patients are unable to work for six weeks after a relapse. This results in a productivity cost of €5,283 occasioned by an MS relapse.

According to the National Health Care Institute's 'Guideline for economic evaluations in healthcare', the cost of productivity loss as a result of sick leave from paid work is calculated using the friction cost method [5]. This method assumes that ill employees can be replaced in case of long-term absence. That is, productivity losses occur mainly during the period the employer requires to replace an ill employee: the friction period. This friction period depends on the number of unfilled and filled vacancies in a calendar year; in 2019, it stood at 15.9 weeks. The costs of this friction period for MS patients amounted to €13,977. For this early HTA, the costs of the use of MS sherpa were estimated to be 480 euros per patient per year.

Table 4. Annual costs per disease severity level (mild, moderate, severe)

Cost input	Mild MS: EDSS 0-3	Moderate MS: EDSS 4-6	Severe MS: EDSS 7-9
Healthcare costs per year	€4,094	€8,033	€10,146
Private care costs per year	€1,009	€4,580	€10,349
Percentage of MS patients with a paid job	51.5%	26.7%	10.4%
Productivity costs of an MS relapse	€5,283		
Productivity costs of loss of job	€13,977		
MS sherpa application costs per year	€480		

QUALITY OF LIFE OF MULTIPLE SCLEROSIS PATIENTS

The quality of life of MS patients in the NICER MS model is also based on results from the survey study by Uitdehaag et al. [35]. This survey used the EQ-5D-3L questionnaire to measure quality of life. The results, displayed in Table 5, show that the quality of life of patients with mild MS is between 0.637 and 0.930, that of patients with moderate MS is between 0.651 and 0.696 and that of patients with severe MS is between 0.041 and 0.528.

Table 5. Quality of life (utility) of MS patients per EDSS score.

Severity of disease	EDSS score	Quality of life (EQ-5D-3L utility)
Mild MS	**0**	0.930
	1	0.858
	2	0.782
	3	0.673
Moderate MS	**4**	0.696
	5	0.690
	6	0.651
Severe MS	**7**	0.528
	8	0.359
	9	0.041

COSTS AND EFFECTS OF MS SHERPA IN THE MODEL

We assume that all patients with an EDSS score under 7 use the MS sherpa app. An EDSS score of 7 or higher means patients are wheelchair dependent, and as a consequence, the MS sherpa app, of which the walking test is an important part, is not suitable for them. The use of the MS sherpa application is expected to cost €480 euros per patient per year. As described, one of the important potential values of MS sherpa is obtaining early insight into oncoming disease activity (in terms of MS relapses or disease progression), allowing for a switch to a different MS drug before the deterioration actually occurs. This was implemented in the NICER MS model as follows. First, the NICER MS model predicts the course of the disease without the use of MS sherpa. For a certain percentage of patients who, according to the NICER MS model, have disease activity in standard of care, we 'retroactively' change the MS drug on the assumption that the use of the app will have this effect in a proportion of the patients with an EDSS <7. As the new MS drug is generally more effective, the risk of actual disease activity is smaller. The percentage of patients who switch to different medication as a result of using MS sherpa is still unknown. It is also possible that the effectiveness of MS sherpa will increase with future versions of the application, in which the algorithm can be attuned to the growing database with the possible result of more accurate insights. Hence, this report tested various assumptions about effectiveness in terms of the proportion of patients who switches to a different drug because of MS sherpa: 5%, 10%, 15% and 20%. We assume here that, if patients switch to a different drug sooner because of the use of MS sherpa, disease progression would actually have occurred if they had not switched to this drug. For patients who do not make an early switch as a result of using the app, we assume that MS sherpa is unable to detect disease activity early.

As the CE-certified product does not contain a self-learning algorithm (i.e., while the algorithm was developed with the help of AI, it does not automatically adapt to the growing database generated by its use by many patients), the analysis assumes a stable effectiveness rate. Given that users of the application will collect more and more data, it is possible that the algorithm will at some point be revised on the basis of a larger database and that this revised algorithm will be used for a new version of the application, with the possible result of more accurate insights (and greater effectiveness).

ANALYSES

The above described information about the costs and effects of standard of care and the costs and various potential effects of MS sherpa has been used for the NICER MS model. The potential cost-effectiveness of MS sherpa was calculated by balancing the costs and effects of MS sherpa against the costs and effects of standard of care – that is, the situation where MS sherpa is not used. This calculation was conducted for the various scenarios of effectiveness of the use of MS sherpa.

In addition, a 'univariate sensitivity analysis' was carried out to obtain an understanding of which parameters have the largest impact on the cost-effectiveness of the AI application. The term univariate means that, in this sensitivity analysis, the value of each individual input parameter of the cost-effectiveness model is adapted, one at a time, to a lower and higher value than that assumed in the base-case analysis. Cost-effectiveness results are reported for each change in the value of an individual input parameter. The results of this analysis are presented in a so-called tornado diagram. They may help providers of AI applications select a focus for future improvements to their applications; these providers would do best to concentrate their attention and available budget on improvements to those elements that have most impact on cost-effectiveness. In this analysis, cost-effectiveness is expressed as net health benefit. Net health benefit is calculated with the below formula. The costs are converted into the health effects that can be achieved against these costs, by dividing the total cost of an intervention by the cost-effectiveness threshold (in this case, €50,000 per QALY). This is then subtracted from the total QALYs that can be achieved with the intervention. The greater the net health benefit, the more QALYs the intervention delivers.

