Assessing what matters: the role of normative commitments in Health Technology Assessment

Bloemen, B.

2025, Dissertation

Assessing what matters: the role of normative commitments in Health Technology Assessment

Bart Bloemen

Assessing what matters: the role of normative commitments in Health Technology Assessment

Bart Bloemen

The work presented in this thesis was performed at the Department for Health Evidence, Radboud University Medical Center, Nijmegen, the Netherlands

ISBN: 978-94-6510-460-7

Cover design: Didi Xie en Phylisha van Heemskerken
Lay-out: Dennis Hendriks, gwhconcepts.nl
Printing: ProefschriftMaken.nl

Assessing what matters: the role of normative commitments in Health Technology Assessment

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 14 maart 2025
om 10.30 uur precies

door

Bart Bloemen
geboren op 30 oktober 1987
te Vierlingsbeek

Promotoren:
Prof. dr. G. J. van der Wilt
Prof. dr. B.G.M. van Engelen

Copromotor:
Dr. W.J. Oortwijn

Manuscriptcommissie:
Prof. dr. M.G.M. Olde Rikkert (voorzitter)
Dr. L. Krabbenborg
Prof. dr. mr. S.M.A.A. Evers (Maastricht University)

TABLE OF CONTENTS

"In an objective system...any mingling of knowledge with values is unlawful, forbidden. But — and here is the crucial point, the logical link which at their core weds knowledge and values together — this prohibition, this 'first commandment' which ensures the foundation of objective knowledge, is not itself objective. It cannot be objective: it is an ethical guideline, a rule for conduct. True knowledge is ignorant of values, but it cannot be grounded elsewhere than upon a value judgment, or rather upon an axiomatic value. It is obvious that the positing of the principle of objectivity as the condition of true knowledge constitutes an ethical choice and not a judgment arrived at from knowledge, since, according to the postulate's own terms, there cannot have been any 'true' knowledge prior to this arbitral choice. To establish the norm for knowledge the objectivity principle defines a value: that value is objective knowledge itself. Thus, assenting to the principle of objectivity one announces one's adherence to the basic statement of an ethical system, one asserts the ethic of knowledge. "

(Jacques Monod, *Chance and Necessity,* 1970)


1

Chapter 1

General introduction

IS THERE A PROBLEM OF HEALTH TECHNOLOGY?

Health technology is advancing rapidly, allowing unprecedented interventions in health and disease processes, and potentially transforming society and humankind in a profound way (Berloznik et al., 2006). Developments in technology promise new methods for disease prevention, diagnosis, and treatment but also evoke complex questions about the nature of human life, the sustainability of healthcare systems (due to rising costs, impact on environment), and the potential impact on society. This issue is exacerbated by the fact that technologies are increasingly aimed at enhancement of human life, rather than (only) combating disease, giving rise to ethical questions (Savulescu & Bostrom, 2009).

These ethical questions include rethinking the value of health technology and health itself. We no longer define health as merely the absence of disease, but as a state of complete physical, mental, and social well-being (World Health Organization, 1995). This perspective recognizes health as a means to achieve valuable states of being (Richardson, 2016). Technologies that intervene in processes constituting health therefore can contribute to the realization of multiple goals that represent different values. For example, wearable devices (e.g., smartwatches, fitness trackers) monitor health and disease, while also providing information to enhance athletic performance. Three-dimensional printed prosthetics and implants assist patients in rehabilitation and serve educational purposes by creating anatomical models for training of healthcare professionals. Furthermore, deep brain stimulation offers prospects for treating depression, but also let us wonder whether we could improve cognitive functions of humans.

Because health technology can serve multiple purposes, there are also different perspectives on what makes a health technology valuable. For example, the introduction of cochlear implants received mixed responses among people within the Deaf community that viewed it as a threat to their culture and identify (Reuzel, 2001).

Uncertainty in predicting what will happen when introducing a health technology can create other ethical questions. Due to their complexity, we will never be sure whether the intended effect will be realized and unintended harm (side effects) can occur. We can take safety measures, for example by postponing the introduction of health technologies into clinical practice until their safety is shown in long-term clinical trials, but this simultaneously delays their availability to patients that need treatment. Therefore, difficult trade-offs need to be made between efforts to reduce or mitigate uncertainty and stimulate valuable innovation.

There can also be *too much health technology* (Hofmann, 2015). Because health technology allows us to see, manipulate, and measure new things (e.g., mutations in our DNA, the microbiome), it also contributes to the recognition and (re-)labeling of certain states of being as a disease. For example, hypertension and hypercholesterolemia gain medical attention because of new abilities to measure and manipulate it. This risks overdiagnosis and unnecessary treatment (Hofmann, 2015).

The increasing availability of health technology also risks outrunning the capacity of our healthcare systems to accomodate the uptake of them. Health technology is a factor in the growth of our expenditures on healthcare (Sorenson et al., 2013). Combined with increased healthcare demand due to an aging population and the rise in chronic diseases, increasing costs could make it unfeasible to introduce new innovations into healthcare systems without jeopardizing universal healthcare coverage. The seriousness of these challenges led the Netherlands Scientific Council for Government Policy to recommend to the Dutch government to make *better choices to steer and limit the growth of healthcare* (Wetenschappelijke Raad voor het Regeringsbebeleid, 2021).

Therefore, despite offering possibilities to reduce suffering from disease, the development of health technology also challenges us to make difficult decisions and to improve our abilities to reflect upon its value to identify those technologies that have *real value for patients and society*.

HEALTH TECHNOLOGY ASSESSMENT: IS IT OVERLOOKING *DESIRABILITY*?

Increasingly, countries worldwide use formal processes of *informed decision-making* regarding the use of health technology. A central element of these processes has become **Health Technology Assessment** (HTA): using *explicit methods to determine the value of a health technology at different points in its lifecycle* (Loblova, 2016; O'Rourke et al., 2020; Teerawattananon et al., 2021; WHO, 2021). HTA can be conducted at universities and hospitals, insurance companies, governmental institutes, or independent research organizations, but must often the responsibility for conducting HTA to inform national decisions on health technology is delegated to a dedicated HTA agency (Fontrier et al., 2022). Experts at these agencies review available evidence on the intended and unintended consequences of using health technology to inform decisions on health benefit packages for reimbursement, priority-setting, resource allocation to different health technologies and programs, and the development of clinical guidelines.

Originally, HTA aimed to broadly assess health technology including safety, clinical effectiveness, cost-effectiveness, and their ethical, legal and social implications, but during its development the focus of HTA was narrowed to cost-effectiveness (Bellemare et al., 2018; Daniels & van der Wilt, 2016; Lehoux, 2006; Lehoux & Williams-Jones, 2007). Consequently, HTA draws attention to questions about *affordability* (e.g., does this health technology provide 'value for money'? Can the healthcare system afford to provide this health technology?). Rising costs (induced by health technology) and subsequent threats to the financial sustainability of healthcare systems are an important matter. However, being affordable is only one aspect that makes health technology *desirable*, and HTA should aim to answer questions about the desirability of health technology to be policy relevant (Lehoux, 2006; Oortwijn et al., 2022). HTA aims to help decision-makers in making *better* decisions. This requires the identification of all *relevant* consequences of health technology, including societal and ethical implications, and the collection of *reliable* information on these consequences. To ground decisions, this information also needs to imply something about the *value* of different choices, i.e., to know what to do, one needs to know which of the (technological) options is *better*. Focusing on cost-effectiveness information presupposes that this is what is important to consider and makes a health technology more desirable.

HTA agencies recognize that cost-effectiveness is not the only or most important aspect that matters when informing decision-making. For example, the National Institute for Health and Care Excellence (NICE) in England states that maximization of quality-adjusted life years (QALYs), a generic measure of the impact of health on *quality of life* that is often used in cost-effectiveness analysis, would be an appropriate standard in HTA *if it covers the only benefits of health technology* worth considering, but also states that there are other important benefits (Richardson, 2016). Furthermore, some HTA agencies, such as the German Institute for Quality and Efficiency in Health Care (IQWiG) and the French Haute Autorité de Santé (HAS) focus on clinical benefits as the most important criterion to be considered in HTA.

Despite its original intent and the variety of policy questions raised by health technology, HTA is still predominantly focused on cost-effectiveness (measured in a particular way). What are the reasons for this? Why do other aspects of value, and questions about the (un)desirability of health technology, receive less attention in assessments (DeJean et al., 2009; Ekmekci & Guner, 2019; Legault et al., 2021)?

One interesting suggestion from literature is that this is the result from the way in which principles of evidence-based medicine guide HTA (Otto et al., 2021). HTA is guided by a set of implicit or explicit norms that render a judgment about the

value of health technology *valid* and *relevant* (Moors & Peine, 2016). Evaluating the validity of claims about the (potential) value of health technology is part of HTA practitioners' expertise, making use of a set of epistemological principles that help determining what can be regarded as *evidence of* valuable properties and consequences of health technology. Driven by ideas about what is reliable evidence, HTA primarily considers those aspects of health technology that are amenable to *objective* and *empirical* inquiry (Moors & Peine, 2016; Otto et al., 2021; Refolo et al., 2016). This strict focus on objective empirical data may render information on cost-effectiveness as reliable, whereas ethical analysis may be seen as unreliable because it is not easily amenable and reducible to empirical inquiry.

If broadening the scope of HTA to ethical and societal implications is seen as introducing unreliable *subjective* elements into assessments, the integration of such methods can receive resistance because they are at odds with the basic idea of HTA to focus on objectively describable aspects of health and health technology (Ducey et al., 2017; Richardson, 2016).

Therefore, to regain the ability of HTA to inform decision-making about the different value questions that arise by the use of health technology, its guiding norms that exclude explicit attention towards value issues should be questioned. A fruitful strategy would be to show that ethical analysis is not something different in nature or antithetical to the epistemology of analyses already conducted in HTA (e.g., safety, clinical effectiveness, cost-effectiveness), but part of a continuum of analyses grounded in *normative* presumptions. This implies acknowledging, and rethinking, the **normativity** of HTA itself, highlighting how HTA is already governed by (implicit) ideas about what is good, reliable, and relevant when it comes to evaluating health technology. By showing how this normativity already plays a role in existing analyses, one can challenge the idea that explicit consideration of value judgments (being normative in nature) would be antithetical to HTA.

WHO IS AFRAID OF NORMATIVITY?

HTA practitioners increasingly acknowledge that HTA is **normative** [1], because HTA is used to address questions concerning the *value* of health technology and to inform decisions such as determining the health benefit package (Charlton et al., 2023). This normativity of HTA also plays a central role in the conduct of HTA itself because it requires a normative framework to identify the facts that *matter* and to interpret those facts in light of the decision at hand (Hofmann et al., 2014; van der Wilt et al., 2022). Normative judgments are needed to decide for example upon which technologies to assess, to select relevant outcome measures to be used in an assessment, and to set thresholds (e.g., to determine when a health technology is considered cost-effective; to determine how much evidence is sufficient to support a conclusion).

Despite this awareness of the normativity of HTA, in practice a careful distinction is maintained between HTA practitioners who are responsible for conducting the *assessment* (i.e., collecting, synthesizing and interpreting available information) and those who are responsible for an *appraisal* of the outcomes of an assessment to inform and make decisions, see Figure 1 (Oortwijn et al., 2022; Sandman & Heintz, 2014; Walley, 2007). This distinction denotes different tasks and responsibilities of different actors involved in the HTA process (see Figure 1). HTA practitioners are responsible for evaluating relevant value dimensions of a health technology *without drawing any conclusions about whether, and how, the health technology should be (de) implemented*, which is often the responsibility of a committee (involving both experts and stakeholders) that draws conclusions or formulates recommendations to the final decision body (often a ministry of health) (Angelis et al., 2018; Fontrier et al., 2022; Kleinhout-Vliek et al., 2021). In other words, the mandate of HTA practitioners is confined to judgments about what *one ought to expect to happen*, whereas normative judgments concerning what *one ought to do* is delegated to those with appropriate authority. This practice makes it difficult for HTA practitioners to acknowledge that they make normative judgments because this may be seen as a threat to their assigned role in decision-making (Boothe, 2019; Ducey et al., 2017).

1 There is no consensus and clarity on the term 'normative', and different associated terms ('value judgment', 'social value judgment', 'preferences') are used to describe this aspect of HTA (Charlton et al., 2023). We use the term to denote the ethical dimension of HTA itself, concerning ideas and decisions regarding how one ought to conduct assessments of health technology, and the implications these have for conclusions on the value of health technology (and how one ought to use them). 'Normative' refers to standards or norms that prescribe how HTA practitioners should do their work, often based on principles of morality and societal expectations. That HTA is normative implies that there are underlying ethical considerations that influence the assessment process and its outcomes, and these considerations may impact decisions regarding the adoption, use, or allocation of health technology.

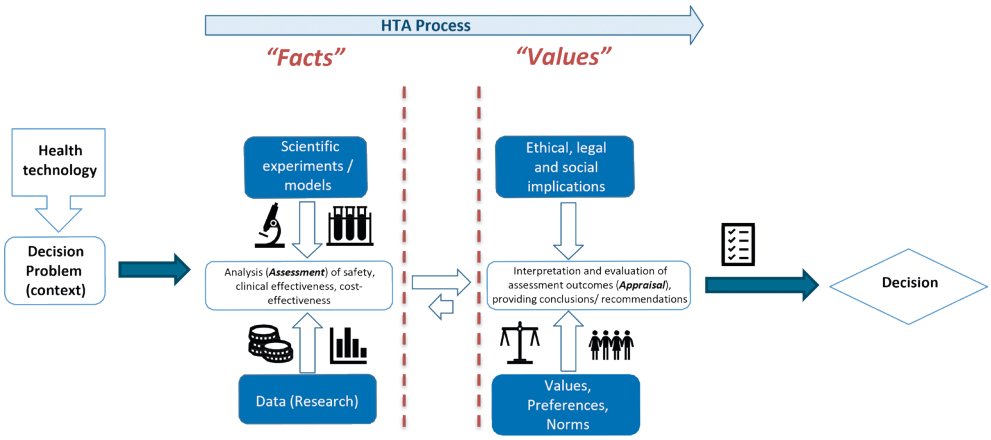


Figure 1. Conventionally, HTA is seen as consisting of a phase in which *factual information* about value dimensions of health technology is evaluated ('assessment'), followed by a phase in which a committee, comprising experts and stakeholders, evaluates the factual information to draw conclusions or make recommendations ('appraisal'). In this view, only the appraisal phase involves normative judgments and is the responsibility of those with the authority of making decisions or recommendations (adapted from (Oortwijn et al., 2022)).

As Hofmann et al noted: "*Many of the value judgments are implicit or tacit, and, by not making them explicit, the illusion of scientific objectivity and neutrality is reinforced*" (Hofmann et al., 2014). A major argument in favor of avoiding or excluding normativity is that it reduces the risk that assessments are influenced by political or partial interests (i.e., skewed to conclusions that decision-makers or stakeholders would prefer), enabling HTA to provide *impartial* information on the *public value* of health technology (Boothe, 2021; Ducey et al., 2017; Sandman & Heintz, 2014). This impartiality of HTA has also been instrumental to legitimizing the role of HTA in public decision-making (Syrett, 2016; Torgersen, 2019).

While it is understandable that HTA practitioners try to confine their contribution to decision-making to the boundaries of their (scientific) expertise, it is questionable whether they do this in practice. Increasingly, social scientists and philosophers highlight that presumably *neutral* scientific 'facts' are always a result of decisions on what is relevant to study and ideas on how that can be done, both presupposing certain values (Putnam, 2002). Decisions on how to perform a study, methods used, and interpretations of data all implicitly invoke value judgments (Alexandrova, 2016; Alexandrova & Fabian, 2022; Douglas, 2009; Kitcher, 2011; Pamuk, 2021).

AIM OF THIS THESIS

This thesis aims to explore the normativity of HTA, with a focus on how norms and evidence become entangled in its practice. By 'norms' we refer to the set of norms and principles that are tacitly understood within HTA as the right way of assessing and interpreting the value of health technologies, including their relevance, feasibility and appropriateness (Lehoux et al., 2009).

The idea of an *entanglement* of norms and evidence poses a challenge to mitigating the influence of the normativity of HTA: if normative presumptions are constitutive in the generation and interpretation of evidence it becomes impractical to separate the tasks of deciding upon the scope of an HTA, collecting evidence, and drawing conclusions based on the evidence.

A challenge in explicating the normativity of HTA is that it often remains invisible. Norms may refer to general desirable features of health technology and are therefore taken for granted (Lehoux et al., 2009). Norms may not be explicated because there is no awareness of them, they are intertwined with methods and evidence, or their explication may be avoided to upheld the illusion of objectivity (Hofmann, 2014; Van Oudheusden et al., 2019). Additionally, the language used to describe normative features of HTA is often unclear or ambiguous, making it difficult to openly discuss it (Charlton et al., 2023).

Therefore, to improve our understanding of the normative aspects of HTA and address associated challenges, this thesis addresses the following research questions:

- How can the normativity of HTA be understood and made visible?
- What is the influence of this normativity on the procedures and methods used in HTA?
- What is the influence of this normativity on conclusions of assessments?

Findings from this thesis will result in a better understanding of the different types, and extent, of normativity in HTA, and how normative considerations influence the assessment process and its conclusions. The overall aim is to contribute to developing ways for making the normativity of HTA visible and integrate normative analysis in its practice, allowing room for the consideration of broader value questions (besides cost-effectiveness) in HTA.

THESIS OUTLINE

In **Chapter 2**, based on an analysis of the literature on normativity in HTA and a case study (assessment of the Non-Invasive Prenatal Test, NIPT), we present the hypothesis that its normativity can be understood as a result of inevitable decisions made in the conduct of an assessment. These decisions (e.g., decisions on the scope, evidence requirements, and presentation of conclusions of an assessment) *commits* the HTA practitioner to moral (regarding what makes a health technology desirable), ontological (regarding which effects of health technology are conceivable), and epistemological (regarding how to obtain reliable information on effects of health technology) norms.