$$\textbf{Net health benefit} = \text{total QALYs} - \frac{total\ costs}{cost\text{-}effectiveness\ threshold\ (€50{,}000/QALY)}$$

COST-EFFECTIVENESS RESULTS

The figures below illustrate the impact of the MS sherpa application on the clinical outcomes of disease progression and the average number of MS relapses per year in a scenario where we assume that 5% of patients who would experience disease progression switch early to a different MS drug as a result of using MS sherpa. Figure 6 shows which proportion of patients has a specific level of severity of MS. In the model, all patients have mild MS at the time of being diagnosed with MS. This is why the non-interrupted line starts with 100% of patients with mild MS in the top left corner of the Figure. As time goes on, some patients have disease progression and move to the group of patients with moderate or severe MS. As a result, the non-interrupted line diminishes over time as fewer people continue to have mild MS. At the same time, this causes the dashed line (moderate MS) and the dotted line (severe MS) to go up over time, as more people develop moderate or severe MS. The red lines reflect these developments over time with the use of MS sherpa, while the black lines reflect them without the use of MS sherpa. **Fout! Verwijzingsbron niet gevonden.**6 shows that, under this assumption, the use of MS sherpa slows down disease progression: a larger proportion of patients has mild MS (EDSS 0-3) for longer, a smaller proportion develops severe MS (EDSS 7-9) and severe MS develops at a later time than it does without the use of MS sherpa. Ten years after diagnosis, for example, the number of patients who continue to have mild MS is higher when using MS sherpa (63.5%) than when not using it (61.7%). The number of patients who progressed to moderate MS within ten years after their diagnosis is lower with the use of MS sherpa (22.8%) than without its use (23.4%). The same goes for the percentage of patients with severe MS: 13.0% if MS sherpa is used compared to 14.2% without its use. In addition, Figure 7 shows that the annual number of MS relapses is somewhat lower among users of MS sherpa. Ten years after the diagnosis of MS, the risk of an MS relapse is 15.7% without the use of MS sherpa, compared to 15.3% if it is used.

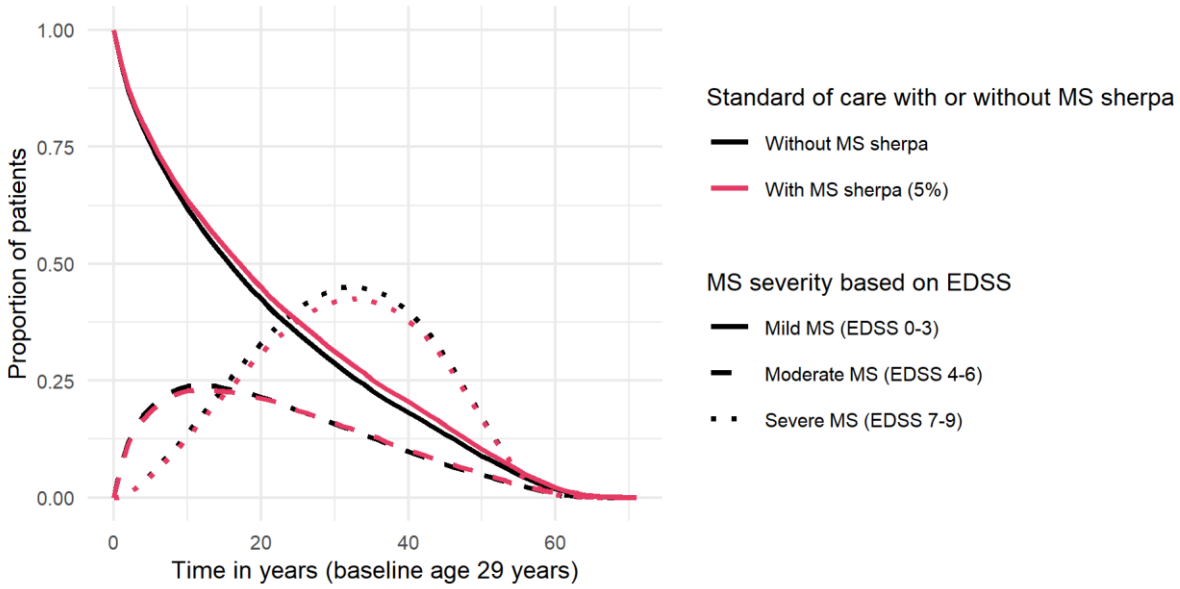


Figure 6. Development of the severity of MS measured over time with EDSS with and without the use of MS sherpa.