In **Chapter 3**, using document analysis, we reconstruct the choices and arguments made in an HTA of NIPT that was produced by an HTA agency in the Netherlands. We show how this assessment involved evaluating *mixed claims*: causal claims, regarding potential consequences of NIPT, in which empirical information becomes entangled with normative presuppositions that are necessary to define desirable outcomes of NIPT.

In **Chapter 4**, we explore how normative commitments shape the procedures and methods used by HTA practitioners for conducting assessments of medical devices. Using an online survey, we map the landscape of HTA processes for medical devices. By conducting interviews with HTA practitioners (working at HTA agencies), with a focus on the case Transcatheter aortic valve implantation (TAVI), we obtain an understanding of the choices they make in conducting assessments and how these are informed by their views about appropriate methodology and the role of HTA. We show that the procedures and methods used for assessing medical devices are still shaped by epistemic norms developed for assessing drugs, impeding the adoption of new methodology proposed for assessing medical devices.

In **Chapter 5**, we perform a mixed-methods study comparing the use of a standardized instrument to measure *capabilities* (ICEpop CAPability measure for Adults, ICECAP-A), interviews with patients, and a standard rehabilitation outcome measure (Canadian Occupational Performance Measure, COPM) in evaluating the impact of rehabilitation on persons with facioscapulohumeral dystrophy (FSHD) or myotonic dystrophy type 1 (DM1).

The capability approach offers an alternative way of measuring the impact of health technology on the *quality of life* of patients, stating that it is the effect of health technology on the opportunity of patients to be or do what they have reason to value

(their *capabilities*) that should be the informational basis for decision-making. This contrasts with the concept of *utility* that guides the QALY measure currently used to quantify quality of life. Utility, in health economics, is the measure of the preferences that individuals (patients or the general population) have for particular health states, leading to a number between 0 (representing death) and 1 (perfect health). This utilitarian approach has the normative starting point that it is preference satisfaction that matters when estimating the quality of life associated with different health states (Ubels, 2021).

The use of the capability approach in HTA is still in development, and our mixed-methods study contributes to our understanding of how to operationalize it in the context of HTA and what the influence is of different ways of operationalizing it. This could also be a basis for making comparisons with utility-based measures to obtain more empirical information about the influence of normative choices in evaluating impact of health technology on quality of life.

In **Chapter 6**, we conclude with summarizing and integrating our main findings and describing our contribution to understanding the normativity of HTA, focusing on the observed mechanisms by which norms and evidence become entangled. We also discuss the implications of our findings for HTA practice and provide recommendations for integrating normative analysis into HTA.

REFERENCES

Alexandrova, A. (2016). Can the Science of Well-Being Be Objective? *The British Journal for the Philosophy of Science*, *69*(2), 421-445. https://doi.org/10.1093/bjps/axw027

Alexandrova, A., & Fabian, M. (2022). Democratising Measurement: or Why Thick Concepts Call for Coproduction. *European Journal for Philosophy of Science*, *12*(1), 1-23. https://doi.org/10.1007/s13194-021-00437-7

Angelis, A., Lange, A., & Kanavos, P. (2018). Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. *The European Journal of Health Economics*, *19*(1), 123-152. https://doi.org/10.1007/s10198-017-0871-0

Bellemare, C. A., Dagenais, P., S, K. B., Beland, J. P., Bernier, L., Daniel, C. E., Gagnon, H., Legault, G. A., Parent, M., & Patenaude, J. (2018). Ethics in Health Technology Assessment: A Systematic Review. *International Journal of Technology Assessment in Health Care*, *34*(5), 447-457. https://doi.org/10.1017/S0266462318000508

Berloznik, R., Casert, R., van Est, R., Enzing, C., van Lieshout, M., & Versleijen, A. (2006). *Technology assessment on converging technologies: literature study and vision assessment*. Brussels: European Parliament

Boothe, K. (2019). "Getting to the Table": Changing Ideas about Public and Patient Involvement in Canadian Drug Assessment. *Journal of Health Politics, Policy and Law*, *44*(4), 631-663. https://doi.org/10.1215/03616878-7530825

Boothe, K. (2021). (Re)defining legitimacy in Canadian drug assessment policy? Comparing ideas over time. *Health Economics, Policy and Law*, *16*(4), 424 - 439. https://doi.org/10.1017/S1744133121000013

Charlton, V., DiStefano, M., Mitchell, P., Morrell, L., Rand, L., Badano, G., Baker, R., Calnan, M., Chalkidou, K., Culyer, A., Howdon, D., Hughes, D., Lomas, J., Max, C., McCabe, C., O'Mahony, J. F., Paulden, M., Pemberton-Whiteley, Z., Rid, A., . . . Wester, G. (2023). We need to talk about values: a proposed framework for the articulation of normative reasoning in health technology assessment. *Health Economics, Policy and Law*, *19*(2), 153 - 173. https://doi.org/10.1017/S1744133123000038

Daniels, N., & van der Wilt, G. J. (2016). Health Technology Assessment, Deliberative Process, and Ethically Contested Issues. *International Journal of Technology Assessment in Health Care*, *32*(1-2), 10-15. https://doi.org/10.1017/S0266462316000155

DeJean, D., Giacomini, M., Schwartz, L., & Miller, F. A. (2009). Ethics in Canadian health technology assessment: a descriptive review. *International Journal of Technology Assessment in Health Care*, *25*(4), 463-469. https://doi.org/10.1017/S0266462309990390

Douglas, H. E. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press. http://www.jstor.org/stable/j.ctt6wrc78

Ducey, A., Ross, S., Pott, T., & Thompson, C. (2017). The moral economy of health technology assessment: an empirical qualitative study. *Evidence & Policy*, *13*(1), 7-27. https://doi.org/10.1332/174426415X14443053123024

Ekmekci, P. E., & Guner, M. D. (2019). Evaluation of Ethical Analyses in Seven Reports from the European Network for Health Technology Assessment. *International Journal of Technology Assessment in Health Care*, 1-7. https://doi.org/10.1017/S0266462319000485

Fontrier, A. M., Visintin, E., & Kanavos, P. (2022). Similarities and Differences in Health Technology Assessment Systems and Implications for Coverage Decisions: Evidence from 32 Countries. *PharmacoEconomics - Open*, *6*(3), 315-328. https://doi.org/10.1007/s41669-021-00311-5

Hofmann, B. (2014). Why not integrate ethics in HTA: identification and assessment of the reasons. *GMS Health Technology Assessment*, *10*.

Hofmann, B., Cleemput, I., Bond, K., Krones, T., Droste, S., Sacchini, D., & Oortwijn, W. (2014). Revealing and acknowledging value judgments in health technology assessment. *International Journal of Technology Assessment in Health Care*, *30*(6), 579-586. https://doi.org/10.1017/S0266462314000671

Hofmann, B. M. (2015). Too much technology. *BMJ (Clinical research ed.)*, *350*(h705). https://doi.org/10.1136/bmj.h705

Kitcher, P. (2011). *Science in a Democratic Society*. Prometheus Books.

Kleinhout-Vliek, T., de Bont, A., & Boer, B. (2021). Necessity under construction - societal weighing rationality in the appraisal of health care technologies. *Health Economics, Policy and Law*, *16*(4), 457-472. https://doi.org/10.1017/S1744133120000341

Legault, G. A., Gagnon, H., Parent, M., Bellemare, C. A., Beland, J. P., Kocsis-Bedard, S., Bernier, L., Dagenais, P., Daniel, C. E., & Patenaude, J. (2021). Integration of ethical considerations into HTA reports: an analysis of integration levels using a systematic review. *International Journal of Technology Assessment in Health Care*, *37*(1), e61. https://doi.org/10.1017/S0266462321000325

Lehoux, P. (2006). *The problem of health technology*. Routledge. https://doi.org/https://doi.org/10.4324/9781315810782

Lehoux, P., Daudelin, G., Demers-Payette, O., & Boivin, A. (2009). Fostering deliberations about health innovation: what do we want to know from publics? *Social science & medicine*, *68*(11), 2002-2009. https://doi.org/10.1016/j.socscimed.2009.03.017

Lehoux, P., & Williams-Jones, B. (2007). Mapping the integration of social and ethical issues in health technology assessment. *International Journal of Technology Assessment in Health Care*, *23*(1), 9-16. https://doi.org/10.1017/S0266462307051513

Loblova, O. (2016). Three worlds of health technology assessment: explaining patterns of diffusion of HTA agencies in Europe. *Health Economics, Policy and Law*, *11*(3), 253-273. https://doi.org/10.1017/S1744133115000444

Moors, E., & Peine, A. (2016). Valuing Diagnostic Innovations: Towards Responsible Health Technology Assessment. In M. Boenink, H. van Lente, & E. Moors (Eds.), *Emerging Technologies for Diagnosing Alzheimer's Disease: Innovating with Care* (pp. 245-261). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-54097-3_13

O'Rourke, B., Oortwijn, W., Schuller, T., & International Joint Task, G. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, *36*(3), 187-190. https://doi.org/10.1017/S0266462320000215

Oortwijn, W., Sampietro-Colom, L., & (Eds.). (2022). *The VALIDATE handbook. An approach on the integration of values in doing assessments of health technologies* (W. Oortwijn & L. Sampietro-Colom, Eds.). Radboud University Press. https://doi.org/10.54195/CKHB1659

Otto, I., Kahrass, H., & Mertz, M. (2021). "Same same but different"? On the questionable but crucial differentiation between ethical and social aspects in health technology assessment. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *164*, 1-10. https://doi.org/10.1016/j.zefq.2021.05.007

Pamuk, Z. (2021). COVID-19 and the Paradox of Scientific Advice. *Perspectives on Politics*, *20*(2), 562-576. https://doi.org/10.1017/s1537592721001201

Putnam, H. (2002). *The collapse of the fact/value dichotomy and other essays*. Harvard University Press.

Refolo, P., Sacchini, D., Brereton, L., Gerhardus, A., Hofmann, B., Lysdahl, K. B., Mozygemba, K., Oortwijn, W., Tummers, M., van der Wilt, G. J., Wahlster, P., & Spagnolo, A. G. (2016). Why

is it so difficult to integrate ethics in Health Technology Assessment HTA The epistemological viewpoint. *European Review for Medical and Pharmacological Sciences*, *20*, 4202-4208.

Reuzel, R. P. B. (2001). *Health technology assessment and interactive evaluation: different perspectives* [Doctoral thesis, Radboud University]. Nijmegen.

Richardson, H. S. (2016). Intelligence and Transparency in Health Technology Assessment. *International Journal of Technology Assessment in Health Care*, *32*(1-2), 3-9. https://doi.org/10.1017/S0266462316000106

Sandman, L., & Heintz, E. (2014). Assessment vs. appraisal of ethical aspects of health technology assessment: can the distinction be upheld? *GMS Health Technology Assessment*, *10*. https://doi.org/https://doi.org/10.3205/hta000121

Savulescu, J., & Bostrom, N. (2009). *Human enhancement*. OUP Oxford.

Sorenson, C., Drummond, M., & Khan, B. B. (2013). Medical technology as a key driver of rising health expenditure: disentangling the relationship. *ClinicoEconomics and outcomes research: CEOR*, *5*, 223 - 234. https://doi.org/10.2147/CEOR.S39634

Syrett, K. (2016). Expanded HTA, Legitimacy and Independence Comment on "Expanded HTA: Enhancing Fairness and Legitimacy". *International Journal of Health Policy and Management*, *5*(9), 565-567. https://doi.org/10.15171/ijhpm.2016.75

Teerawattananon, Y., Painter, C., Dabak, S., Ottersen, T., Gopinathan, U., Chola, L., Chalkidou, K., & Culyer, A. J. (2021). Avoiding health technology assessment: a global survey of reasons for not using health technology assessment in decision making. *Cost Effectiveness and Resource Allocation*, *19*(1), 62. https://doi.org/10.1186/s12962-021-00308-1

Torgersen, H. (2019). Three myths of neutrality in TA - How different forms of TA imply different understandings of neutrality. *Technological Forecasting and Social Change*, *139*, 57-63. https://doi.org/10.1016/j.techfore.2018.06.025

Ubels, J. (2021). The assessment of value in health economics: utility and capability. In Schildmann, Buch, & Zerth (Eds.), *Defining the value of medical interventions* (pp. (69 - 81). Kohlhammer.

van der Wilt, G. J., Oortwijn, W., & Consortium, V.-H. (2022). Health technology assessment: A matter of facts and values. *International Journal of Technology Assessment in Health Care*, *38*(1), e53. https://doi.org/10.1017/S0266462322000101

Van Oudheusden, M., Delvenne, P., & Lucivero, F. (2019). Making the invisible visible. *TATuP Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis*, *28*(1), 21-26. https://doi.org/10.14512/tatup.28.1.21

Walley, T. (2007). Health technology assessment in England: assessment and appraisal. *The Medical Journal of Australia*, *187*(5), 283-285. https://doi.org/10.5694/j.1326-5377.2007.tb01244.x

Wetenschappelijke Raad voor het Regeringsbebeleid. (2021). *Kiezen voor houdbare zorg. Mensen, middelen en maatschappelijk draagvlak* (WRR-Rapport 104). Den Haag: WRR

WHO. (2021). *Health Technology Assessment Survey 2020/21. Main findings.* https://www.who.int/data/stories/health-technology-assessment-a-visual-summary

World Health Organization. (1995). *Constitution of the world health organization*. World Health Organization.


2

Chapter 2

Understanding the normativity of Health Technology Assessment: Ontological, Moral, and Epistemological commitments

Bloemen, B., Oortwijn, W., & van der Wilt, G. J. (2024). Understanding the Normativity of Health Technology Assessment: Ontological, Moral, and Epistemological Commitments. *Health Care Analysis*, 1-17. https://doi.org/10.1007/s10728-024-00487-x

ABSTRACT

The inherent normativity of HTA can be conceptualized as a result of normative commitments, a concept that we further specify to encompass moral, epistemological and ontological commitments at play in the practice of HTA. Based on examples from literature, and an analysis of the example of assessing Non-Invasive Prenatal Testing (NIPT), we will show that inevitable normative decisions in conducting an assessment commits the HTA practitioner to moral (regarding what makes a health technology desirable), ontological (regarding which effects of health technology are conceivable), and epistemological (regarding how to obtain reliable information about health technology) norms. This highlights and supports the need for integrating normative analysis and stakeholder participation, providing guidance to HTA practitioners when making normative choices. This will foster a shared understanding between those who conduct, use, or are impacted by assessments regarding what are conceivable and desirable outcomes of using health technology, and how to collect reliable information to assess whether these outcomes are (going to be) realized. It also provides more insight into the implications of different normative choices.

INTRODUCTION

Health Technology Assessment (HTA) *informs* decision-making on health technology *to promote* an equitable, efficient, and high-quality health system (O'Rourke et al., 2020). Hence, HTA claims to *improve* decision-making *by providing and assessing information* on (un)intended consequences of health technology. This aim makes HTA inherently normative (Charlton et al., 2023; Hofmann et al., 2018; Legault et al., 2018). This normativity of HTA has been described before, and HTA practitioners (i.e., those responsible for conducting assessments, comprising scoping the research question; collecting, synthesizing and appraising evidence; and reporting findings and implications) increasingly acknowledge that their practice has normative aspects (Gagnon et al., 2020; Hofmann et al., 2018).

Despite widespread recognition of HTA's normativity, the adoption of approaches to address it, such as integrating ethical analysis to clarify value judgments made in HTA, remains rather limited (Bellemare et al., 2018). A recent paper by Charlton et al suggests that this results from a lack of exactness and consistency in the language used to articulate HTA's normative aspects (Charlton et al., 2023). According to the authors, this leads to ambiguities in approaches to address normativity, undermining their ability to make normative reasonings explicit and scrutinize substantive rationales underlying HTA. Therefore, they propose a conceptual framework for articulating normative aspects of HTA.

We agree with Charlton et al that HTA is inherently normative, that this normativity is often left implicit or underspecified, and that addressing this can enhance legitimacy of HTA. We propose to extend their framework by explicating their concept of *normative commitments,* further specifying it to encompass different types of normativity.

In this paper, we analyze different examples from literature, and provide an in-dept analysis of Non-Invasive Prenatal Testing (NIPT) to illustrate that the normativity of HTA can be understood as resulting from commitments regarding what are *conceivable* and *desirable* outcomes of using a health technology, and *how to gather reliable information* to establish whether these outcomes are (going to be) realized. These normative commitments guide HTA practitioners in making methodological decisions, *committing* them to underlying *moral, epistemological,* and *ontological* norms. Before presenting our analysis of normative commitments, we will briefly describe the complexity and importance of explicating normative aspects of HTA.

MAKING THE INVISIBLE VISIBLE [2]

…By having a language that describes normative aspects of HTA

Normative aspects of HTA are often unarticulated and therefore hidden for all involved stakeholders (Charlton et al., 2023). HTA practitioners employ (in)formal norms that specify abstract values worth pursuing, in measuring and evaluating health technology's impact. While guidelines may codify these norms, individual cases may require re-specifications or difficult trade-offs. Moreover, the relation between norms and their underlying values is often left implicit, obscuring their normative nature. It would enhance the legitimacy of HTA if this *invisible* normativity would be made *visible*. We agree with Charlton et al that a clear language is needed to explicate this normativity, and propose to specify their concept of *normative commitments* to show that it encompasses various types of normativity in HTA and acknowledge their complex interrelations (Charlton et al., 2023).

…That acknowledges the entanglement of norms and evidence

Charlton et al argue that judgements about what is valuable ('*ethical judgments*') operate *alongside* judgments about what one ought to believe given the available information ('*evidential judgments*'). According to them, judgments on meaning and quality of evidence do not *determine* recommendations and decisions. Despite this being true in principle, evidential judgments, such as concluding whether a drug is effective based on available evidence, may have important implications that HTA bodies and/or decision-makers may be required to consider when making recommendations or decisions (Fontrier et al., 2022; Janoudi et al., 2016; WHO, 2021).

Charlton et al also state that evidential judgements incorporate normative commitments which can ground knowledge claims (Charlton et al., 2023). By noticing this, they acknowledge a certain entanglement of norms and evidence in HTA. However, norms do not only play a role in evidence interpretation and judgements on its validity (e.g., thresholds for statistical significance), but also in evidence selection (what is considered evidence) and evidence generation (e.g., outcome measures used in HTA incorporate values by how they are measured) (Schroeder, 2016).