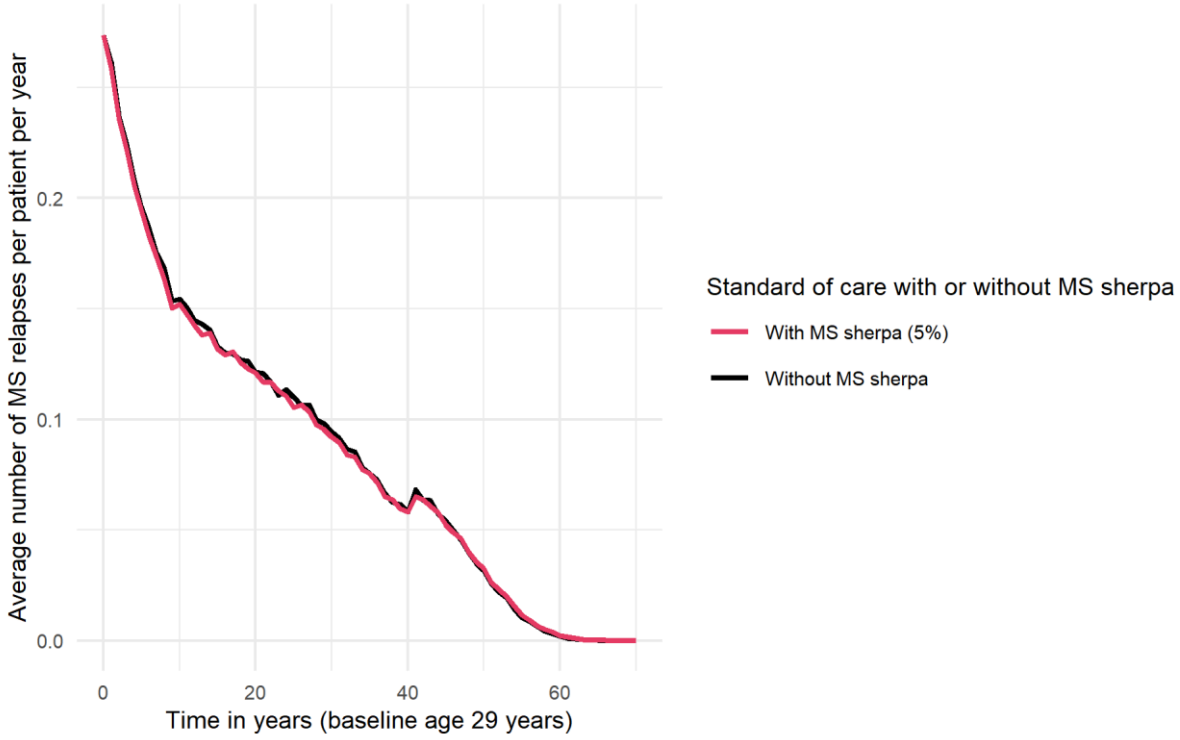


Figure 7. Average number of MS relapses per patient per year over time with and without the use of MS sherpa.

In the Netherlands, the threshold value of costs per QALY depends on the disease burden of the disease in question. The disease burden is a measure that indicates whether a disease leads to significant loss of health compared to a 'normal', healthy life. In case of a high disease burden, the treatment of the disease receives high priority, and a treatment (given equal effectiveness) is allowed to cost more than the treatment of a lower-priority disease. The disease burden is calculated by comparing the number of quality-adjusted life years (QALYs) a patient loses because of the disease with the number of QALYs of the average healthy Dutch person (for more information, see Chapter 2.2). The threshold value for MS is determined by calculating the number of QALYs in case of standard of care (in this case, the 'current treatment') without MS sherpa, using the MS model (without discounting). The total number of QALYs in case of standard of care without MS sherpa is 27.23. The quality-adjusted life expectancy of a Dutch population that does not have MS and that is comparable to the MS population with regard to age and gender (29 years of age and 74% women) is 46.74. This means that MS causes an absolute loss ('absolute shortfall') of 46.74-27.23 = 19.51 QALYs, or a proportional loss ('proportional shortfall') of 19.51/46.74 = 0.42, i.e. 42%. In other words, the disease burden of MS is 0.42. This number falls within the range of 0.41 to 0.70, which according to the standards of the National Health Care Institute corresponds to a reference value of €50,000 per QALY (see Chapter 2.2).[2]

The cost-effectiveness results of MS sherpa compared to standard of care are shown in Table 6 for all four scenarios of effectiveness of MS sherpa. In case of 5% or 10% assumed effectiveness, additional QALYs are gained (0.43 and 0.87, respectively), but this does involve higher costs in return. Assuming a reference value of €50,000, as calculated above, MS sherpa is cost-effective in these cases. In the scenarios in which the MS sherpa application is able to detect disease activity early in 15% or 20% of patients, the use of MS sherpa is cost-saving, while an average of 1.33 or 1.78 additional QALYs are gained compared to standard of care (MS sherpa becomes 'dominant' in these scenarios).

[2]There is a free tool to help calculate the disease burden of a condition: http://imta.nl/idbc

Table 6. Cost-effectiveness results of MS sherpa compared to standard of care from a societal perspective.

	Total		Difference between standard of care and MS sherpa		
Scenario	Costs	QALYs	Costs	QALYs	ICER
Standard of care	€614,732	20.51			
MS sherpa 5%	€620,990	20.94	€6,258	0.43	€14,535
MS sherpa 10%	€618,288	21.38	€3,556	0.87	€4,069
MS sherpa 15%	€614,538	21.84	€-194	1.33	D
MS sherpa 20%	€611,073	22.29	€-3,659	1.78	D

D = 'dominant'; this means the MS sherpa application delivers more health (in terms of QALYs) against fewer costs compared to standard of care.

The results of the base-case analysis as reported in Table 6 include costs outside of the healthcare sector – that is, the costs of informal care and of lost productivity. Table 7 displays the results of the cost-effectiveness analysis without these costs – in other words, from a healthcare perspective rather than a societal perspective. The results demonstrate lower total costs, but a greater difference in costs between healthcare with MS sherpa and healthcare without MS sherpa. This is due to the fact that certain benefits of the use of MS sherpa, such as a reduction in informal care and in loss of productivity thanks to reduced disease progression and fewer relapses, are not included in this cost calculation. The choice of perspective does not influence the calculation of QALYs. As the additional costs of MS sherpa are higher and the number of QALYs gained remains the same, the ICER of MS sherpa increases from the healthcare perspective compared to the societal perspective.

Table 7. Cost-effectiveness results of MS sherpa compared to standard of care from a healthcare perspective.

	Total		Difference between standard of care and MS sherpa		
Scenario	Costs	QALYs	Costs	QALYs	ICER
Standard of care	€530,345	20.51			
MS sherpa 5% effectiveness	€539,528	20.94	€9,183	0.43	€21,328
MS sherpa 10% effectiveness	€539,803	21.38	€9,458	0.87	€10,822
MS sherpa 15% effectiveness	€539,101	21.84	€8,756	1.33	€6,574
MS sherpa 20% effectiveness	€538,703	22.29	€8,358	1.78	€4,696

The tornado diagram in Figure 8 shows the results of the univariate sensitivity analysis. The central line in the diagram reflects the net health benefit if we fill in the base-case value for all parameters, with an assumed effectiveness of MS sherpa of 5%. The bars reflect the sensitivity of the cost-effectiveness results to various parameters. The longer the bar, the greater the sensitivity. The base-case value of most parameters was decreased or increased by 20% in order to test the parameter's impact. The blue bars show the net health benefit at the lowest value of the parameter in question (80% of the base case), while the red bars show the net health benefit at the parameter's highest value (120% of the base case). In the effectiveness parameter, the lowest value reflects 0% effectiveness and the highest value reflects 20% effectiveness. The parameters high up in the Figure have the greatest impact on net health benefit. The results show that the assumed effectiveness of MS sherpa has a considerable impact on net health benefit. The greater the share of patients who switch to a

different drug as a result of MS sherpa, thus preventing disease activity, the greater the net health benefit. This means that the provider of this AI application could focus on further improving the detection of disease activity, so that the greatest possible share of patients can switch to a more effective MS drug before the disease activity actually occurs. The tornado diagram also shows that the quality of life and costs of healthcare in the case of mild MS (EDSS 0-3) have a considerable impact on net health benefit. This occurs because patients have mild MS for most of their life, and this period is extended by the use of MS sherpa. If the quality of life in the case of mild MS increases or healthcare costs for mild MS decrease, the net health benefit naturally increases. The costs of the MS sherpa app are also varied, ranging from 200 euros per user per year (lowest value) to 1,000 euros per user per year (highest value). At the highest value of 1,000 euros, the net health benefit decreases. In the scenario with 5% effectiveness, the incremental cost-effectiveness ratio at this price is €30,100 per QALY.

Figure 8. Tornado diagram showing the results of the univariate sensitivity analysis

3.4 POTENTIAL BROADER EFFECTS

With the help of self-learning software, self-monitoring applications may be able to detect patterns, thus improving insight into progression of the disease. The dialogue between healthcare professionals and patients about the insights gathered via the application can increase patients' self-efficacy. Patients may also feel empowered by the data from the app, as their individual experience can be rendered objectively and over time. The information can hereby encourage shared decision-making in the consultation room. MS sherpa may also contribute to more appropriate care if the insights are used to decide which consultations are necessary and

which are not. Furthermore, it will also be investigated whether monitoring via MS sherpa can contribute to the safe discontinuation of medication.

With the large-scale roll-out of a self-monitoring applications such as MS sherpa, data are collected that can contribute to more accurate insights provided by the AI application. The developer of MS sherpa intends for the application to enable short-term predictions about users' disease progression or well-being and to link these to actions or advices, allowing patients and healthcare professionals to anticipate these developments. This would allow patients to prepare mentally for oncoming disease activity and to make adaptations to the home. The planned studies of the use of MS sherpa have yet to demonstrate these and other effects of the application.

INSIGHTS FROM OTHER STUDIES

Other studies on health applications for MS already offer some insight into the potential effects of the use of MS sherpa. For example, a study on the experiences of seven patients with the use of MS sherpa's predecessor, 'Mijn Kwik', showed that the application raised patient awareness [38]. On the one hand, this had benefits such as a better understanding of their illness and the lifestyle factors that could influence their functioning. Some patients took action based on the data from the app, such as going to bed earlier or trying to exercise more. Greater awareness about one's condition may also have a positive effect on therapy compliance [39]. On the other hand, patients also experienced drawbacks, such as greater anxiety, confrontation with one's limitations, and pressure to effect lifestyle changes where this is physically impossible. At times, patients also found it difficult to interpret the application's findings and act accordingly. Participants in the study stated they would like more support with interpreting the application's findings and acting correspondingly, in terms of both information in the application itself and guidance from healthcare workers. Lastly, patients wished for the application to be integrated into the neurologist's personal treatment plan. The information from the studies of the Mijn Kwik application was taken into account in the design and development of MS sherpa. Accordingly, the experiences with Mijn Kwik cannot be translated directly into the expected experiences with MS sherpa. Measurements with Mijn Kwik were more frequent (daily) than those with MS sherpa will be (at least once a month). The Mijn Kwik application was also combined with an activity tracker, FitBit, yielding more insight into physical activity and sleep.