This entanglement of norms and evidence is not uncommon in HTA. When evaluating the effectiveness of health technology an implicit connection is made between a descriptive statement (e.g., 'this drug will lower blood pressure') and an evaluative premise (e.g., 'by lowering blood pressure the risk of developing cardiovascular dis-

2 This metaphor, which we think nicely describes the task at hand in addressing the normativity of HTA, was inspired by an interesting paper discussing the normativity of Technology Assessment (TA): Lucivero, F., Delvenne, P., & Van Oudheusden, M. (2019). Making the invisible visible. *TATuP*, *28*(1).

ease will be reduced, and this is a good thing to do') (Hofmann et al., 2018; Mertz et al., 2023; Stegenga, 2015). HTA evaluates such mixed claims because it relies on concepts and classifications that are descriptive *and* evaluative. For example, what is considered a disease is based on descriptions of symptoms and an interpretation of these symptoms as something that requires (medical) treatment. Another example is that measuring *health-related quality of life* (HRQoL) requires empirical information on impact of a condition on aspects of life held to be valuable (Bloemen et al., 2021; Hofmann et al., 2018). The connection between a descriptive and evaluative premise may not be obvious, may also not be intended by HTA practitioners, but it should be there to inform decision-makers on whether a health technology is a 'better' option than alternatives (Legault et al., 2018; Mertz et al., 2023). Consequently, the relation between evidence and norms in HTA is not one of independency but of entanglement (Bloemen et al., 2021; Hofmann et al., 2018; van der Wilt et al., 2022).

…And acknowledges normative features of health technology

Moreover, social science has shown that health technology is not neutral. Using health technology often requires changes in how we do or organize healthcare practice; shapes the way we treat and think about health and disease; and creates new situations that challenge or demand rethinking of existing norms (Lehoux, 2006). These normative features of health technology should be accounted for in HTA (Boenink, 2012; Giacomini et al., 2013; Smits et al., 2022). For example, when assessing the *effectiveness* of a health technology, it is important to consider that its ability to realize the intended effect depends on the context in which it is applied, but also that the intended effect itself could be different by re-designing the technology to realize another (valuable) effect. Because there may be disagreement about the desirable effect, assessing a given configuration of a technology may lead to assessing a health technology on effects deemed undesirable by some stakeholders. Therefore, HTA also needs to conceptualize *how* certain outcomes *can be realized* by using a health technology. These *ontological assumptions* are also part of the (implicit) normativity of HTA [3].

…And reflects the role of HTA in decision-making

Finally, because of the role of HTA in decision-making, it can be used to prioritize (a particular use of) certain health technologies (Lehoux, 2006). Therefore, HTA is an actor in decision-making, and its practitioners have a responsibility in providing justifications for this role. It can certainly not be expected from HTA practitioners alone to identify and justify all normative choices guiding HTA, but by being part of its practice they are *committed* to underlying rationales.

3 Ontology refers to the set of categories used to describe the nature of objects, their relations, and phenomena held to exist.

Considering this broader understanding of the normativity of HTA, making it visible is one aspect, but it is also important to acknowledge that normative commitments are *jointly produced in assessments* (Delvenne & Parotte, 2019). Therefore, a language that helps anyone involved in recognizing and explicating this normativity would be helpful, but it should encompass *all types of normativity* involved, their interrelatedness, and its actual use needs a consideration of *who is involved*.

NORMATIVE COMMITMENTS IN HTA

We define **normative commitments** as [4]: "*any commitment to a norm that guides, and is further specified in, the assessment of information about the properties, effects or impacts of health technology. Different types of norms jointly provide justifications for conclusions about whether, and how, health technologies should be used. Commitments to these norms can be implicit but are shown by a willingness to deploy and defend them in justifying the rationales underlying an assessment*".

Our concept of normative commitments highlights that every assessment is an expression of normative commitments of those involved in the process. There are three types of normative commitments, regarding how the world ought to be (*moral commitment*), which role health technology could play in realizing this (*ontological commitment*), and which information reliably indicates that intended outcomes are (going to be) realized (*epistemological commitment*). We will further define and illustrate these types of normative commitments and their implications for HTA by providing examples from literature. We start with a brief description of moral commitments and provide more extensive descriptions of epistemological and ontological commitments because moral commitments are already extensively discussed in literature on normativity of HTA (Charlton et al., 2023; Hofmann et al., 2018; Legault et al., 2018). Although the different types of normative commitments are introduced separately, in practice there will be many dependencies among them as discussed in the next section.

4 With 'normative' we refer to ideas about how things *should be* or how people should behave. These ideas can be formalized by norms and are expressed by normative judgments or acting in accordance with certain norms. In the context of HTA, 'normative' refers to ideas, norms, and standards that are tacitly understood within its practice as the right way to conduct assessments of health technology (Lehoux et al., 2009). There are different types of norms that share a prescriptive and evaluative function, and moral norms (value judgments) are a subset of this (Reiss, 2017). For example, consider the difference between 'this is a fair distribution of healthcare resources' and 'this body of evidence can be considered reliable'. Both express an evaluation, but the goal of the evaluation is different, i.e., one prescribes that you ought to consider a particular distribution of resources as morally acceptable, whereas the other prescribes that you ought to consider particular evidence as truthful. In HTA, because it informs public decision-making that has moral consequences, the justification of different norms always includes references to moral norms.

Moral commitments in HTA

Moral commitments are "*commitments to norms concerning what is desirable and acceptable*". In HTA, these norms guide which effects of health technology are considered important based on commitments to values such as avoiding harm or promoting well-being, and norms concerning what constitutes legitimate decision-making (e.g., transparent and inclusive processes).

In HTA, not all conceivable effects of health technology are assessed, the focus is on consequences that matter to us. Moral considerations guide distinctions between beneficial and harmful consequences to assess desirable qualities of health technology (Legault et al., 2018). For example, cost-effectiveness analyses often use quality-adjusted life years (QALYs) to measure benefits which entails a commitment to maximizing HRQoL for recipients of care. However, alternative conceptions of the good that include broader aspects of well-being (beyond health) or benefits beyond recipients of care may also inform such analyses (Coast et al., 2008; Engel et al., 2021; Wilson, 2023).

Epistemological commitments in HTA

Epistemological commitments are "*commitments to norms concerning the nature of knowledge, how it can be acquired, and how its reliability can be evaluated and justified*". In HTA, these norms define which types of information can support conclusions on effects of health technology, based on commitments to theories about what is reliable knowledge. Where moral commitments guide what we consider desirable consequences of health technology, epistemological commitments concern what can be reliably stated about these consequences. For example, although a shared moral commitment to maximizing HRQoL can lead us to consider the QALY as outcome measure, diverging epistemological commitments can lead to different ideas about how QALYs should be estimated (e.g., is a patient a reliable judge on his or her quality of life? Can quality of life be reliably quantified?).

Important normative epistemological questions in HTA concern: *who decides what forms of information count as appropriate evidence in HTA*, and *who is recognized as a credible source of knowledge or information*? (Moes et al., 2020; Staniszewska & Soderholm Werko, 2021). Although the epistemology of HTA (aligned with principles of evidence-based medicine) does not exclude consideration of (qualitative) experiential knowledge, HTA agencies prefer quantitative types of evidence (Moes et al., 2020; Staniszewska & Soderholm Werko, 2021; Szabo et al., 2024). Also, when assessing impact on quality of life, outcome measures constructed without patient involvement are used (Wiering et al., 2017).

The link between who is involved in evidence-making processes and fair decision-making has drawn recent attention in discussions about evidence use in HTA (Michaels, 2020; Moes et al., 2020). HTA agencies have also been challenged, even taken to court, regarding their interpretation and use of evidence, mainly taking from RCTs. For example, the Dutch National Health Care Institute was sued for misconduct against interstitial cystitis patients because, according to the patient's association, it attributed *too little credibility* to the experience of doctors and patients concerning potential effects of bladder instillations (Moes et al., 2017). The exclusion of experiences and perspectives as an (unintended) result of epistemic norms upheld by HTA agencies is also reported in studies that evaluated stakeholder participation at HTA agencies. Because experiential knowledge contributed by patients and clinicians does not fit prevailing ideas of robust evidence, their ability to engage in discussions on HTA outcomes is limited (Hashem et al., 2018; Lips et al., 2022; Steffensen et al., 2022).

The role of epistemological commitments in HTA is also evident in discussions on the use of real-world data (RWD). Despite scholars highlighting inherent limitations of RCTs such as generalizability to real-life clinical contexts, and situations in which RCTs are difficult to conduct (e.g., due to small population sizes or characteristics of a technology), the validity of RWD is debated (Makady et al., 2017). Even when consideration of RWD in assessments is required by law, some HTA practitioners assign a low credibility to it (effectively excluding it from impacting recommendations) (Makady et al., 2017). However, in some cases (e.g., orphan diseases, end-of-life treatments) HTA agencies and practitioners are willing to consider other types of evidence or thresholds, based on overriding moral commitments (Stafinski et al., 2022).

Ontological commitments in HTA

Ontological commitments are "*commitments to norms regarding the nature of reality, the existence of certain entities or phenomena, and plausible causal mechanisms*". In HTA, these are ideas about which types of interventions could impact health, and which effects are conceivable, based on theories about mechanisms of health and disease, the working mechanisms of technology, and the organization of healthcare. Where moral commitments concern what are desirable consequences of health technology, and epistemological commitments guide ways of finding out whether these consequences do happen, ontological commitments concern what could happen. For example, when estimating QALYs it is important to consider what constitutes HRQoL and identify aspects changeable by technology. This involves defining the nature of health (e.g., from a view that health is the absence of disease a health technology can improve HRQoL by eliminating symptoms and disabilities, whereas from a conception of health as the ability to experience a state of well-being a health

technology can also improve HRQoL by improving social relations or living conditions) (Stegenga, 2015; Wilson, 2023).

An implicit ontology also guides the organization of HTA processes and the conduct of assessments.

Firstly, the decision problem in HTA is framed as a question concerning the use of a health technology. This is reflected in the assigned remit of HTA agencies that is often linked to a single class of technologies (e.g., drugs, medical devices), and in the scope of assessments (e.g., relative effectiveness of a *single* technology compared to a standard of care) (Fontrier et al., 2022; Richardson et al., 2023). Evaluating health technologies individually may neglect undesirable cumulative effects, such as shifts in societal norms regarding responsibility for one's own health due to the omnipresence of health checks (e.g., screening programs, diagnostic tests) (Stol et al., 2016).

Secondly, outcome measures used in HTA, like the QALY, embed ontological presumptions. Measuring QALYs relies on questionnaires by which patients can report the impact of their health condition on aspects of life deemed important for evaluating *quality of life*. This not only implies a moral commitment to which aspects of life are desirable, but also what *constitutes* quality of life. Although there are theories that focus on other constituents of quality of life, HTA has predominantly relied on a utilitarian conception (Kennedy-Martin et al., 2020). The use of instruments that are based on different conceptualizations (e.g., the capability approach) may lead to different assessments and priorities in healthcare (Mitchell et al., 2015).

Finally, health technology may imply something about the nature of a health problem. Increasingly, technologies are used in diagnosing and treating people, distinguishing between healthy and unhealthy, making visible biological processes and functions held to be constitutive of a disease. These are not just ontological classifications, but also invoke ideas about what is normal and desirable. For example, cochlear implants imply, by aiming to restore normal hearing, that deafness is a disability that should be resolved; and interventions for autism suggest that its associated behavior is not part of normal functioning (Oortwijn et al., 2022). Assessing such technologies on intended effects, without questioning the underlying problem definition, risks evaluating them on outcomes that are unacceptable to certain stakeholders.

THE INTERRELATEDNESS OF NORMATIVE COMMITMENTS IN HTA

While our description above intentionally separated normative commitments for clarity, in practice they are interdependent and may align or be in conflict. For example, there is a broadly shared moral commitment to involve those that are potentially impacted by an HTA informed decision, creating tensions with epistemological commitments to only consider reliable evidence, excluding experiential knowledge as formal evidence. It is challenging to resolve these tensions. Realizing that the epistemology of HTA is grounded in moral commitments (i.e., objective knowledge is needed to evaluate the *public* value of health technology) may be a constructive way forward (Ducey et al., 2017). It acknowledges that *epistemic acts* (generating and interpreting evidence) do not merely capture aspects of reality but also shape reality (Wehrens & de Graaff, 2024).

Establishing an a priori hierarchy for balancing commitments is challenging, as commitments need case-specific specifications. However, articulating normative commitments during an assessment can reveal their connections and facilitate coherence.

An example of how articulating different commitments can facilitate coherence is an HTA conducted by the Finish Council for Choices in Health Care (COHERE) (Saarni et al., 2022). Staff of COHERE was asked to conduct an assessment to give recommendations to the Ministry of Social Affairs and Health about public funding of medical treatments for gender dysphoria. Their approach was to integrate ethical analysis in every step of the assessment (scoping, evidence gathering, and appraisal), by embedding ethicists in a multidisciplinary HTA team and organizing stakeholder hearings. This led to an identification of different normative issues and revealed relations between different commitments:

- Moral commitments to autonomy (people should be able to decide a gender identity themselves) are related to epistemological commitments regarding high quality of evidence because autonomous decision making requires having reliable information on life-long consequences of choices.
- Ontological commitments to classifications of gender dysphoria (distinguishing between people with transgender identity, people with a nonbinary gender identity, and minors with variations in gender identity) potentially conflict with a moral commitment to equity (why should subgroups receive different treatment?).

Bringing these normative considerations together led to a recommendation for psychosocial counseling and being cautious about offering permanent treatments to ado-

lescents (considering the uncertainty in predicting long-term outcomes which also impedes autonomy), and highlighted equity issues regarding the (currently) different treatment of people with transgender or non-binary gender identity.

We will further demonstrate the usefulness of our concept of normative commitments in explicating normative aspects in the assessment of NIPT.

EXAMPLE: ASSESSMENT OF NIPT

In this example, HTA informs a decision on implementing NIPT in a national prenatal screening program. NIPT is a procedure that analyses fetal DNA in the mother's blood to obtain information about the fetal genotype (Gadsboll et al., 2020). The procedure only requires a blood sample of the mother and does not pose risks of a miscarriage associated with other (invasive) prenatal tests (e.g., amniocentesis, chorionic villus sampling). It offers other potential advantages such as early testing (around 10th week of pregnancy) and more reliable and comprehensive information on genetic conditions of the unborn child.

Given NIPT's diverse potential uses in prenatal screening, decisions are needed concerning its implementation, e.g., offer it commercially or within a national screening program, as a first line test or supplement to other prenatal tests, to all pregnant women or women with particular risks, with or without reimbursement by health insurance, and screening specific genetic conditions or providing whole genome coverage (Gadsboll et al., 2020). This multiple realizability of NIPT is mirrored by different views people may have on how NIPT should be used, which effects are (un) desirable, and how to evaluate whether these effects are (going to be) realized (Kibel & Vanstone, 2017).

The HTA practitioner's task is to demonstrate, by collecting, synthesizing, and appraising (quality of) evidence, the potential effects of different implementations of NIPT and how these compare with alternative prenatal screening options. Despite legal and methodological guidelines on how an assessment should be done, several decisions must be made.

Firstly, it should be decided whether NIPT needs an assessment and what the relevant comparators are, reflecting *prior moral and ontological commitments*. For example, the decision to assess NIPT classifies it as a *health* technology, useful in realizing *health-related* benefits. This classification is not unquestionable, as shown by critical responses that the HTA agency of Germany (the Federal Joint Committee, G-BA) received. Because the G-BA only has a mandate for assessing technologies

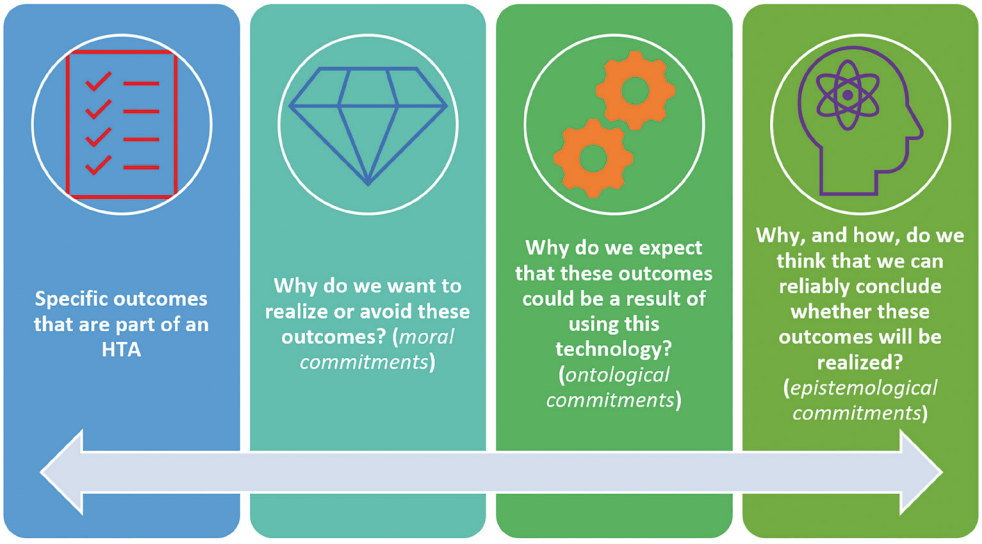
with a *medical* purpose, their decision to assess it implied a purpose of NIPT that was contested by some stakeholders (Braun & Könninger, 2018). Moreover, using prenatal screening technologies such as NIPT can be seen as an expression of *disvalue* for people with disabilities (Hofmann, 2017). It also raises ontological questions concerning whether the disabilities themselves make disabled people worse off or the societal conditions (e.g., prejudices and discrimination, lack of societal support), questioning the focus on technological solutions.