Another application for MS self-monitoring, Floodlight, not only uses the collected data to gain insights at the patient level but also creates a fairly accessible database, Floodlight Open, to add to scientific knowledge and stimulate research [40]. This initiative demonstrates that the data from a self-monitoring application can have broader value. The data collected via MS sherpa may also be of value outside of the AI application, for instance through integration with the electronic patient record or through inclusion in a public database. The added value of the data collected via MS sherpa compared to existing data sources has not yet been studied. When data are used for such purposes, patients will need to provide their consent. A survey among 143 MS patients found that about 60% of respondents feel the need to better understand their individual disease progression, but that almost half of them do not want to disclose personal data [29]. A recent study by the Netherlands Patients Federation found that, when it comes to consenting to sharing personal data and body material, patients care about for what purpose the data will be used: 16% of the respondents would provide consent to organisations that develop non-medical aids such as food supplements, lifestyle support, sports products and health apps. More patients are willing to share data with organisations that develop medical aids or carry out medical scientific research (62% and 92%, respectively) [41].

The extent to which self-monitoring applications are used will also depend on patients' individual wishes. The need for broader benefits such as personal control and a better understanding of one's own health is a personal one and may not be equally strong in every patient. For instance, a study on an information portal for patients with a congenital heart defect found that only half of the patients actually made use of the portal [42]. The ease of use of new healthcare tools also plays an important role. As an example, a significant impediment to the use of the ABC tool to monitor disease progression in patients with COPD was that the software had not been integrated into the electronic patient record of the GP practices [43]. When patients notice that the information

they supply via the application is seen by the neurologist and can lead to new insights or actions, they will be encouraged to use the application and to complete measurements with certain frequency. To optimise appropriate use of the application, and consequently its effectiveness, integration into the healthcare process will be crucial.

3.5 INSIGHTS FROM NEUROLOGISTS

Two neurologists who are involved with the planned and ongoing clinical studies of MS sherpa were asked about their ideas regarding the (added) value of MS sherpa, as well as the value of HTA for this case study.

Both neurologists stated they found the new AI applications for monitoring MS patients, such as MS sherpa, to be promising developments. Registering clinical outcome measures on an ongoing basis, rather than obtaining snapshots during clinical visits, generates a lot of additional information. The healthcare process could eventually be attuned to this, for instance with regard to planning consultations and making MRI scans. It would be a positive development if patients could be monitored without requiring a hospital visit. The neurologists doubt whether the information collected by MS sherpa in and of itself will in the short term lead to different treatment choices with regard to medication, as is assumed in the early HTA. Although the measurements have been proved to be relevant and insightful with regard to the progression of MS symptoms, more certainty is needed about the way measurements in MS sherpa reflect disease activity. While the neurologists see potential in MS sherpa, they emphasise that its value has yet to be proved by currently planned clinical studies. The current clinical outcome measures are not ideal either and leave much room for improvement. The neurologists expect that a multi-domain approach, in which the outcome measures obtained with MS sherpa are combined with other elements such as MRI scans or perhaps other non-invasive measurements, can yield the greatest value.

Once the effectiveness of the AI application in terms of impact through changes in medication is sufficiently demonstrated by published scientific studies, the neurologists consider it highly likely that they would recommend that patients switch to a different drug or even discontinue medication on based on, or informed in part by, information from MS sherpa. It is likely that most neurologists will first want to gain sufficient experience with the AI application before initiating a switch in treatment without also consulting current diagnostics (MRI scans). As yet, the neurologists cannot estimate for which percentage of patients in whom MS sherpa detects a deterioration they would recommend a switch in treatment, given that not enough is known at this point about how the data from MS sherpa reflect disease activity. Consequently, it is not possible to choose a preferred scenario in our early HTA on the basis of the assessment by the neurologists. They do consider the chosen bandwidth to be realistic. Decision-making analysis techniques can be applied to gain a better understanding of the possible impact in clinical practice of AI applications such as MS sherpa, by estimating how the information from the AI application influences treatment decisions and calculating the subsequent effects on clinical outcomes [32].

The neurologists state they are increasingly aware of the costs of the healthcare they provide. They therefore believe it is useful to carry out an HTA for an AI application. It is expected that the importance of such analyses will only grow. An HTA of MS sherpa is considered a bit premature in view of the remaining uncertainty about the extent to which MS sherpa can detect disease activity and since it is not known how neurologists will use this information in clinical practice. As such, they confirm that MS sherpa is promising, but the results of the early HTA must be interpreted in the light of assumed effectiveness, which will require additional proof obtained from clinical studies and clinical practice. The use of MS sherpa also has added value outside of the case study presented, namely through home monitoring and self-monitoring, which are increasingly important in view of time and staff shortages in healthcare.

4. DISCUSSION

The results of the economic evaluation show that, under certain assumptions on effectiveness, the AI application MS sherpa can be cost-effective or even cost saving. A societal perspective was used here, meaning that, apart from health benefits and costs within healthcare, the analyses also considered effects in the areas of informal care, paid work and unpaid work. The actual effectiveness of MS sherpa has yet to be demonstrated in clinical studies and clinical practice. The effectiveness will be driven by on the one hand the accuracy with which MS sherpa can detect disease activity at an early stage, and on the other hand from the extent to which this information will influence treatment decisions. The effects of MS sherpa may be greater than is shown by the results of the economic evaluation, as the calculations only consider the potential influence of MS sherpa on the switch to a different MS drug in case of oncoming disease activity.

HTA OF AI INTERVENTIONS

The economic evaluation of healthcare interventions is always complex, but the evaluation of AI interventions involves additional challenges if the intervention's primary goal is not to improve health but to improve broader outcomes such as self-efficacy, shared decision-making or well-being. In standard economic evaluations, such as the one carried out for MS sherpa in this report, these values are not included in the calculations. However, these standard analyses for economic evaluations remain essential for decision-making in view of their uniformity and comparability across other healthcare interventions, as the analyses make it possible to inform decisions with the aim to achieve maximum health benefit using a limited budget. There is a growing recognition of the broader benefits that healthcare interventions may have, and the expectations regarding broader benefits are increasing within the society. To establish the full picture of an AI intervention's value, it is therefore important to ascertain whether there are any benefits in other dimensions than healthcare. This would make it possible to weigh these benefits as contextual factors in HTA decision-making. These broader effects were explored as part of this HTA and as an addition to the calculations in the economic evaluation. For example, the literature and insights from the supplier of the AI application show that the use of MS sherpa may also have effects outside of health dimensions, such as greater self-reliance, learning to cope better with one's condition, anticipating a deterioration in health and a richer database on MS disease activity. However, there may also be negative consequences, such as being regularly reminded of one's illness or being confronted with physical deterioration that would have remained undetected without an application due to its gradual onset [38]. Some patients may also have privacy concerns related to the use of the application. In a general sense, the risk that HTA decision-making on the basis of broader benefits will be at the expense of the budget for interventions that deliver more health benefits should also be kept in mind. The risk of sacrificing health as a result of weighing broader benefits does not apply, or not as much, if the intervention is funded through other channels, such as a personal contribution on the part of the patient.

PRICE-INCREASING EFFECT OF AN EARLY HTA

The use of an economic evaluation as part of an early HTA can help set fair prices in case other pricing mechanisms do not apply the way they do in regulated competition in Dutch healthcare [44]. Early HTA makes it possible to estimate the health value of AI applications and to express this as a price, even when there are still few data on effectiveness. This price based on the economic evaluation is also called the value-based price, as it represents the monetary value of the health that gets created, given a certain willingness to pay for health. The price of an AI application that is 'cost-effective' is therefore a price that is in balance with the application's health value but may not be proportionate to the development costs or the customary market prices of similar technology. In order to determine whether the price of an AI application in healthcare is appropriate, it is not just the value-based price coming out of a decision model that should be taken into consideration. It should also be considered whether the price of the service or product fits in with that of similar services and products that have a consumer market of similar size. If pricing in comparable markets is lost sight of, an early HTA may have a price-increasing effect, given that all created health value gets included in the price.

4.1 CONCLUSION

Similar to other healthcare interventions, AI applications can contribute to the length and quality of patients' lives, but they often also entail additional costs. The extent to which these additional costs lead to savings elsewhere in healthcare or in society, or the extent to which they are proportionate to the health benefits achieved, can be examined by way of an HTA. Investing in an HTA has most added value if it needs to be determined whether the intervention can be reimbursed from the basic health insurance package. An HTA may also be useful when benefits occur outside of the purchasing healthcare institution, for instance to demonstrate an AI application's added value to funders such as health insurers.

Carrying out an HTA will require making decisions, among others, on the perspective used, the choice of analysis technique and the structure of the data collection.

For this report, a case study of an AI application in MS was assessed with an HTA. The HTA demonstrates that the intervention under evaluation is potentially cost-effective or cost saving, provided that clinical studies and clinical practice confirm the expected effectiveness. The use of the AI application is only cost-effective if broader societal benefits such as productivity gains and informal care are weighed in an analysis that uses a societal perspective. If only costs and benefits within healthcare are weighed, the intervention is potentially cost-effective, but not cost saving.

There are limits to the added value of an HTA for assessing the value of AI applications. First, the information provided by an HTA is not the most important information to all parties involved. Sometimes, a business case will suffice. Second, AI applications may have benefits that are usually not included or are only included to a limited extent, as they have an effect on well-being but do not have a demonstrable effect on *health-related* quality of life or on costs. These limitations of the HTA are not unique to AI applications.

The added value of AI applications has already been demonstrated in various sectors. The extent to which the use of an AI application in healthcare has added value can be evaluated with an HTA. The case study carried out here shows that a promising AI application may be cost-effective and may even be cost saving.

REFERENCES

[1] Ministerie van Economische Zaken en Klimaat, „Strategisch Actieplan voor Artificiële Intelligentie," 2019. [Online]. Available: https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/beleidsnotas/2019/10/08/strategisch-actieplan-voor-artificiele-intelligentie/Rapport+SAPAI.pdf. [Geopend 15 Mei 2021].

[2] H. Koffijberg, B. van Zaane en K. Moons, „From accuracy to patient outcome and cost-effectiveness evaluations of diagnostic tests and biomarkers: an exemplary modelling study," *BMC Med Res Methodol.,* vol. 13, nr. 12, 2013.

[3] L. G. Jenniskens K, C. Naaktgeboren, L. Hooft, K. Moons, J. Poldervaart, H. Koffijberg en J. Reitsma, „Decision analytic modeling was useful to assess the impact of a prediction model on health outcomes before a randomized trial.," *J Clin Epidemiol,* vol. 115, pp. 106-115, 2019.

[4] B. L. D. W. G. M. K. Lagerweij GR, L. Benschop, A. Maas, A. Franx, M. Wermer, J. Roeters van Lennep, B. van Rijn en H. Koffijberg, „Impact of preventive screening and lifestyle interventions in women with a history of preeclampsia: A micro-simulation study," *Eur J Prev Cardiol,* vol. 27, nr. 13, pp. 1389-1399, 2020.