Secondly, scoping includes identifying outcomes to consider in the assessment which reflects *prior moral commitments* about what makes a health technology desirable (Mitchell et al., 2019; van der Wilt et al., 2022). In the case of NIPT, its purpose may be seen as *providing information* on genetic abnormalities, or *prevent the birth of children* with genetic abnormalities, or *enhance reproductive autonomy* (Kibel & Vanstone, 2017). These ideas about the primary purpose of NIPT implicitly guide the assessment of its safety, clinical effectiveness, and cost-effectiveness. To assess effectiveness, decisions are necessary regarding which and whose benefits to consider (e.g., only the benefits for prospective parents or also for the unborn child, which genetic anomalies are important to detect) focusing on either its ability to effectively help prospective parents making decisions or to prepare a life for the unborn child. Assessing cost-effectiveness also leads to methodological issues that imply value judgments. For example, using QALYs as an outcome would imply, depending on whether maternal or fetal QALYs are considered, that NIPT is cost-effective to the extent that it prevents a sufficient number of births affected by genetic disability or helps supporting the lives of children with disabilities (Kibel & Vanstone, 2017; Nshimyumukiza et al., 2018).

Once the assessment scope is determined, evidence collection and appraisal occur. Decisions must be made, guided by *epistemological commitments*, on the types of information to include and to what extent they constitute *evidence of* NIPT achieving specific outcomes. This may result in the exclusion of conceivable and important outcomes (e.g., changing social norms and attitudes towards people with disabilities) due to the unavailability of information that meets certain methodological criteria to be counted as evidence.

To conclude, our analysis highlights different ontological, moral, and epistemological commitments that jointly shape an assessment of NIPT, but also illustrates their close interrelatedness. Because HTA informs public decision-making, which can have moral consequences such as impacting availability of NIPT, a justification of using certain (even epistemological and ontological) norms will include references to moral ideas concerning justice and the good life (e.g., a particular operationalization of the QALY

will partly be justified by reference to ideas about what makes this technology acceptable). Despite this interrelatedness among types of normative commitments, there are relevant differences. For example, although it is conceivable that NIPT leads to more terminations of pregnancies, and reliable data about this effect can be obtained, some HTA agencies have decided to ignore this as outcome and focus on enhancing reproductive autonomy because of overriding moral commitments. Therefore, a particular co-specification of normative commitments takes place in an assessment, which may look like the schematic example in Figure 1 (Bloemen et al., 2021).



Test characteristics of NIPT (e.g. positive predictive value)	We want prospective parents to be able to trust the information that is provided about the health status of their fetus; we want to prevent unnecessary worries or abortion of pregnancies based on misinformation	NIPT provides information in a different way than existing prenatal tests; Any genetic test can provide wrong information	We can collect data on results of NIPT and compare them with other tests and clinical information
Cost-effectiveness of NIPT	We want to ensure the financial sustainability of the health care system and equal access to health services; public resources should be spent efficiently	Implementing health technology often requires investments; resources are limited, and the introduction of new health technologies may impede the offering of other technologies	We can estimate the cost-effectiveness of NIPT by estimating its costs (based on data of costs associated with comparable tests) and its benefits (expressed in QALYs or number of genetic conditions detected)
Reproductive autonomy	We want to facilitate prospective parents in making their own choices concerning the continuation of pregnancy	The availability of information on health conditions of the unborn child helps prospective parents in making choices	We can perform qualitative research (surveys, interviews) to retrieve preferences of prospective parents and their experiences with prenatal tests

Figure 1. The role of normative commitments in assessing NIPT, grounding the use of particular outcomes.

DISCUSSION

Implications for the practice of HTA

We have argued that the normativity of HTA is a result of moral, epistemological, and ontological normative commitments that guide its practice. These commitments are often left implicit but express themselves by what HTA practitioners do. What are the practical implications of this? How can we make these normative commitments more visible? As we will argue, this requires integrating normative analysis and stakeholder participation.

Integrating normative analysis in HTA

Normative analysis should be integral to HTA, explicating and justifying the moral, epistemological, and ontological commitments underpinning the assessments, and explaining how they jointly produce a reliable, relevant, and coherent assessment of the potential value of a health technology (see Figure 2). Traditionally, assessments are divided into evaluations of different 'aspects' of health technology, e.g., safety, effectiveness, cost-effectiveness, and ethical analyses. Although they emphasize different qualities or consequences of health technology, these distinctions disguise their dependencies (e.g., improving safety may require safety measures that make using a technology more time consuming, reducing cost-effectiveness) (Mitchell et al., 2019). This division into different evaluations may also lead to conflicting results. For example, whereas ethical analysis may suggest that NIPT's purpose should be to facilitate informed choice, cost-effectiveness analysis may use outcome measures (e.g., cost per additional chromosomal abnormality detected) that frame the purpose of NIPT as preventing a sufficient number of affected births (i.e., because that makes NIPT more 'cost-effective'). Consequently, decision-makers using the outcomes of such HTA are confronted with different results that imply conflicting ideas about the desirable use of NIPT (Kibel & Vanstone, 2017).

To enhance coherence of assessments, we should clarify the moral commitments underlying different evaluations. For example, the safety of health technology is evaluated *because we are committed* to the moral principle of doing no harm, and it requires *moral distinctions* between intended effects (which improve clinical effectiveness) and undesirable side effects (threatening safety) (Oortwijn et al., 2022; van der Wilt et al., 2022). Explicating these moral commitments aids in making connections between different analyses, as safety, clinical effectiveness, and costs-effectiveness all ascribe qualities to a health technology that make a discrete contribution to its (perceived) value (Legault et al., 2018; Mitchell et al., 2019). An articulation of how these different qualities contribute to ends that are being pursued by using that technology improves coherence between different analyses (van der Wilt et al., 2017). For

example, normative analysis may show that NIPT is broadly valued for its ability to expand parental choices during the prenatal care pathway. Using outcome measures in safety, effectiveness and cost-effectiveness analysis aligned with this ability ensures coherence between these different analyses (Kibel & Vanstone, 2017).

Additionally, the ontological and epistemological commitments underlying analyses should be clarified. Clarifying why specific effects of a health technology are evaluated involves moral ideas about what is relevant as well as background ideas about what is conceivable (i.e., which effects are, given known mechanisms underlying a disease and the workings of a technology, plausible), and about how to obtain reliable information about these effects.

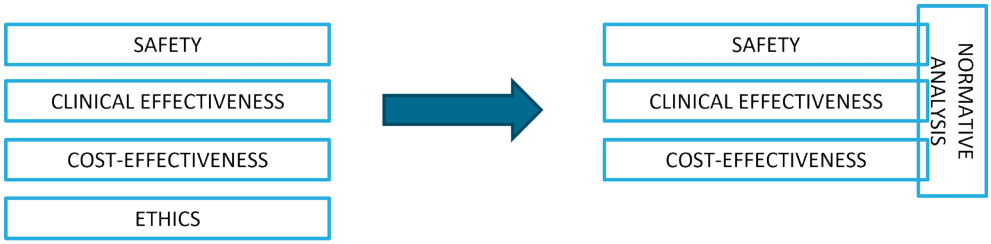


Figure 2. Instead of viewing normative analysis as a separate analysis conducted in HTA (e.g., 'ethics'), it should be seen as an integral part of assessments. Adapted from (Oortwijn et al., 2022).

Integrating and conducting this normative analysis in HTA requires embedding ethicists, with experience in health policy or HTA processes, in the HTA team. These experts could help in articulating the normative commitments at play (Hofmann et al., 2018; Refolo et al., 2020; Saarni et al., 2022).

Stakeholder participation in HTA

Because normative commitments influence *what* is assessed in HTA, and *how,* and impact public decision-making, an important question is *whose* commitments are explored in HTA. This foregrounds the importance of stakeholder participation in HTA to ensure that all relevant epistemological, ontological, and moral views are considered, and commitments are *jointly* produced. Because commitments are partly established during assessments, stakeholders should not just be consulted but participate in all phases of HTA processes. Moreover, given the often-implicit nature of normative commitments in HTA, the involvement of different perspectives may also help in making them visible. Therefore, someone's voice may be relevant because he or she is affected by outcomes of an assessment, and/or because of the enlightenment and broadening of *perspectives* such voices provide (van der Wilt et al., 2022). The focus should not be primarily on *stakeholder* involvement as such, but on ensuring

that all feasible and relevant *perspectives* are included in assessments (van der Wilt et al., 2022).

Engaging stakeholders early in HTA processes could help determining the scope of assessments, clarifying what the specific health problem is that should be resolved by a health technology, identify alternative solutions and qualities that desirable solutions should have, and how it can be established whether desirable outcomes are (going to be) realized (Oortwijn et al., 2022). This should establish from which shared moral, ontological, and epistemological commitments the assessment can proceed.

During assessment, stakeholders can have a more substantial role besides providing comments. For example, in building models for cost-effectiveness analysis, stakeholders can provide normative guidance by co-deciding on what to include in the model and how to interpret its findings, which are normative choices (Harvard & Winsberg, 2023). Conducting cost-effectiveness analyses starting from different normative premises, for example with and without assuming equivalent value of QALYs (i.e., irrespective of characteristics of patients), may provide empirical data on the sensitivity of outcomes to different normative presumptions (Luyten & van Hoek, 2021). This can either lead to a conclusion that outcomes of the HTA are invariant to different normative premises, improving its robustness, or lead to more insight into what the normative disagreements are and their impact on conclusions of HTA.

Responsibility and accountability of HTA

Making normative commitments of HTA visible is a shared responsibility among all those who are involved in, or make use of, HTA. Although the normativity of HTA is shaped by contextual factors (e.g., available time, capacity, existing laws and guidelines), HTA practitioners can make important contributions to explicating and justifying this normativity. Integrating stakeholder participation and normative analysis into HTA processes provides practitioners with normative guidance, distributing the responsibility for justifying normative choices among all involved parties.

Limitations

Our conceptualization of the normativity of HTA draws from literature and analysis of an example (NIPT) (Bloemen et al., 2021). Because NIPT is a morally challenging health technology, the generalizability of our conclusions may be limited because we have identified issues specific for NIPT. Some of the normative aspects of HTA that we have highlighted may be less salient in other health technologies (e.g., drugs and medical devices) that are mostly assessed by HTA agencies. However, as health technology always has implications for the lives of people, assessing these implications will always touch upon normative issues, invoking judgments about which benefits

represent societal value, as described previously in literature (Hofmann et al., 2018; Lehoux, 2006; van der Wilt et al., 2022). Also, in analyzing normative commitments we have drawn on literature describing normative aspects of HTAs of different types of technologies (including drugs). Therefore, we are confident that we have described normative aspects of HTA that have a generic nature.

We are aware that included literature may not fully represent the views of HTA practitioners, given also the highly contestable nature of concepts like 'value judgment', 'objectivity', 'normativity' etc., and different views about the nature and purpose of HTA, which requires further research and debate.

CONCLUSION

We have argued that HTA is a normative practice, being an expression of ontological, moral, and epistemological commitments. Our specification of normative commitments provides a conceptual tool that could facilitate integration of normative analysis and stakeholder participation in HTA, bridging the gap between normative analysis and empirical inquiry. It also explains why stakeholder involvement is of intrinsic value to HTA, beyond its instrumental role in gaining acceptance. Moreover, it integrates insights from social science, philosophy of science and technology, concerning the complex interplay of technology, evidence, values, and science in evidence-based decision-making.

REFERENCES

Bellemare, C. A., Dagenais, P., S, K. B., Beland, J. P., Bernier, L., Daniel, C. E., Gagnon, H., Legault, G. A., Parent, M., & Patenaude, J. (2018). Ethics in Health Technology Assessment: A Systematic Review. *International Journal of Technology Assessment in Health Care*, *34*(5), 447-457. https://doi.org/10.1017/S0266462318000508

Bloemen, B., Jansen, M., Rijke, W., Oortwijn, W., & van der Wilt, G. J. (2021). Mixed claims in Health Technology Assessment: The case of Non-Invasive Prenatal Testing. *Social science & medicine*, *270*. https://doi.org/10.1016/j.socscimed.2021.113689

Boenink, M. (2012). Debating the Desirability of New Biomedical Technologies: Lessons from the Introduction of Breast Cancer Screening in the Netherlands. *Health Care Analysis*, *20*(1), 84-102. https://doi.org/10.1007/s10728-011-0173-0

Braun, K., & Könninger, S. (2018). Realizing responsibility. Institutional routines, critical intervention, and the "big" questions in the controversy over non-invasive prenatal testing in Germany. *New Genetics and Society*, *37*(3), 248-267. https://doi.org/10.1080/14636778.2018.1495555

Charlton, V., DiStefano, M., Mitchell, P., Morrell, L., Rand, L., Badano, G., Baker, R., Calnan, M., Chalkidou, K., Culyer, A., Howdon, D., Hughes, D., Lomas, J., Max, C., McCabe, C., O'Mahony, J. F., Paulden, M., Pemberton-Whiteley, Z., Rid, A., . . . Wester, G. (2023). We need to talk about values: a proposed framework for the articulation of normative reasoning in health technology assessment. *Health Economics, Policy and Law*, *19*(2), 153 - 173. https://doi.org/10.1017/S1744133123000038

Coast, J., Smith, R. D., & Lorgelly, P. (2008). Welfarism, extra-welfarism and capability: the spread of ideas in health economics. *Social science & medicine*, *67*(7), 1190-1198. https://doi.org/10.1016/j.socscimed.2008.06.027

Delvenne, P., & Parotte, C. (2019). Breaking the myth of neutrality: Technology Assessment has politics, Technology Assessment as politics. *Technological Forecasting and Social Change*, *139*, 64-72. https://doi.org/10.1016/j.techfore.2018.06.026

Ducey, A., Ross, S., Pott, T., & Thompson, C. (2017). The moral economy of health technology assessment: an empirical qualitative study. *Evidence & Policy*, *13*(1), 7-27. https://doi.org/10.1332/174426415X14443053123024

Engel, L., Bryan, S., & Whitehurst, D. G. T. (2021). Conceptualising 'Benefits Beyond Health' in the Context of the Quality-Adjusted Life-Year: A Critical Interpretive Synthesis. *Pharmacoeconomics*, *39*(12), 1383-1395. https://doi.org/10.1007/s40273-021-01074-x

Fontrier, A. M., Visintin, E., & Kanavos, P. (2022). Similarities and Differences in Health Technology Assessment Systems and Implications for Coverage Decisions: Evidence from 32 Countries. *PharmacoEconomics - Open*, *6*(3), 315-328. https://doi.org/10.1007/s41669-021-00311-5

Gadsboll, K., Petersen, O. B., Gatinois, V., Strange, H., Jacobsson, B., Wapner, R., Vermeesch, J. R., Group, N. I.-m. S., & Vogel, I. (2020). Current use of noninvasive prenatal testing in Europe, Australia and the USA: A graphical presentation. *Acta Obstetricia et Gynecologica Scandinavica*, *99*(6), 722-730. https://doi.org/10.1111/aogs.13841

Gagnon, H., Legault, G. A., Bellemare, C. A., Parent, M., Dagenais, P., S, K. B., Tapin, D., Bernier, L., Beland, J. P., Daniel, C. E., & Patenaude, J. (2020). How does HTA addresses current social expectations? An international survey. *International Journal of Technology Assessment in Health Care*, *37*, e9. https://doi.org/10.1017/S0266462320000793

Giacomini, M., Winsor, S., & Abelson, J. (2013). Ethics in Health Technology Assessment: Understanding Health Technologies as Policies. *Healthcare Management Forum*, *26*(2), 72-76. https://doi.org/10.1016/j.hcmf.2013.04.004

Harvard, S., & Winsberg, E. (2023). Patient and Public Involvement in Health Economics Modelling Raises the Need for Normative Guidance. *Pharmacoeconomics*, *41*, 733-740. https://doi.org/10.1007/s40273-023-01274-7

Hashem, F., Calnan, M. W., & Brown, P. R. (2018). Decision making in NICE single technological appraisals: How does NICE incorporate patient perspectives? *Health Expectations*, *21*(1), 128-137. https://doi.org/10.1111/hex.12594

Hofmann, B. (2017). 'You are inferior!' Revisiting the expressivist argument. *Bioethics*, *31*(7), 505-514. https://doi.org/10.1111/bioe.12365

Hofmann, B., Bond, K., & Sandman, L. (2018). Evaluating facts and facting evaluations: On the fact-value relationship in HTA. *Journal of evaluation in clinical practice*, *24*, 957-965. https://doi.org/10.1111/jep.12920

Janoudi, G., Amegatse, W., McIntosh, B., Sehgal, C., & Richter, T. (2016). Health technology assessment of drugs for rare diseases: insights, trends, and reasons for negative recommendations from the CADTH common drug review. *Orphanet Journal of Rare Diseases*, *11*(1), 164. https://doi.org/10.1186/s13023-016-0539-3

Kennedy-Martin, M., Slaap, B., Herdman, M., van Reenen, M., Kennedy-Martin, T., Greiner, W., Busschbach, J., & Boye, K. S. (2020). Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. *The European Journal of Health Economics*, *21*(8), 1245-1257. https://doi.org/10.1007/s10198-020-01195-8

Kibel, M., & Vanstone, M. (2017). Reconciling ethical and economic conceptions of value in health policy using the capabilities approach: A qualitative investigation of Non-Invasive Prenatal Testing. *Social science & medicine*, *195*, 97-104. https://doi.org/10.1016/j.socscimed.2017.11.024

Legault, G.-A., Suzanne, K., Bellemare, C. A., Béland, J.-P., Bernier, L., Dagenais, P., Daniel, C.-É., Gagnon, H., Parent, M., & Patenaude, J. (2018). Revisiting the Fact/Value Dichotomy: A Speech Act Approach to Improve the Integration of Ethics in Health Technology Assessment. *Open Journal of Philosophy*, *8*(5), 578-593. https://doi.org/10.4236/ojpp.2018.85042

Lehoux, P. (2006). *The problem of health technology*. Routledge. https://doi.org/https://doi.org/10.4324/9781315810782

Lehoux, P., Daudelin, G., Demers-Payette, O., & Boivin, A. (2009). Fostering deliberations about health innovation: what do we want to know from publics? *Social science & medicine*, *68*(11), 2002-2009. https://doi.org/10.1016/j.socscimed.2009.03.017

Lips, P., Timmers, L., Bal, R., & Delnoij, D. (2022). Involvement of Patients and Medical Professionals in the Assessment of Relative Effectiveness: A Need for Closer Cooperation. *Value in Health*, *25*(9), 1480-1488. https://doi.org/10.1016/j.jval.2022.03.021

Luyten, J., & van Hoek, A. J. (2021). Integrating Alternative Social Value Judgments Into Cost-Effectiveness Analysis of Vaccines: An Application to Varicella-Zoster Virus Vaccination. *Value in Health*, *24*(1), 41-49. https://doi.org/10.1016/j.jval.2020.07.011

Makady, A., Ham, R. T., de Boer, A., Hillege, H., Klungel, O., Goettsch, W., & GetReal Workpackage 1. (2017). Policies for Use of Real-World Data in Health Technology Assessment (HTA): A Comparative Study of Six HTA Agencies. *Value in Health*, *20*(4), 520-532. https://doi.org/10.1016/j.jval.2016.12.003

Mertz, M., Prince, I., & Pietschmann, I. (2023). Values, decision-making and empirical bioethics: a conceptual model for empirically identifying and analyzing value judgements. *Theoretical Medicine and Bioethics*, *44*, 567-587. https://doi.org/10.1007/s11017-023-09640-4

Michaels, J. A. (2020). Potential for epistemic injustice in evidence-based healthcare policy and guidance. *Journal of Medical ethics*, *47*, 417-422. https://doi.org/10.1136/medethics-2020-106171

Mitchell, P., Cribb, A., & Entwistle, V. A. (2019). Defining What is Good: Pluralism and Healthcare Quality. *Kennedy Institute of Ethics Journal*, *29*(4), 367-388. https://doi.org/10.1353/ken.2019.0030

Mitchell, P. M., Al-Janabi, H., Richardson, J., Iezzi, A., & Coast, J. (2015). The Relative Impacts of Disease on Health Status and Capability Wellbeing: A Multi-Country Study. *PLoS One*, *10*(12), e0143590. https://doi.org/10.1371/journal.pone.0143590

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2017). Contested evidence: a Dutch reimbursement decision taken to court. *Health Economics, Policy and Law*, *12*(3), 325-344. https://doi.org/10.1017/S1744133116000281

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2020). Questions regarding 'epistemic injustice' in knowledge-intensive policymaking: Two examples from Dutch health insurance policy. *Social science & medicine*, *245*, 112674. https://doi.org/10.1016/j.socscimed.2019.112674

Nshimyumukiza, L., Menon, S., Hina, H., Rousseau, F., & Reinharz, D. (2018). Cell- free DNA non-invasive prenatal screening for aneuploidy versus conventional screening: a systematic review of economic evaluations. *Clinical Genetics*, *94*(1), 3-21. https://doi.org/10.1111/cge.13155

O'Rourke, B., Oortwijn, W., Schuller, T., & International Joint Task, G. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, *36*(3), 187-190. https://doi.org/10.1017/S0266462320000215

Oortwijn, W., Sampietro-Colom, L., & (Eds.). (2022). *The VALIDATE handbook. An approach on the integration of values in doing assessments of health technologies* (W. Oortwijn & L. Sampietro-Colom, Eds.). Radboud University Press. https://doi.org/10.54195/CKHB1659

Refolo, P., Bond, K., Bloemen, B., Autti-Ramo, I., Hofmann, B., Mischke, C., Mueller, D., Nabukenya, S., Oortwijn, W., Sandman, L., Stanak, M., Steele, D., van der Wilt, G. J., & Sacchini, D. (2020). Core competencies for ethics experts in health technology assessment. *International Journal of Technology Assessment in Health Care*, *36*(6), 534-539. https://doi.org/10.1017/S0266462320001968

Reiss, J. (2017). Fact-value entanglement in positive economics. *Journal of Economic Methodology*, *24*(2), 134-149. https://doi.org/10.1080/1350178x.2017.1309749

Richardson, M., Sander, B., Daneman, N., Mighton, C., & Miller, F. A. (2023). Moving from intervention management to disease management: a qualitative study exploring a systems approach to health technology assessment in Canada. *International Journal of Technology Assessment in Health Care*, *39*(1), e67. https://doi.org/10.1017/S0266462323002696

Saarni, S. I., Uusitalo, S., & Autti-Ramo, I. (2022). The role of ethical analysis in conducting a health technology assessment of medical treatments for gender dysphoria. *International Journal of Technology Assessment in Health Care*, *38*(1), e82. https://doi.org/10.1017/S0266462322003257

Schroeder, S. A. (2016). Value Choices in Summary Measures of Population Health. *Public Health Ethics*, *10*(2), 176 - 187. https://doi.org/10.1093/phe/phw032

Smits, M., Ludden, G. D. S., Verbeek, P.-P., & van Goor, H. (2022). Responsible design and assessment of a SARS-CoV virtual reality rehabilitation programme: guidance ethics in context. *Journal of Responsible Innovation*, *9*(3), 344-370. https://doi.org/10.1080/23299460.2022.2076986

Stafinski, T., Glennie, J., Young, A., & Menon, D. (2022). HTA decision-making for drugs for rare diseases: comparison of processes across countries. *Orphanet Journal of Rare Diseases*, *17*, 258. https://doi.org/10.1186/s13023-022-02397-4

Staniszewska, S., & Soderholm Werko, S. (2021). Mind the evidence gap: the use of patient-based evidence to create "complete HTA" in the twenty-first century. *International Journal of Technology Assessment in Health Care*, *37*, e46. https://doi.org/10.1017/S026646232100012X

Steffensen, M. B., Matzen, C. L., & Wadmann, S. (2022). Patient participation in priority setting: Co-existing participant roles. *Social science & medicine*, *294*, 114713. https://doi.org/10.1016/j.socscimed.2022.114713

Stegenga, J. (2015). Effectiveness of medical interventions. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *54*, 34-44. https://doi.org/10.1016/j.shpsc.2015.06.005

Stol, Y. H., Schermer, M. H. N., & Asscher, E. C. A. (2016). Omnipresent Health Checks May Result in Over-responsibilization. *Public Health Ethics*, *10*(1), 35 - 48. https://doi.org/10.1093/phe/phw034

Szabo, S. M., Hawkins, N. S., & Germeni, E. (2024). The extent and quality of qualitative evidence included in health technology assessments: a review of submissions to NICE and CADTH. *International Journal of Technology Assessment in Health Care*, *40*(1), e6. https://doi.org/10.1017/S0266462323002829

van der Wilt, G. J., Bloemen, B., Grin, J., Gutierrez-Ibarluzea, I., Sampietro-Colom, L., Refolo, P., Sacchini, D., Hofmann, B., Sandman, L., & Oortwijn, W. (2022). Integrating Empirical Analysis and Normative Inquiry in Health Technology Assessment: The Values in Doing Assessments of Health Technologies Approach. *International Journal of Technology Assessment in Health Care*, *38*(1), e52. https://doi.org/10.1017/s0266462321001768

van der Wilt, G. J., Gerhardus, A., & Oortwijn, W. (2017). Toward Integration in the Context of Health Technology Assessment: The Need for Evaluative Frameworks. *International Journal of Technology Assessment in Health Care*, *33*(5), 586-590. https://doi.org/10.1017/S0266462317004354

Wehrens, R., & de Graaff, B. (2024). Working with epistemic uncertainties: Emerging entanglements within conditional reimbursement practices. *Health Policy and Technology*, *13*(2). https://doi.org/10.1016/j.hlpt.2024.100850

WHO. (2021). *Health Technology Assessment Survey 2020/21. Main findings.* https://www.who.int/data/stories/health-technology-assessment-a-visual-summary

Wiering, B., de Boer, D., & Delnoij, D. (2017). Patient involvement in the development of patient-reported outcome measures: a scoping review. *Health Expectations*, *20*(1), 11-23. https://doi.org/10.1111/hex.12442

Wilson, J. (2023). What makes a health system good? From cost-effectiveness analysis to ethical improvement in health systems. *Medicine, Health Care and Philosophy*, *26*, 351-365. https://doi.org/10.1007/s11019-023-10149-9


3

Chapter 3

Mixed claims in Health Technology Assessment: The case of Non-Invasive Prenatal Testing

Bloemen, B., Jansen, M., Rijke, W., Oortwijn, W., & van der Wilt, G. J. (2021). Mixed claims in health technology assessment: the case of non-invasive prenatal testing. *Social science & medicine*, *270*, 113689. https://doi.org/10.1016/j.socscimed.2021.113689

ABSTRACT

Health Technology Assessment (HTA) uses explicit methods to determine the value of a health technology. This typically results in several claims regarding the effects that are expected to follow from the use of a health technology in a particular context. These claims seem to capture conclusions based solely on facts, but they often combine empirical information with normative presuppositions. Claims that have this character reflect (implicit) value judgments and have been labelled *mixed claims.* Not recognizing these normative components of such claims risks *value inattention* and *value imposition*, presenting results as self-evident and not in need of any moral justification. As proposed by Anna Alexandrova, to avoid these risks of value inattention and imposition we need rules to deal with mixed claims. According to her, when producing and evaluating mixed claims we need to unearth the invoked value presuppositions and check whether these presuppositions are invariant to disagreements. By applying these rules, the *robustness* of mixed claims can be checked: it can be evaluated whether their truth value is independent from the way in which their components, involving normative presuppositions, are conceptualized. This paper aims to illustrate the role of mixed claims in HTA, and expand upon the work by Alexandrova, by analyzing claims and recommendations presented in an HTA report on the introduction of Non-Invasive Prenatal Testing (NIPT) in The Netherlands. Our results show that the report contains mixed claims, and that a normative analysis of these claims can help to clarify the normativity of HTA and evaluate the robustness of claims on alleged effects of a health technology.

INTRODUCTION

According to a widely held view, scientists, when informing policy, should refrain from making value judgments (Douglas, 2009). Although, on this account, values admittedly play a role in policy making, scientific research itself is regarded as value-neutral. In the present paper, this view will be challenged. In particular, it will aim to show that in a specific type of policy analysis, Health Technology Assessment (HTA), value judgments cannot be avoided. HTA is a specific type of policy analysis, aimed at clarifying, through empirical inquiry, the value of health technologies (O'Rourke et al., 2020). Typically, it is conducted in the context of public policy making, providing guidance as to how public resources are to be used in funding health services. Because of its functioning in value declaration and in public policy making, the role of values is of specific importance to HTA. In the current practice of HTA, two features stand out: [1] the distinction between *assessment* on the one hand, and *appraisal* on the other, and [2] the distinction between ethical, legal and social issues ('ELSI') on the one hand, and safety, clinical effectiveness and cost-effectiveness on the other hand. 'Assessment' is held to be the value-neutral collection of facts about a health technology; 'appraisal' refers to the value-laden process of reaching decisions (e.g., reimbursement) on the basis of those facts. The separate inquiry into 'ELSI' suggests that the other aspects (safety etc.) can be insulated from these value issues (Ducey et al., 2017; Hofmann et al., 2018; Legault et al., 2018). Both features may be considered as attempts to prevent values from unduly biasing the evidential basis for policy making.

The question is, however, whether the passionately sought separation of facts from values can be obtained in the first place. And, if not, whether we are not making things worse, and would be better advised to squarely face the unavoidable entanglement between facts and values without sacrificing scientific rigor. The present paper explores whether the concept of mixed claims, as developed by Alexandrova (Alexandrova, 2016), can be used for this purpose.

This paper aims to investigate the role of mixed claims in HTA and apply the methodology presented by Alexandrova (Alexandrova, 2016) to deal with these claims. We use the introduction of Non-Invasive Prenatal Testing (NIPT) in the Netherlands as a case study. Our findings will be discussed in terms of whether the methodology of mixed claims provides a means to get a firmer grip on the role of values in HTA, without sacrificing scientific rigor and introducing undue bias.

In the following, we will present a more detailed account of the concept of mixed claims and explain how it may play a role in the context of HTA.

A potential role for mixed claims in HTA

According to Lucivero (Lucivero, 2016), the general structure of claims that are being produced and evaluated in the context of HTA is as follows:

$$\textbf{Technology} \xrightarrow{[\textit{Conditions}]} \textbf{Effect}$$

They relate the introduction or use of a technology, given certain conditions, to alleged effects. For example, a claim that a new antihypertensive drug ('Technology'), prescribed for people with high blood pressure ('Conditions'), will lead to a reduced risk of cardiovascular disease ('Effect'). Such claims can be made by either developers of a health technology, professionals who wish to use a health technology, or other stakeholders having certain expectations. A central task for HTA is to assess the plausibility of such claims. How these components (Technology, Conditions, Effect) are defined will determine what sort of evidence needs to be collected (as part of the assessment to substantiate, or challenge, such claims).

The reason why the work by Alexandrova on mixed claims (Alexandrova, 2016) is interesting to HTA, and to health policy analysis generally, is that she has demonstrated that components of a causal claim can be defined in a way that involves value judgments, potentially affecting the causal relation itself. In fact, she has shown that examining causal claims may involve value judgments in two ways: (i) a researcher may adopt a given effect measure because it is held to more adequately reflect a certain quality (e.g., health) than other measures; (ii) and a researcher may adopt a particular methodology for measuring an effect that implies a normative commitment to the validity of certain conceptualizations of that effect. For exampling, when analyzing the effectiveness of an antihypertensive drug its effect can be defined and measured in terms of lowering blood pressure, reducing the risk at cardiovascular disease and co-morbidities, or in terms of the impact on quality of life. A choice between these measures already suggests a judgment on the relative importance of certain outcomes. But some of these measures, like quality of life, demand additional decisions, in terms of how to apply them, which imply a normative commitment to what constitutes desirable outcomes. Therefore, when examining claims on the effectiveness of antihypertensive drugs, empirical inquiry and normative presuppositions become entangled, resulting in mixed claims.

Hence, values may become entangled with empirical information in a more complex way when claims concerning the effects of a health technology are being produced or critically examined. This means that claims that appear to capture conclusions based solely on facts actually partly depend on value judgments. Not recognizing this may lead policy makers and the general public to assume, wrongly, neutrality of the claims

at stake. Instead of trying to prevent values from influencing the evidential basis for policy making, we should acknowledge this entanglement between facts and values. By this, we can avoid the risk of *value imposition*, importing substantive moral views into science, by controlling the risk of *value inattention*, failing to notice the value judgments involved in science and presenting results as self-evident (Alexandrova, 2016).

Aim of this paper

This paper aims to explore in more detail what such entanglement between value judgments and empirical inquiry looks like, by applying the concept of mixed claims and the methodology proposed by Alexandrova (Alexandrova, 2016) to an HTA conducted on NIPT.

Alexandrova proposed the following steps to deal with mixed claims: (i) unearth value presuppositions in methods and measures; (ii) check whether these presuppositions are invariant to disagreements; (iii) in case of disagreements, consult relevant parties (Alexandrova, 2016). A claim can be regarded robust when it stands up to a range of different ways of conceptualizing its components. Thus, a claim that states that a specific health technology is safe, under particular conditions, is robust when its truth value is independent from the way in which safety is conceptualized. If it is not robust, the particular choice of how to conceptualize such an element (e.g., safety) should be discussed in a deliberative setting involving stakeholders.

We explored the use of the first two steps proposed by Alexandrova, using an HTA report on the introduction of NIPT in the Netherlands. NIPT is a prenatal screening procedure that analyzes cell-free fetal DNA, circulating in the mother's blood, to obtain information about the fetal genotype (Hui & Bianchi, 2017). Its introduction raised a broad range of questions beyond clinical effectiveness, including societal and ethical implications (Kibel & Vanstone, 2017). Assessing its potential added value raises questions on how to define its effects because these are not reducible to health-related outcomes and requires an evaluation of normative concepts such as reproductive autonomy (Kessels et al., 2019; Stapleton et al., 2019). Consequently, it is to be expected that an assessment of this technology invokes value judgments, making it a suitable case to investigate how mixed claims may play a role in HTA.

METHODS

We have used a case study approach to obtain an in-depth understanding of mixed claims in HTA. First, we selected an HTA report on NIPT in the Netherlands as an instrumental case to gain a broader understanding of the issue of mixed claims in

HTA (see subsection '*Case study: introduction of NIPT in the Netherlands*'). To identify mixed claims, we collected causal claims from this report and analyzed them in terms of how alleged effects of NIPT were conceptualized (see subsection '*Identifying mixed claims in an HTA report on NIPT*'). The robustness of these identified mixed claims was then evaluated by discussing different ways of conceptualizing and operationalizing these effects (see subsection '*Evaluating the robustness of mixed claims*').

Case study: introduction of NIPT in the Netherlands

Until recently, prenatal screening for trisomy 13 (Patau syndrome), 18 (Edwards syndrome), and 21 (Down syndrome), consisted of a combined test (i.e. a first trimester screening test based on blood serum markers and an ultrasound scan) and, after a positive test result, a choice between two invasive tests, amniocentesis or chorionic villus sampling, to confirm diagnosis (Hui & Bianchi, 2017). Since 2011, this practice of prenatal screening has changed due to the introduction of NIPT. Because only a blood sample of the mother is needed, it is a non-invasive procedure that does not pose any risks of procedure-related miscarriage. In addition, it can be performed early during pregnancy (around the 10th week of pregnancy), it might have an even higher reliability than existing tests, and could potentially be used to analyze the fetal genome which provides the option of detecting conditions for which as yet no screening protocol exists (Hui & Bianchi, 2017). In The Netherlands, the current use of NIPT targets screening for trisomy 13, 18, 21 (van der Meij et al., 2019). After a positive test result, parents are offered the choice to take an invasive test to confirm diagnosis.

A task of The Health Council of the Netherlands is to advice the Ministry of Health on population screening programs. In 2013, the Health Council was asked to produce a report exploring potential future uses of NIPT. The goal was to provide recommendations with respect to future uses of NIPT and to explore whether the current evaluative framework suffices to provide guidance on novel screening technologies such as NIPT. The report analyzes safety, reliability, and social and ethical implications of NIPT, in an attempt to determine whether NIPT could contribute to improved prenatal screening by providing respective parents with 'meaningful reproductive choices' (Health Council of the Netherlands, 2013).The report concluded that NIPT could indeed lead to improved reproductive choices, but also raises questions concerning the appropriateness of the current evaluative framework.