[5] Zorginstituut Nederland, „Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg," 2016. [Online]. Available: https://www.zorginstituutnederland.nl/binaries/zinl/documenten/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg.pdf. [Geopend 15 Mei 2021].

[6] EuroQol, „EuroQol is developing a new instrument – the EQ-HWB," 2021 Februari 2021. [Online]. Available: https://euroqol.org/euroqol-is-developing-a-new-instrument-the-eq-hwb-2/. [Geopend 12 Maart 2021].

[7] EuroQol, „What's the Difference: Comparing societal value sets for the 5L and 3L using the same protocol," 6 Maart 2018. [Online]. Available: https://euroqol.org/wp-content/uploads/2018/07/2018-05-06-EQ5DUS-5Lvs3L.pdf . [Geopend 12 Maart 2021].

[8] Zorginstituut Nederland, „Ziektelast in de praktijk," 2018. [Online]. Available: https://www.zorginstituutnederland.nl/binaries/zinl/documenten/rapport/2018/05/07/ziektelast-in-de-praktijk/Ziektelast+in+de+praktijk_definitief.pdf. [Geopend 15 Mei 2021].

[9] J. Wolff, J. Pauling, A. Keck en J. Baumbach, „ The Economic Impact of Artificial Intelligence in Health Care: Systematic Review," *Journal of medical Internet research,* vol. 22, nr. 2, 2020.

[10] ZonMw, „Programmavoorstel HTA Methodologie 2021-2024," 2021. [Online]. Available: https://publicaties.zonmw.nl/programmatekst-hta-methodologie-2021-2024/ . [Geopend 15 Mei 2021].

[11] M. Drummond, M. Sculpher, G. Torrance, B. O'Brien en G. Stoddart, Methods for the Economic Evaluation of Health Care Programmes, Oxford University Press, 2005.

[12] K. Payne, M. McAllister en L. Davies, „Valuing the economic benefits of complex interventions: when maximising health is not sufficient," *Health economics,* vol. 22, nr. 3, pp. 258-271, 2013.

[13] T. Kidd, N. Carey, F. Mold, S. Westwood, M. Miklaucich, E. Konstantara, A. Sterr en D. Cooke, „A systematic review of the effectiveness of self-management interventions in people with multiple sclerosis at improving depression, anxiety and quality of life," *PLoS One,* vol. 12, nr. 10, p. e0185931, 2017.

[14] Centraal Bureau voor de Statistiek, „Overkoepelend beeld Monitor Brede Welvaart & de Sustainable Development Goals 2020," 2020. [Online]. Available: https://www.cbs.nl/-/media/_pdf/2020/21/cbs-mbw-2020.pdf. [Geopend 15 Mei 2021].

[15] H. Al-Janabi, T. Flynn en J. Coast, „ Development of a self-report measure of capability wellbeing for adults: the ICECAP-A," *Quality of Life Research,* vol. 21, pp. 167-176.

[16] E. de Bekker-Grob, C. Berlin, B. Levitan, K. Raza, K. Christoforidi, I. Cleemput, J. Pelouchova, H. Enzmann, N. Cook en M. Hansson, „Giving Patients' Preferences a Voice in Medical Treatment Life Cycle: The PREFER Public-Private Project," *Patient,* vol. 10, nr. 3, pp. 263-266, 2017.

[17] D. Lakdawalla, J. Doshi, L. Garrison, C. Phelps, A. Basu en P. Danzon, „Defining Elements of Value in Health Care—A Health Economics Approach: An ISPOR Special Task Force Report [3]," *Value in Health,* vol. 21, nr. 2, pp. 131-139, 2018.

[18] G. Rosati, „The prevalence of multiple sclerosis in the world: an update," *Neurological sciences,* vol. 22, nr. 2, pp. 117-139, 2001.

[19] Zorginstituut Nederland, „Monitor MS-Geneesmiddelen 2019," 2019. [Online]. Available: https://www.zorginstituutnederland.nl/binaries/zinl/documenten/rapport/2019/06/13/monitor-ms-geneesmiddelen-2019/Monitor+MS+geneesmiddelen+2019.pdf. [Geopend 15 Mei 2021].

[20] Z. Salimzadeh, S. Damanabi, L. Kalankesh en R. Ferdousi, „Mobile Applications for Multiple Sclerosis: a Focus on Self-Management," *ACTA INFORM MED,* vol. 27, nr. 1, pp. 12-18, 2019.

[21] N. Griffin en M. Kehoe, „A questionnaire study to explore the views of people with multiple sclerosis of using smartphone technology for health care purposes," *Disabil Rehabil,* vol. 40, nr. 12, pp. 1-9, 2017.

[22] M. Cohen, „Connected health and multiple sclerosis," *Rev Neurol,* vol. 174, nr. 6, pp. 480-485, 2018.

[23] ZorgInnovatie, „MS sherpa," December 2020. [Online]. Available: https://www.zorginnovatie.nl/innovaties/ms-sherpa.

[24] H. Khalid, E. Shihab, M. Nagappan en A. Hassan, „What do mobile app users complain about?," *IEEE Software,* vol. 32, nr. 3, pp. 70-77, 2015.

[25] G. Giunti, J. Kool, O. Romero en E. Zubiete, „Exploring the specific needs of persons with multiple sclerosis for mHealth solutions for physical activity: mixed-methods study," *JMIR mHealth and uHealth,* vol. 6, nr. 2, p. e37, 2018.