This report was used to analyze the role of mixed claims in assessing NIPT. It qualifies as an HTA report according to the typology of the International Network of Agencies for HTA (INAHTA), and is included in the international HTA database coordinated by INAHTA (https://database.inahta.org/article/15018).

Identifying mixed claims in an HTA report on NIPT

To assess whether the Health Council report on NIPT contains mixed claims, the following steps were taken:

(i) The first author (BB) performed a mapping exercise, identifying causal and correlational claims in the report that relate the introduction and use of NIPT, under certain conditions, to a certain effect (e.g., 'The use of NIPT will lead to less false-negative test results for trisomy 21'):

$$\text{NIPT} \xrightarrow{[\textit{Conditions of use}]} \text{Effect}$$

As a first step, such causal and correlational claims were identified on the basis of the report's recommendations. Next, the full report was searched for definitions of their components (*NIPT*, *conditions of use*, *effect*), and for the arguments and evidence presented in evaluating the claim. Face validity of the results of this analysis was independently checked by a second author (GJvdW).

(ii) To assess whether these identified claims represent '*mixed* claims', two authors (BB, GJvdW) independently analyzed every claim by answering the following question: *Does the conceptualization of the alleged effect of NIPT presuppose a value judgment about its nature?*

According to the definition given by Alexandrova (Alexandrova, 2016), *all components* of a causal claim (e.g., *NIPT, Conditions, Effect*), may presuppose value judgments and, therefore, make the claim *mixed*. Given that NIPT is a prenatal strategy that can be implemented and used in many ways (Vanstone, 2015), examining claims on alleged effects requires assumptions on the presumed implementation and use of NIPT. These assumptions may presuppose value judgments on the desirability of particular uses of NIPT. Although this may be interesting to analyze with respect to identifying mixed claims, we have chosen to restrict our analysis to the conceptualization of alleged effects. The Health Council of The Netherlands did not explicitly assess different ways of using NIPT, besides a broader (i.e., genome-wide testing) use of NIPT, because its use is constrained by the Dutch Population Screening Act. Therefore, many decisions concerning its presumed use were already made when the report was commissioned.

Although 'values' may be conceptualized differently, also in the context of health policy (Giacomini et al., 2001; Shams et al., 2016), a common denominator is that it concerns ideas about what is right and wrong, and what situations in life we should aim to realize or avoid. A value judgment, then, is a judgment on whether a particular situation (or

act, or event) satisfies these ideas (i.e. a judgment about what is good) (Hofmann et al., 2014). A value judgment does not have to explicitly declare that we 'ought' to support a certain situation. For example, the statement 'clean needle programs reduce the incidence of AIDS' implies an evaluation and a prescription without declaring that we ought to support clean needle programs (Giacomini et al., 2001).

(iii) The results were discussed by these two authors (BB, GJvdW) to arrive at a consensus on the interpretation of the claims and whether they truly represent *mixed* claims.

Evaluating the robustness of mixed claims

To unearth the value judgments invoked by the identified mixed claims, all authors independently analyzed these claims. For every mixed claim, an author was asked to identify and explicate the values invoked by its conceptualization of an alleged effect of NIPT. This required them to provide different ways in which a particular effect could be conceptualized, the consequences of such a conceptualization in terms of how the effect may be observed and measured (operationalized), and whether these conceptualizations invoke different values and / or different conceptualizations of these values. For example, a claim that relates the use of NIPT to enhanced safety invokes the value of avoiding harm to people (i.e., the ethical norm of non-maleficence), which may be defined in different ways. Every author listed their identification of invoked values and alternative conceptualizations of these values.

To analyze the robustness of the identified mixed claims, the invoked values and their conceptualizations, as listed by the authors, were discussed in a joint meeting. During this meeting, we identified the values, and their conceptualizations, that could be recognized by all authors. Based on that, we evaluated whether alternative conceptualizations of these values would lead to different ways of conceptualizing and operationalizing alleged effects of NIPT, and whether that may lead to different conclusions concerning the plausibility of NIPT realizing that effect.

RESULTS

Identified mixed claims in an HTA report on NIPT

Identified causal and correlational claims

The report on NIPT presents four recommendations for introducing NIPT in the Netherlands (see Table 1). Based on these recommendations, six causal claims are distilled from the report (Table 1). For these six claims it was examined whether they can be regarded as a mixed claim, by determining whether the conceptualization of the alleged effect of NIPT invokes value judgments.

Table 1. Recommendations and claims in the Health Council of the Netherlands report on NIPT (Health Council of the Netherlands, 2013), components of the claims, and our assignment of their status as mixed claim.

Recommendation	Claim	Components	Mixed claim
(1) "NIPT as an early, completely safe and direct diagnostic test is not currently a feasible option. A realistic scenario is NIPT as the first follow-up test after a combined test indicating an elevated risk of trisomy. In time, NIPT may prove an alternative to the combined test. Although screening with NIPT can currently be considered, at most, safer and more reliable than the current approach, the Committee believes this suffices for consideration of NIPT in the screening programme. It particularly means that disadvantages of the screening for participants decrease, increasing the appeal of the advantages of participation (the intended reproductive choice) for a larger group. Removing impediments can lead to improved achievement of the screening goals. The Committee does feel that implementation studies should explicitly examine the potential disadvantages, such as the fact that NIPT sometimes fails."	**(1) NIPT** $\xrightarrow{[second\ tier\ test]}$ **safety**	*Second tier test* = NIPT as the first follow-up test after a combined test indicating an elevated risk of trisomy (13, 18, 21). *Safety* = safer outcomes = avoidance of invasive prenatal test procedures and procedure-related miscarriages.	Yes
	(2) NIPT $\xrightarrow{[second\ tier\ test]}$ **more reliable test results**	*Second tier test* = NIPT as the first follow-up test after a combined test indicating an elevated risk of trisomy (13, 18, 21). *More reliable test results* = higher predictive value for the fetus having trisomy 21 than the existing prenatal screening procedure.	Yes
	(3) NIPT $\xrightarrow{[second\ tier\ test]}$ **higher uptake**	*Second tier test* = NIPT as the first follow-up test after a combined test indicating an elevated risk of trisomy (13, 18, 21). *Higher uptake* = increased appeal of participation = less barriers to participation.	No
(2) "To date, the worries discussed in the literature on NIPT are primarily concerned with the possibility of routinization as a downside to an improved screening procedure. This may come at the cost of careful decision making and makes screening options vulnerable to the criticism that what is labelled reproductive autonomy is in fact nothing more than preventing the birth of children with conditions and handicaps that are costly for society. The Committee does not consider this a valid reason to reject improvement of current screening processes with NIPT. However, it does mean the development of daily practice must be carefully examined within the context of screening objectives. And that providing careful information and counseling will be no less important as screening tests become simpler, more reliable and safer."	**(4) NIPT** $\xrightarrow{[second/first\text{–}tier\ test]}$ **routinization**	*Second/first-tier test* = NIPT as second tier test (the first follow-up test after a combined test indicating an elevated risk of trisomy 13, 18, 21), or as a first-tier screening test. *Routinization* = the risk that the choice of taking the NIPT test will be presented as self-evident, which could potentially lead to situations in which parents do not sufficiently realize the disadvantages and consequences of taking this test.	Yes
(3) "An unavoidable question for the future is what the scope of screening tests should ultimately be, who should make the decisions, and based on what criteria. The Committee notes that assuming the goal of 'providing meaningful reproductive choices' is best served by screening that is as broad as possible, is too simplistic."	**(5) NIPT** $\xrightarrow{[broad\ scope]}$ **meaningful reproductive choices**	*Broad scope* = a screening test that is as broad as possible, given technological and economical possibilities. *Meaningful reproductive choices* = choices that concerns (severe) health problems (of the fetus), is informed (respects the autonomy of the parents), and proportional (respects the anticipatory autonomy rights of the child).	Yes
(4) "The Committee expects that the developments outlined in this monitoring report will have a major impact on the daily practice of prenatal screening in the coming decade. There is a strong need for steering focused on a responsible and timely transition to prenatal personalised medicine. How future proof the current normative framework is, as well as the role of the WBO in this area, must also be examined critically."	**(6) NIPT** $\xrightarrow{[broad\ scope]}$ **prenatal personalized medicine**	*Broad scope* = a screening test that includes mapping gene expression patterns and genetic variations that predict pregnancy complications and fetal developmental disorders, as well as findings that can lead to medical treatment for the foetus. *Prenatal personalized medicine* = offering prevention and treatment interventions, based on genetic information, to treat conditions of the fetus.	Yes

Conceptualizations of the effects of NIPT

The potential effects that the Health Council evaluated included the impact of NIPT on the safety, reliability and uptake of prenatal screening; routinization, meaningful reproductive choices; and prenatal personalized medicine (Table 1). The way in which these effects were conceptualized (Table 1), as well as their operationalization and arguments and evidence presented to assess them (see Appendix 1), were analyzed to determine whether they invoked value judgments and can, therefore, be classified as being part of mixed claims.

Safety was evaluated with respect to avoidance of procedure-related miscarriages (Table 1, Appendix 1; claim #1). Because NIPT is a non-invasive test its main contribution to safety would be to avoid using invasive tests and their associated risk of miscarriage. In addition, the Health Council also considered avoidance of unnecessary worries concerning the health of the child, due to false positive test results, as part of the impact of NIPT on safety (Appendix 1). Other possible influences on psychological well-being, such as decisional regret, societal pressure to take the test, and distress related to difficult decisions that need to be made as a consequence of test results, were not taken into account in the evaluation of safety. These consequences relate to alternative ways of defining 'safety' that involve value judgments on its scope and nature.

Reliability of NIPT was evaluated in terms of its predictive value for the fetus having trisomy 21 (Table 1, Appendix 1; claim #2). Although the Dutch Population Screening Act permits screening for severe conditions with no existing treatment or prevention options, and screening for Down's syndrome is already current practice in the Netherlands, it still requires value judgments to determine whether Down's syndrome is severe enough to offer information and facilitate parents in making decisions about continuing a pregnancy with this condition. Moreover, it is not a neutral exercise to define what constitutes an accurate and reliable test. Acceptable thresholds of different components of test accuracy and reliability (e.g. sensitivity, specificity) invoke value judgments concerning the (un)desirability of certain outcomes, the acceptance of uncertainty, and the severity of conditions being tested for.

Higher uptake was evaluated with respect to whether introducing NIPT would lead to more parents willing to participate in prenatal screening (Table 1, Appendix 1; claim #3). Although the desirability of a higher uptake of prenatal screening tests is a normative issue, defining higher uptake itself seems straightforward, not involving any particular value judgments.

Another examined claim was the alleged relation between introducing NIPT and routinization (Table 1, Appendix 1; claim #4). According to the Health Council, routinization refers to a specific threat to careful decision making: the risk that the choice of taking NIPT will be presented as self-evident, potentially leading to situations in which parents insufficiently realize the consequences of taking the test. Therefore, the assessment of claims on the relation between NIPT and routinization implies value judgments on what constitutes careful decision making to identify potential barriers introduced and denoted by 'routinization', making this another example of an effect that invokes value judgments in its conceptualization.

A central claim in the report on NIPT is the relation between offering this test and providing meaningful reproductive choices (Table 1, Appendix 1; claim #5). A meaningful reproductive choice is defined by the Health Council as a choice that concerns severe health problems (of the fetus), is made in an informed way (respecting autonomy of the parents), and proportional – meaning that it respects the anticipatory autonomy rights of the child. The report concludes that offering NIPT with a broad scope, providing genome wide information, does not necessarily serve the goal of supporting meaningful reproductive choices. This conclusion is based on concerns related to the requirements of informed choice, and that offering information on late-onset disease may not necessarily be in the interest of the fetus. This implies value judgments on which sorts of conditions are sufficiently severe enough to neglect the rights of the fetus to make its own decisions. In addition, determining the scope of meaningful reproductive choices also invokes value judgments.

Finally, we examined the claim that NIPT could be used to enable prenatal personalized medicine (Table 1, Appendix 1; claim #6). According to the Health Council, this means that NIPT does not only offer information to make decisions on the continuation of pregnancy but also on conditions that could be treated during pregnancy and birth (e.g., fetal developmental disorders). These decisions on prenatal prevention and therapy would align with a second objective of prenatal screening: ensuring a healthy outcome of pregnancy for mother and child. This could be conflicting with the first objective of prenatal screening (facilitating reproductive choice) when taking a test would be necessary to enable treatment of the fetus. The Health Council concludes that this raises questions on the adequacy of the current normative framework, which is based on an ethics of non-directive reproductive counseling. Although it is plausible that the development of prenatal personalized medicine would lead to potential conflicts between facilitating reproductive choice and ensuring a healthy outcome for the child, it could be argued that prenatal screening already leads to such conflicts. For example, does information on trisomy 21 not already allow the adoption of preventive measures (e.g., adapting the environment in which the child will be born) that could enhance the future health of the child? Therefore, the conceptualization of prenatal

personalized medicine as a development that would lead to future conflicts is based on value judgments related to the scope of parental autonomy in relation to the rights of the future child. This makes it another example of a mixed claim.

Evaluation of the robustness of identified mixed claims

The results of the identification of the values invoked by the identified mixed claims, and alternative conceptualizations of these values, are summarized in Table 2. The results of the discussion on whether these alternative conceptualizations influence the robustness of these claims, leading to different conclusions concerning the plausibility of NIPT realizing these effects, are described below.

Table 2. Results of the identification and evaluation of the values invoked by the identified mixed claims in the NIPT report (Health Council of the Netherlands, 2013).

Mixed claim	Invoked values	Alternative conceptualizations
(1) NIPT $\xrightarrow{[second\ tier\ test]}$ **safety**	Non-maleficence	Non-maleficence: *avoid harm done by taking the test; avoid harm done by decisions made during pregnancy; avoid premature death of a viable fetus; avoid harm done to users and non-users of the test.*
(2) NIPT $\xrightarrow{[second\ tier\ test]}$ **more reliable test results**	Autonomy Non-maleficence	Autonomy: *respect the autonomy of prospective parents by offering them reliable information about the health of their fetus*; *respect the autonomy of prospective parents by offering them all relevant information concerning the reliability of a prenatal test; respect the autonomy of prospective parents by facilitating informed decisions during pregnancy that do not clearly neglect or abuse the rights of the fetus.* Non-maleficence: *avoid unnecessary worries about the health of the fetus; avoid premature death of a viable fetus.*
(3) NIPT $\xrightarrow{[second/first-tier\ test]}$ **routinization**	Autonomy Non-maleficence	Autonomy: *respect the autonomy of prospective parents by offering prenatal tests in such a way that they are informed about the consequences of taking a test and able to make their own decisions; respect the autonomy of prospective parents by offering prenatal tests in such a way that they are able to make a decision that is in their own interest.* Non-maleficence: *avoid the risk that the choice of taking the NIPT test will be presented as self-evident, which could potentially lead to parents insufficiently realizing the consequences of the test and a violation of disability rights.*
(4) NIPT $\xrightarrow{[broad\ scope]}$ **meaningful reproductive choices**	Autonomy	Autonomy: *respect the autonomy of prospective parents by offering prenatal tests that provide parents with the most reliable information about the future health of their fetus and help them making decisions during pregnancy and childhood that promote a desirable future for the fetus.*
(5) NIPT $\xrightarrow{[broad\ scope]}$ **prenatal personalised medicine**	Autonomy Beneficence	Autonomy: *respect the autonomy of prospective parents by allowing them to make decisions during pregnancy that do not jeopardize the viability of the fetus; respect the autonomy of prospective parents by allowing them to make decisions during pregnancy that do not jeopardize the wellbeing of the fetus.* Beneficence: *you should promote the wellbeing of persons who have been entrusted to your care by acting in a way that ensures the highest possibility of a healthy outcome of pregnancy for mother and child; you should promote the wellbeing of persons who have been entrusted to your care by acting in a way that ensures a minimal quality of life of the fetus.*

The results show that safety (Table 2, claim #1) was conceptualized by the Health Council primarily in terms of avoidance of procedure-related miscarriages and unnecessary worries due to false positive test results. It could, however, be argued that other possible consequences of taking a prenatal test should be taken into account when considering its safety. For instance, potentially harmful effects on the fetus, on people declining the use of such a test despite societal pressure, and changing societal views related to people still being born with conditions that are screened, could be taken into account. Such considerations regarding the safety of a prenatal test may be held to relate to the value of non-maleficence, to avoid harm being done. Depending on how this is conceptualized, whose safety should be considered and which negative impacts are important, other aspects of using NIPT become relevant to take into account when assessing its safety. Depending on the outcome of such an assessment, this may also influence conclusions regarding the safety of NIPT (that is, the causal relation between the technology and the outcome).

The evaluation of the relation between introducing NIPT and the reliability of prenatal testing, routinization, facilitating meaningful reproductive choices, and the development of prenatal personalized medicine (Table 2, claims #2,3,4,5) all appear to invoke the value of autonomy. The relevance of these effects of NIPT is determined by the goal of enhancing parental autonomy, which is realized, according to the Health Council, by enabling parents to make well-informed decisions concerning severe health problems of the fetus and respecting the child's autonomy. The concept of autonomy can also be understood in different ways. Does it refer only to decision-making capacity or also to be able to realize decisions? Is it realized by increasing knowledge or by encouraging self-reflection on how NIPT could help in realizing life goals of prospective parents (Kater-Kuipers et al., 2020)? Consequently, depending on the underlying concept of autonomy, providing information to prospective parents (by offering them NIPT) may not be sufficient nor necessary. In addition, values such as non-maleficence and beneficence are important with respect to determining which choices and information should be offered to prospective parents. Depending on the conceptualization of these values, and their relations, other aspects should be taken into account when examining alleged effects of NIPT. This could influence conclusions regarding the plausibility that NIPT leads to routinization and / or meaningful reproductive choices. They also could influence conclusions regarding the relation between NIPT and prenatal personalized medicine, given that it relates to defining the scope of parental autonomy with respect to the autonomy of the child. Claims on the reliability of NIPT could be influenced because the underlying value conceptualizations determine ideas on which conditions should be screened by using NIPT.

DISCUSSION

Mixed claims and the normativity of HTA

Our results show that some of the claims examined in the HTA report on NIPT, produced by the Health Council of the Netherlands, can be perceived as mixed claims. In addition, these claims may not stand up against different ways of conceptualizing and operationalizing the effects of NIPT, possibly involving value judgments that are not agreed upon by different stakeholders.

These results imply that values play a much larger role in identifying and collecting empirical information, that needs to be taken into account in an HTA, than is often acknowledged. When examining the plausibility of claims concerning alleged effects of a health technology, decisions have to be made about how to conceptualize these effects and which methods to be used in measuring these effects. These decisions may involve value judgments but are necessary to conduct an assessment. Evidence on effects of a health technology needs to be actively collected and taken into account. Acknowledging and explicating the role of values in this process helps in identifying relevant evidence and empirical research that may be needed to draw conclusions on the robustness of these claims.

Therefore, a normative analysis aimed at explicating values that should be realized by the use of a health technology is an integral part of an assessment because it influences the informational requirements that an assessment needs to meet to inform decision making. This implies that the distinctions between assessment and appraisal, and between 'value-laden' ELSI issues and 'value-neutral' analyses of safety, effectiveness, and cost-effectiveness of a health technology, cannot be obtained. Instead of trying to separate facts from values, we should develop methodology to get a firmer grip on the role of values in HTA and maintain scientific rigor. Incorporating the work performed by Alexandrova (Alexandrova, 2016) in the practice of HTA would be an important first step in realizing this.

In the remainder of the discussion, we will address issues related to the generalizability of our findings and the consequences for the practice of HTA.

Generalizability of results

Mixed claims in different types of analyses of health technology

Our results show that assessing safety and clinical effectiveness may involve mixed claims, and this points towards a role for value judgments in all types of analyses (safety, effectiveness, costs-effectiveness, ELSI) associated with HTA. Although the report on NIPT that we have analyzed does not include a cost-effectiveness analysis,

there are reasons for suggesting that mixed claims are also involved in examining cost-effectiveness. This can be seen in the Belgian Health Care Knowledge Centre report on health economic aspects of NIPT (Hulstaert et al., 2014). The authors of the report explicitly state that they deviate from Belgian guidelines for economic evaluations by not expressing outcomes in terms of euro per quality-adjusted life years (QALYs) gained. They prefer to use the average cost per trisomy 21 detected for different screening scenarios. According to the authors, this choice is not based on methodological considerations but based on their view that the use of NIPT should focus on providing correct information to parents. In addition, they also discuss challenges with translating outcomes into QALYs concerning whose QALYs should be taken into account and defining the appropriate time horizon (e.g., should only impacts up to detection of affected pregnancies be taken into account, or also long-term implications). This shows that analyzing NIPT in terms of cost-effectiveness also involves value judgments. Although NIPT may raise specific challenges, a role for value judgments in health economics in general is already described (Harvard et al., 2020). Further research on the role of mixed claims in this area is needed to demonstrate how mixed claims are constitutive of claims concerning cost-effectiveness, and its implications for the role of values in health economic modeling.

Mixed claims and context

Our analysis was based on a published policy report which may not capture all concerns, questions and decisions encountered and addressed by a committee responsible for drafting the report. Despite this, the report embodies the public manifestation and justification of decisions made and its results. Therefore, an analysis based on such a report does provide valuable insight on how mixed claims play a role, and are justified, in HTA. To also uncover how practical aspects of an assessment (e.g., context, material- and non-verbal elements) structure and influence this normativity of HTA, it may be fruitful to conduct interviews or take an ethnographic approach in future research on mixed claims.

Because we have only analyzed an assessment of NIPT conducted in the Netherlands, contextual aspects of the Dutch healthcare system may have influenced our conclusions. Despite this, we believe that the nature of the normativity involved in the assessment process is similar in other contexts. For example, an analysis of the governance of NIPT in Germany shows similar challenges related to normative questions that are raised by informing and making decisions on the use of NIPT (Braun & Könninger, 2018). One especially contested issue was the question whether NIPT can be seen as a *medical procedure*, because German law states that the statutory health insurance should guarantee access to *medical necessary* services. The Federal Joint Committee (G-BA) is the supreme decision-making body, and decides on therapeutic usefulness,

cost-effectiveness, and medical necessity of a health technology. The G-BA decided to define the medical purpose of NIPT in terms of its ability to save healthy fetuses from procedure-related miscarriages. By doing this, it implicitly defined the purpose of NIPT (i.e., as a medical procedure) in a way that was highly contested by different stakeholders and organizations in German society. Although the G-BA tried to take responsibility for these ethical and social implications by inviting other institutions to address them, the normativity involved in conducting an assessment of medical and scientific aspects of NIPT could not be eliminated. The decision itself to conduct an assessment of NIPT, recognizing it as a medical procedure, and decisions on what effects to take into account and how to define them, are already normative decisions.

Mixed claims and other health technologies

Because our findings are based on a single case study. it may be influenced by the specific nature of NIPT. NIPT is an example of a morally challenging health technology that raises challenges because its purpose can be conceptualized in many ways, it is aimed at non-curative goals, and demands the use of outcome measures that are able to capture valuable effects beyond health-related goals (Kessels et al., 2019). Despite these particular characteristics of NIPT, the claims assessed in HTA have a generic nature that makes it likely that mixed claims are involved in a wide variety of HTAs. The assessment of the impact of a health technology in terms of safety, clinical effectiveness, cost-effectiveness, or social and ethical implications, always has a normative dimension. The relevance and scope of these impacts is related to normative commitments. For example, 'safety' refers to the sort of outcomes that we wish to avoid *because of our commitment to avoiding harm*. And 'clinical effectiveness' refers to the sort of outcomes that we wish to achieve *because of our commitment to doing good*. These normative commitments guide the collection of information needed to assess whether a health technology is able to realize these effects. And concluding that a health technology realizes a certain effect implies ascribing a certain quality to this technology (Legault et al., 2018).

Moreover, for a health technology to be *effective* it must be able to reduce disease-related disabilities or enhance health, which refer to states of being that are disvalued or desirable (Stegenga, 2018). Therefore, a health technology not only needs to target just any constitutive, physiological, basis of health and disease, but also realize a change that is regarded an improvement. Consequently, an assessment of the effectiveness of a health technology always depends on specific accounts of normative concepts, such as health and wellbeing (Hofmann et al., 2018), and an evaluation of different states of being. Given the centrality of claims of effectiveness in HTA, it is highly likely that mixed claims are present in assessments of a wide variety of health technologies.

Other approaches to normative analysis

In our identification of values, we have tried to stay as close as possible to the reasoning of the Health Council of the Netherlands. Because their approach is guided by the Dutch Population Screening Act, which provides criteria that an acceptable screening program needs to satisfy, it may be seen as an instance of *principlism*. It could be the case that taking another approach, such as virtue ethics or a phenomenological approach, leads to the identification of other value concerns related to the use of NIPT (Svenaeus, 2018). It would be interesting to see how an analysis of mixed claims would work out when these other approaches to ethics are taken into account.

CONCLUSION

Our analysis of an HTA report on NIPT shows that some of the claims examined are *mixed claims*, because the way in which alleged effects of NIPT are conceptualized invokes value judgments on desirable consequences of using NIPT. This illustrates that facts and values become entangled in assessing potential effects of a health technology. Recognizing this, by identifying and scrutinizing mixed claims in HTA, is important to avoid value imposition and inattention and get a firmer grip on the role of values in HTA. Developing methods for evaluating mixed claims could enhance transparency and robustness of the results of HTA.

CREDIT AUTHOR STATEMENT

Bart Bloemen: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Project administration; Maarten Jansen: Formal analysis, Writing – review & editing; Wouter Rijke: Formal analysis, Writing – review & editing; Wija Oortwijn: Methodology, Formal analysis, Writing – review & editing; Gert Jan van der Wilt: Conceptualization, Methodology, Formal analysis, Writing – review & editing.

DECLARATION OF COMPETING INTEREST

Gert Jan van der Wilt was member of the Committee on Population Screening of the Health Council of the Netherlands at the time the NIPT report was drafted. He participated in the research that is reported in the present paper as an independent scholar affiliated to Radboud University Medical Centre, not bound by this erstwhile committee membership.

ACKNOWLEDGEMENTS

We would like to thank Anna Alexandrova for her thoughtful comments on earlier versions of this paper.

APPENDIX

The Appendix for this chapter can be found online as part of the published article in *Social Science & Medicine*: https://doi.org/10.1016/j.socscimed.2021.113689.

REFERENCES

Alexandrova, A. (2016). Can the Science of Well-Being Be Objective? *The British Journal for the Philosophy of Science*, *69*(2), 421-445. https://doi.org/10.1093/bjps/axw027

Braun, K., & Könninger, S. (2018). Realizing responsibility. Institutional routines, critical intervention, and the "big" questions in the controversy over non-invasive prenatal testing in Germany. *New Genetics and Society*, *37*(3), 248-267. https://doi.org/10.1080/14636778.2018.1495555

Douglas, H. E. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press. http://www.jstor.org/stable/j.ctt6wrc78

Ducey, A., Ross, S., Pott, T., & Thompson, C. (2017). The moral economy of health technology assessment: an empirical qualitative study. *Evidence & Policy*, *13*(1), 7-27. https://doi.org/10.1332/174426415X14443053123024

Giacomini, M., Hurley, J., Gold, I., Smith, P., & Abelson, J. (2001). 'Values' In Canadian Health Policy Analysis: What Are We Talking About? *Canadian Health Services Research Foundation*, 13.

Harvard, S., Werker, G. R., & Silva, D. S. (2020). Social, ethical, and other value judgments in health economics modelling. *Social science & medicine*, *253*, 112975. https://doi.org/10.1016/j.socscimed.2020.112975

Health Council of the Netherlands. (2013). *NIPT: dynamics and ethics of prenatal screening* (2013/34). H. C. o. t. Netherlands.

Hofmann, B., Bond, K., & Sandman, L. (2018). Evaluating facts and facting evaluations: On the fact-value relationship in HTA. *Journal of evaluation in clinical practice*, *24*, 957-965. https://doi.org/10.1111/jep.12920

Hofmann, B., Cleemput, I., Bond, K., Krones, T., Droste, S., Sacchini, D., & Oortwijn, W. (2014). Revealing and acknowledging value judgments in health technology assessment. *International Journal of Technology Assessment in Health Care*, *30*(6), 579-586. https://doi.org/10.1017/S0266462314000671

Hui, L., & Bianchi, D. W. (2017). Noninvasive prenatal DNA testing: the vanguard of genomic medicine. *Annual review of medicine*, *68*, 459-472. https://doi.org/10.1146/annurev-med-072115-033220

Hulstaert, F., Neyt, M., & Gyselares, W. (2014). *The non-invasive prenatal test (NIPT) for trisomy 21 - health economic aspects* (KCE Reports 222. D/2014/10.273/36). Health Technology Assessment (HTA) Brussels: Belgian Health Care Knowledge Centre (KCE).

Kater-Kuipers, A., de Beaufort, I. D., Galjaard, R. H., & Bunnik, E. M. (2020). Rethinking counselling in prenatal screening: An ethical analysis of informed consent in the context of non-invasive prenatal testing (NIPT). *Bioethics*, *34*(7), 671 - 678. https://doi.org/10.1111/bioe.12760

Kessels, S. J. M., Carter, D., Ellery, B., Newton, S., & Merlin, T. L. (2019). Prenatal genetic testing for cystic fibrosis: a systematic review of clinical effectiveness and an ethics review. *Genetics in Medicine*, *22*, 258-267. https://doi.org/10.1038/s41436-019-0641-8

Kibel, M., & Vanstone, M. (2017). Reconciling ethical and economic conceptions of value in health policy using the capabilities approach: A qualitative investigation of Non-Invasive Prenatal Testing. *Social science & medicine*, *195*, 97-104. https://doi.org/10.1016/j.socscimed.2017.11.024

Legault, G.-A., Suzanne, K., Bellemare, C. A., Béland, J.-P., Bernier, L., Dagenais, P., Daniel, C.-É., Gagnon, H., Parent, M., & Patenaude, J. (2018). Revisiting the Fact/Value Dichotomy: A Speech Act Approach to Improve the Integration of Ethics in Health Technology Assessment. *Open Journal of Philosophy*, *8*(5), 578-593. https://doi.org/10.4236/ojpp.2018.85042

Lucivero, F. (2016). *Ethical Assessments of Emerging Technologies*. Springer.

O'Rourke, B., Oortwijn, W., Schuller, T., & International Joint Task, G. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, *36*(3), 187-190. https://doi.org/10.1017/S0266462320000215

Shams, L., Akbari Sari, A., & Yazdani, S. (2016). Values in Health Policy - A Concept Analysis. *International Journal of Health Policy and Management*, *5*(11), 623-630. https://doi.org/10.15171/ijhpm.2016.102

Stapleton, G., Dondorp, W., Schroder-Back, P., & de Wert, G. (2019). A Capabilities Approach to Prenatal Screening for Fetal Abnormalities. *Health Care Analysis*, *27*, 309-321. https://doi.org/10.1007/s10728-019-00379-5

Stegenga, J. (2018). *Medical nihilism*. Oxford University Press.

Svenaeus, F. (2018). Phenomenology of pregnancy and the ethics of abortion. *Medicine, Health Care and Philosophy*, *21*(1), 77-87. https://doi.org/10.1007/s11019-017-9786-x

van der Meij, K. R. M., Sistermans, E. A., Macville, M. V. E., Stevens, S. J. C., Bax, C. J., Bekker, M. N., Bilardo, C. M., Boon, E. M. J., Boter, M., Diderich, K. E. M., de Die-Smulders, C. E. M., Duin, L. K., Faas, B. H. W., Feenstra, I., Haak, M. C., Hoffer, M. J. V., den Hollander, N. S., Hollink, I., Jehee, F. S., . . . Dutch, N. C. (2019). TRIDENT-2: National Implementation of Genome-wide Non-invasive Prenatal Testing as a First-Tier Screening Test in the Netherlands. *The American Journal of Human Genetics*, *105*(6), 1091-1101. https://doi.org/10.1016/j.ajhg.2019.10.005

Vanstone, M., Yacoub, K., Winsor, S., Giacomini, M., Nisker, J. (2015). What is "NIPT"? Divergent characterizations of noninvasive prenatal testing strategies. *AJOB Empirical Bioethics*, *6*(1), 54-67. https://doi.org/10.1080/23294515.2014.993102


4

Chapter 4

Assessing medical devices: A qualitative study from the VALIDATE perspective

Bloemen, B., & Oortwijn, W. (2024). Assessing medical devices: a qualitative study from the validate perspective. *International Journal of Technology Assessment in Health Care*, *40*(1), e29.
https://doi.org/10.1017/s0266462324000254

ABSTRACT

Objectives: Our objective was to explore procedures and methods used at health technology assessment (HTA) agencies for assessing medical devices, and underlying views of HTA practitioners about appropriate methodology, to identify challenges in adopting new methodology for assessing devices. We focused on the role of normative commitments of HTA practitioners in the adoption of new methods.

Methods: An online survey, including questions on procedures, scoping and assessments of medical devices, was sent to members of the International Network of Agencies for Health Technology Assessment (INAHTA). Interviews were conducted with survey respondents, and HTA practitioners involved in assessments of Transcatheter Aortic Valve Implantation, to gain an in-depth understanding of choices made in, and views about, assessing medical devices. Survey and interview questions were inspired by the *VALues In Doing Assessments of health TEchnologies* (VALIDATE) approach towards HTA that states that HTA addresses value-laden questions and information.

Results: Current practice of assessing medical devices at HTA agencies is predominantly based on procedures, methods and epistemological principles developed for assessments of drugs. Both practical factors (available time, demands of decision-makers, existing legal frameworks and HTA guidelines), as well as commitments of HTA practitioners to principles of evidence-based medicine make adoption of new methodology difficult.

Conclusions: There is a broad recognition that assessments of medical devices may need changes in HTA methodology. In order to realize this, the HTA community may require both a discussion on the role, responsibility, and goals of HTA, and resulting changes in institutional context to adopt new methodologies.

INTRODUCTION

Health Technology Assessment (HTA) aims to *inform* decision-makers by assessing the *potential value* of health technologies (O'Rourke et al., 2020). Therefore, HTA practitioners (those responsible for conducting assessments, including scoping, collecting, synthesizing, and interpreting available evidence) need to identify evidence that can answer policy-relevant questions about the potential value of health technology, requiring decisions on which information can be regarded reliable and relevant. Current discussions about appropriate HTA methodology for assessing (high-risk) medical devices show that this is not an easy task. Based on differences between medical devices and drugs, scholars argue that HTA methodology for medical devices should be adapted to 1) *integrate other types of evidence* (e.g., real-world evidence) to address the lack of evidence from randomized clinical trials, and capture the impact of iterative developments of devices on outcomes; 2) *broaden the scope of assessments* to capture organizational aspects (e.g., impact on healthcare capacity); and 3) *involve stakeholders in assessments* (e.g., making methodological decisions) to address context-dependence of outcomes and gather information on user experiences and preferences (J. J. Enzing et al., 2021; Joost J. Enzing et al., 2021; Fuchs et al., 2017; Ming et al., 2022; Pomey et al., 2020; Tarricone et al., 2017; Torbica et al., 2022).

Despite these calls to assess medical devices differently, previous studies have shown that HTA agencies use similar methodology for assessing drugs and medical devices (Bluher et al., 2019; Ciani et al., 2015; Joost J. Enzing et al., 2021; Fuchs et al., 2017; Ming et al., 2022). Although practical reasons like capacity problems and existing regulatory frameworks contribute to this uniformity, we argue that *normative commitments* of HTA agencies and practitioners also play a role. Inspired by the VALIDATE (*VALues In Doing Assessments of health TEchnologies)* approach, which emphasizes that the relevance and meaning of evidence considered in HTA depends on underlying values, we reasoned that both the value perspectives of stakeholders and HTA practitioners are instrumental in conducting assessments (Gert Jan van der Wilt et al., 2022; G. J. van der Wilt et al., 2022). This implies that activities of HTA agencies and practitioners are not solely guided by established HTA guidelines but are also influenced by practitioners' views on how HTA can improve outcomes of health technology for society. Given that HTA is often presumed to provide information about the *public value* of health technology, transcending particular interests, HTA practitioners and agencies are committed to methodological principles presumed to guarantee a *neutral* or *unbiased* evidence base for decision-makers (Boothe, 2021; Ducey et al., 2017; Gagnon et al., 2020). These commitments may conflict with new types of evidence, outcome measures and methodologies proposed for assessing medical devices.