[26] K. Costello, P. Kennedy en J. Scanzillo, „Recognizing nonadherence in patients with multiple sclerosis and maintaining treatment adherence in the long term," *The Medscape Journal of Medicine,* vol. 10, nr. 9, p. 225, 2008.

[27] P. van Oirschot, M. Heerings, K. Wendrich, B. den Teuling, M. Martens en P. Jongen, „Symbol Digit Modalities Test Variant in a Smartphone App for Persons With Multiple Sclerosis: Validation Study," *JMIR Mhealth Uhealth,* vol. 8, nr. 10, 5 Oct 2020.

[28] Lam et al., „Reliability, concurrent and ecological validity of smartphone-based cognition and walking tests ," in *(presented at ECTRIMS 2020)*, 2020.

[29] Nationaal MS Fonds, „Digitale zelfmonitoring," [Online]. Available: https://nationaalmsfonds.nl/onderzoek/digitale-zelfmonitoring/. [Geopend 4 Maart 2021].

[30] Amsterdam Universitair Medisch Centra, „DOT-MS: Stoppen van medicatie bij jarenlang stabiele multiple sclerose," [Online]. Available: https://www.vumc.nl/zorg/expertisecentra-en-specialismen/ms-centrum-amsterdam/informatie-voor-patienten-ms-centrum/meedoen-aan-onderzoek/dot-ms-stoppen-van-medicatie-bij-jarenlang-stabiele-multiple-sclerose.htm. [Geopend 4 Maart 2021].

[31] ZonMw, „Stoppen met medicatie bij jarenlange stabiele MS: DOT-MS (discontinuation of therapy in multiple sclerosis)," [Online]. Available: https://www.zonmw.nl/nl/actueel/nieuws/detail/item/stoppen-met-medicatie-bij-jarenlange-stabiele-ms-dot-ms-discontinuation-of-therapy-in-multiple-scl/. [Geopend 4 Maart 2021].

[32] T. Kappen, W. van Klei, L. van Wolfswinkel, C. Kalkman, Y. Vergouwe en K. Moons, „ Evaluating the impact of prediction models: lessons learned, challenges, and recommendations," *Diagn Progn Res,* vol. 2, nr. 11, 2018.

[33] Zorginstituut Nederland, „GIPdatabank.nl," [Online]. Available: www.gipdatabank.nl. [Geopend September 2020].

[34] Zorginstituut Nederland, „Farmacotherapeutisch kompas," 2019. [Online]. Available: wwww.farmacotherapeutischkompas.nl.

[35] B. Uitdehaag, G. Kobelt, J. Berg, D. Capsa, J. Dalén en E. M. S. Platform, „European Multiple Sclerosis Platform. New insights into the burden and costs of multiple sclerosis in Europe: Results for the Netherlands.," *Multiple Sclerosis,* vol. 23, nr. Suppl 2, pp. 117-129, 2017.

[36] T. Kanters, C. Bouwmans, N. van der Linden, S. Tan en L. Hakkaart-van Roijen, „Update of the Dutch manual for costing studies in health care," *PLoS One,* vol. 12, nr. 11, 2017.

[37] E. van Egmond, D. van Gorp, C. Honan, M. Heerings, P. Jongen, J. van der Klink, M. Reneman en E. Beenakker, „. A Dutch validation study of the Multiple Sclerosis Work Difficulties Questionnaire in relapsing remitting multiple sclerosis," *Disability and Rehabilitation,* 2019.

[38] K. Wendrich, P. van Oirschot, M. Martens, M. Heerings, P. Jongen en L. Krabbenborg, „Toward Digital Self-monitoring of Multiple Sclerosis: Investigating First Experiences, Needs, and Wishes of People with MS," *International Journal of MS Care,* vol. 6, 2019.

[39] K. Stockl, J. Shin, S. Gong, A. Harada, B. Solow en H. Lew, „Improving patient self-management of multiple sclerosis through a disease therapy management program," *Am J Manag Care,* vol. 16, nr. 2, pp. 139-144, 2010.

[40] Roche, „Floodlight," [Online]. Available: https://floodlightopen.com/nl-NL/. [Geopend Februari 2021].

[41] Patiëntenfederatie Nederland, „Delen van data in gezondheidszorg," 2021. [Online]. Available: https://www.patientenfederatie.nl/downloads/monitor/847-monitor-delen-van-data/file. [Geopend 15 Mei 2021].

[42] J. Etnel, L. Bons, F. de Heer, D. Robbers-Visser, I. van Beynum, B. Straver, M. Jongbloed, P. Kiès en M. Slieker, „Patient information portal for congenital aortic and pulmonary valve disease: a stepped-wedge cluster randomized trial," *Open Heart,* vol. In press.

[43] A. Slok, M. Twellaar, L. Jutbo, D. Kotz, N. Chavannes, S. Holverda, P. Salomé, P. Dekhuijzen, M. Rutten-van Mölken, D. Schuiten, J. In 't Veen en O. van Schayck, „'To use or not to use': a qualitative study to evaluate experiences of healthcare providers and patients with the assessment of burden of COPD (ABC) tool," *Prim Care Respir Med,* 2016.

[44] M. Versteegh, „Prijsbepaling van medicijnen: Wat is redelijk?," *Nederlands Tijdschrift voor Geneeskunde,* vol. 164, nr. 34, 2020.