To explore the significance of these commitments, besides practical challenges, in the adoption of new methodology (e.g., real-world data, stakeholder involvement) for (high-risk) medical devices assessments, we conducted a survey and interview study among relevant HTA agencies. Our objective was to map the procedures and methodologies currently used by these HTA agencies, and to retrieve the views of HTA practitioners about the role of HTA, stakeholder involvement, and appropriate evidence in HTA.

METHODS

We used a semi-structured survey to gather information on current practice of assessing (high-risk) medical devices by HTA agencies (i.e., legal frameworks, procedures, methods). We defined high-risk medical devices as Class IIb and Class III medical devices according to the European Regulation on Medical Devices – Regulation (EU) 2017/745. Additionally, via semi-structured interviews with HTA practitioners we explored, building on previous findings in literature, whether changes in HTA methodology may conflict with their views (Ducey et al., 2017). Specifically, we were interested in their perspectives on the role of HTA in decision-making, their responsibilities in the conduct of HTA, stakeholder involvement, and what constitutes appropriate evidence, particularly for assessing medical devices. Both survey and interview questions, inspired by the VALIDATE approach and literature on HTA for medical devices, also delved into the value-laden aspects of HTA procedures and methodology. See also Supplementary Figure 1 for a schematic illustration of the qualitative approach taken in this study.

Survey

The online survey was developed based on our previous work regarding deliberative HTA processes (targeting stakeholder involvement), normative analysis, and desk research on challenges in assessing medical devices (Bluher et al., 2019; Ciani et al., 2015; J. J. Enzing et al., 2021; Joost J. Enzing et al., 2021; Fuchs et al., 2017; Ming et al., 2022; Oortwijn et al., 2022; Oortwijn et al., 2020; Pomey et al., 2020; Tarricone et al., 2017; Torbica et al., 2022; Gert Jan van der Wilt et al., 2022). Questions focused on *institutional context* and *current HTA processes*; *scoping*; and *assessing medical devices* (the types of evidence used, aspects assessed, stakeholder involvement). A draft version was tested by an HTA practitioner at a national HTA agency from our network. Based on received feedback, minor changes were introduced to clarify questions. The survey (and invitation email) is provided as Supplementary file 1.

We invited members of the International Network of Agencies for Health Technology Assessment (INAHTA), except research organizations and regulatory agencies

(n=3), and one institute which we know does not assess medical devices. We targeted specific persons, known from our networks and/or who assess medical devices; otherwise contact persons mentioned on the INAHTA website (www.inahta.org) were approached. Data collection occurred via the online tool CheckMarket, between January-February 2023, including two biweekly reminders. We asked respondents for consent to analyze results and assured confidentiality (no attribution is made to specific persons). We also asked consent to contact them for an interview.

Descriptive statistics (frequencies, presented as percentages) derived from the CheckMarket tool were used to summarize findings. When needed, websites, literature, and publicly available guidelines and HTA reports from HTA agencies (retrieved by manually searching on their websites) were reviewed to clarify responses and gain an in-depth understanding of processes and methodology used for assessing medical devices, see also Supplementary file 2.

Interviews

We invited (via email) HTA practitioners that responded to the survey and indicated to be contacted, and specifically invited HTA practitioners involved in assessing Transcatheter Aortic Valve Implantation (TAVI), to explore choices made in real-world assessments. TAVI was chosen as example because it is a high-risk medical device, already implemented in clinical practice, and full HTAs are conducted in different jurisdictions. It is a minimally invasive technology aimed at inoperable patients with symptomatic severe aortic valve stenosis. Since its Conformité Européene (CE) marking in 2007, usage expanded to patients at high, intermediate, and low surgical risk. We focused on assessments of TAVI for patients at low risk for surgical complications (i.e., eligible for the standard treatment, Surgical Aortic Valve Replacement, SAVR) which became standard care for patients 75 years old and above (Vahanian et al., 2022). In November 2022, the HTA database (https://database.inahta.org/) was used to search for full HTA reports, using the MeSH term 'Transcatheter Aortic Valve Replacement', which retrieved available HTA reports (on TAVI for low risk patients) from Health Information and Quality Authority – HIQA (Ireland), Ontario Health (Canada), and the Norwegian institute of Public Health (Health Information and Quality Authority (HIQA), 2019; Himmels et al., 2021; Ontario Health, 2020a). In addition, a manual search retrieved a report by Haute Autorité de Santé (France) (Haute Autorité de Santé (HAS), 2020).

We developed a semi-structured interview guide based on relevant literature on normativity in HTA, challenges in assessing medical devices / TAVI, and the VALIDATE approach. Interviews comprised three parts: (i) professional background, experience, and current position of the HTA practitioner; (ii) questions on context and deci-

sions made in developing the respective HTA report on TAVI, or questions to clarify answers given to survey questions; (iii) personal views of the HTA practitioner on roles and responsibilities of HTA, and methodological issues in assessments of medical devices. The interview guide was iteratively updated based on experiences with conducting the interviews. Given the explorative nature of our study, data saturation was not a target.

The lead author (BB; PhD candidate in HTA) conducted online interviews (using Microsoft Teams) between February and May 2023, having a duration between 1-1.5 hours. All interviews were audio-recorded and summarized; interviewees were asked to provide feedback on the summary to clarify any misunderstandings. Prior to participation, oral consent was obtained from all interviewees, who were informed about the study objectives through invitation mails and the concept interview guide.

More information about the preparation of interviews, and the interview guide, can be found in Supplementary file 3.

The basis for analyzing the interviews were the updated summaries (based on feedback from the interviewees), including information retrieved from websites of respective HTA agencies, HTA reports and publicly available guidelines. Thematic analysis was used, which is a method for identifying, analyzing, and reporting themes within the data. Because interviews were conducted to provide in-depth information, complementary to the surveys, about the context and reasons (including views of HTA practitioners) behind current processes and methodology for assessing medical devices (see also Supplementary Figure 1), main themes from the survey (scoping, types of evidence, aspects of devices being assessed, stakeholder involvement) were the starting point for analyzing the interviews. The lead author used a process of inductive comparison and reasoning to identify subthemes that reflect the content of conducted interviews.

The Consolidated criteria for reporting qualitative research (COREQ) checklist was used to ensure methods, results and discussion were reported appropriately (Tong et al., 2007).

RESULTS

Study participants

We invited fifty contact persons of INAHTA member agencies, of which twenty-two (response rate of 44 percent) responded to the survey. Two respondents answered less than 50 percent of the main questions and were excluded from the analysis. In addi-

tion, five respondents were excluded as they were not involved in the assessment of medical devices. In total, we analyzed fifteen survey responses, including twelve fully completed surveys and three agencies that provided meaningful answers (answering more than 50 percent of questions on either scoping and / or assessment). Among these, eight were willing to be interviewed (53 percent).

Four accepted our invitation for an interview (50 percent) from HTA agencies in the Netherlands, Spain, Taiwan, and Colombia. Of the authors of the four retrieved HTA reports on TAVI who were invited for an interview (n = 9), two accepted our invitation, one did initially agree to be interviewed but did not respond after sending multiple reminders to set an interview date, one declined participation, two referred to a co-author, and three did not respond at all. When an author of an HTA report on TAVI accepted the invitation, other authors of the same HTA report were not invited.

Table 1 provides an overview of participating HTA agencies. Additional information about interview participants is reported in Supplementary Table 1. Most participating agencies are governmental institutions (29 percent), or institutes with a government function (47 percent, independent from a Ministry of Health), advising policy makers on national policy decisions (e.g., allocation of public resources, reimbursement by health insurance) on medical devices.

Table 1. Overview of HTA agencies that (partially) completed the survey and / or participated in the interviews.

Institution, country / region	Type of institution [a]	Completed the survey?	Participated in interviews?
Avalia-t / ACIS, Spain (Galician region)	3	Yes	Yes (on medical devices)
AQuAS, Spain, Catalonia	3	Yes	No
CADTH, Canada	4	Yes (partial response)	No
CDE / HTA, Taiwan	2a	Yes	Yes (on medical devices)
FOPH, Switzerland	2a	Yes	No
G-BA, Germany	5	Yes	No
Health Technology Wales, Wales	4	Yes (partial response)	No
IECS, Argentina	1	Yes	No
IETS, Colombia	4	Yes	Yes (on medical devices)
IQWiG, Germany	4	Yes	No
MaHTAS, Malaysia	2a	Yes	No
NECA, South Korea	4	Yes	No
NIPH, Norway	2a	Yes	No
SR-NRCHD, Kazakhstan	2a	Yes (partial response)	No
ZIN, The Netherlands	4	Yes	Yes (on medical devices)
Ontario Health, Canada	4	No	Yes (on TAVI)
HIQA, Ireland	4	No	Yes (on TAVI)

Notes: [a] categorization based on Fuchs et al 2017: 1 = independent academic research entity, 2 = Governmental institutions (a. national, b. regional), 3 = Regional Ministries of Health / Social Affairs including a related department, 4 = Independent entities with function as governmental institution, 5= Non-departmental public body with legislative function.

Abbreviations: Avalia-t / ACIS: Unidad de Asesoramiento Científico-técnico (Avalia-t), Axencia Galega de Coñecemento en Saúde (ACIS); AQuAS: Agència de Qualitat I Avaluació Sanitàries de Catalunya; CADTH: Canadian Agency for Drugs and Technologies in Health; CDE/HTA: Center for Drug Evaluation Health Technology Assessment; FOPH: Federal Office of Public Health; G-BA: Gemeinsamer Bundesausschuss; IECS: Instituto de Efectividad Clínica y Sanitaria; IETS: Instituto de Evaluación Tecnológica en Salud; IQWiG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; MaHTAS: Malaysian Health Technology Assessment Section; NECA: National Evidence-based healthcare Collaborating Agency; NIPH: Norwegian Institute of Public Health; SK-NRCHD: Salidat Kairbekova National Research Center for Health Development; ZIN: Zorginstituut Nederland; HIQA: Health Information and Quality Authority.

Institutional context, procedures for assessing medical devices

Survey respondents and interviewees were asked about how assessments of medical devices are initialized and differences with HTA processes for drugs (see Supplementary Table 1 and 2).

In general, agencies have similar procedures for assessing devices and drugs, but processes may differ in duration, initialization of assessments, and evidential requirements, being more heterogeneous for devices. The definition of medical devices varies widely: five agencies use EU directives that include specific definitions of (classes of) medical devices, three agencies use a definition from their national law, while five agencies report a broader definition of *health technology* that includes devices.

When a medical device is introduced to a market (after regulatory approval), HTA agencies are mostly asked to conduct assessments that inform re-imbursement decisions at the request by decision-makers (73 percent), followed by an application of the manufacturer and identification via horizon scanning (47 percent). Although there are experiments with involving stakeholders in deciding which devices need an assessment, this is often limited to proposing topics or providing feedback on a draft HTA protocol, and the final decision rests with decision-makers and sometimes HTA practitioners. Interviewees also mentioned that decision-makers' needs often determine which assessments are initiated (see also Table 3).

Scoping

Nine survey respondents (60 percent) reported that their agency has (publicly available) guidelines or documents on scoping applicable to medical devices, see Table 2. Guiding principles of the scoping process are transparency (78 percent), overarching goals of the HTA agency or healthcare system, impartiality, consistency, verifiability (all 67 percent), whereas inclusivity (44 percent), timeliness (44 percent) and efficiency (33 percent) are less frequently mentioned. Scoping often focuses on defining the health technology and its comparators needing an assessment (67 percent), whereas defining the health problem is rarely the objective of scoping (22 percent).

Table 2. Overview of answers provided to survey questions on scoping.

Question	Answers	Percentage
Are guidelines / documents describing the process of scoping applicable to the evaluation of high-risk medical devices present in your country / region? (n=15)	Present and publicly available	27%
	Present but not publicly available	33%
	Not present	40%
What are the guiding principles of the scoping process described in the guidelines? [multiple answers possible] (n=9)	Transparency	78%
	Overarching goals of HTA agency or health system	67%
	Impartiality	67%
	Consistency	67%
	Verifiability	67%
	Inclusivity	44%
	Timeliness	44%
	Efficiency	33%
What is the main focus of the scoping process described in the guidelines? (n=9)	Defining the health technology and the alternative technology(s) against which the health technology under assessment should be compared	67%
	Defining to what extent the health problem under study can be addressed (i.e., are non-technological interventions that could be proposed to address the health problem being considered)	22%
	Other, please specify: - In relation with the health condition, we used to define the baseline characteristics of population; moreover, we defined the outcomes that will be assessed in the report (n=1)	11%
How are stakeholders selected to be involved in the scoping process (if described in the guidelines)? (n=8)	By invitation or appointment (closed procedure)	50%
	Using a hybrid approach	38%
	Open to all who qualify (application process)	13%
	Open to all (public call)	0%
	Nominated by relevant interest groups (nomination process)	0%
Which input is requested from stakeholders in the scoping process? [multiple answers possible] (n=8)	Background information provided by stakeholders (e.g., experiential knowledge that can help in defining the research question; ideas about the plausibility of different interventions in addressing the health problem; different views on how to define the health problem)	88%
	The contribution of stakeholders is primarily focused on providing value perspectives and selecting relevant outcomes	63%
	Stakeholders are explicitly involved in determining the objectives of the assessment	50%

Table 2. *Continued.*

Question	Answers		Percentage
Which stakeholders are explicitly involved via consultation (i.e., structured process to collect feedback among groups of stakeholders on specific decisions via e.g., surveys, interviews, expert panels, patient testimonies); and which stakeholders are involved via participation (i.e., active engagement in deliberations and open exchange on argumentation and evidence)? [multiple answers possible] (n=8)	***Stakeholder***	***Consultation (relative position)***	***Participation (relative position)***
	Providers of care (e.g., clinician, nurse, hospital board member etc.)	88% (1)	88% (1)
	Experts in medicine	88% (1)	88% (1)
	Patient's organization	75% (2)	75% (2)
	Experts in (health) Economics	63% (3)	88% (1)
	Policy makers	63% (3)	50% (4)
	Experts in Epidemiology	50% (4)	63% (3)
	Manufacturers	50% (4)	50% (4)
	Experts in Ethics	38% (5)	50% (4)
	Experts in Healthcare Administration	38% (5)	38% (5)
	Payers / purchasers (e.g., health insurer, HMO etc.)	38% (5)	0% (8)
	Patients with the disease but not yet treated	25% (6)	13% (7)
	Patients with the disease and already treated with the comparator	25% (6)	25% (6)
	Patients treated with the new intervention	25% (6)	13% (7)
	Informal caregivers	25% (6)	13% (7)
	Experts in Patient / Public involvement	25% (6)	25% (6)
	Experts in Bioengineering	25% (6)	38% (5)
	Experts in Statistics	25% (6)	25% (6)
	Experts in Law	13% (7)	38% (5)
	Experts in Psychology	13% (7)	13% (7)
	Public / (organized) group of citizens	13% (7)	13% (7)
Which tool(s) are used for scoping (if described in guidelines)? [multiple answers possible] (n=8)	Population Intervention Comparators Outcomes (PICO) tool		100%
	Technology Indication Comparison Outcome (TICO) tool		13%
	Other, please specify: - We also use the PICOD (D=design) tool (n=1)		13%

Table 2. *Continued.*

Question	Answers	Percentage	
Which methods are used for selecting comparators and outcome measures to be considered in an assessment? [multiple answers possible] (n=8)		***Comparators (relative position)***	***Outcome measures (relative position)***
	Literature or document review	100% (1)	88% (1)
	Interviews with health professionals relevant to the disease under study	63% (2)	50% (2)
	Interviews with other relevant experts	50% (3)	25% (4)
	Focus groups with a mix of relevant experts, including health professionals and / or patients	38% (4)	38% (3)
	Interviews with patients suffering from the disease under study	25% (5)	25% (4)
	Surveys of relevant stakeholders	25% (5)	38% (3)
	Other, please specify: - Interviews used to be doing by telephone or email (n=1) - We have an evidence assessment group and patient and public involvement group that consider and agree on relevant outcomes and methods (n=1)	25% (5)	25%
	Focus groups with health professionals relevant to the disease under study	13% (6)	25% (4)
	Focus groups with patients suffering from the disease under study	13% (6)	13% (5)
	Focus groups with other relevant experts	13% (6)	25% (4)

Eight agencies (53 percent) have a description of stakeholder involvement included in their guidelines for scoping. Input requested from stakeholders is primarily providing background information (88 percent), and information on their value perspectives and ideas about relevant outcome measures (63 percent). Stakeholders are recruited by invitation (50 percent) or a combination of closed and open procedures (38 percent). The stakeholders mostly involved in scoping are providers of care, experts in medicine, patients' organizations, experts in health economics, and policy makers, whereas involvement of patients themselves (not represented via a patients' organization), informal caregivers, and the public (organized group of citizens) is low (25 percent or less). Some groups of stakeholders are mostly involved in a specific way: payers and purchasers primarily via *consultation* (i.e., asked to provide written feedback); experts in law primarily via *participation* (i.e., involved in deliberations and meetings).

When it comes to methodology used in scoping, the Population Intervention Comparators Outcomes (PICO) tool is always used. This tool structures the scoping

process, focusing on specifying the research question. Comparators and outcomes are primarily selected based on literature reviews, interviews with health professionals and other relevant experts, and focus groups with a mix of experts (including health professionals and patients). In some cases, relevant outcome measures are selected by surveying relevant stakeholders.

Scoping was also discussed during interviews, confirming that it is often technology-focused, based on literature and expert opinion (see also illustrative fragments from interviews in Table 3 and Supplementary Table 3). At some agencies, stakeholders are consulted about whether they agree with the scope and to raise comments about whether there is anything missing. Interviews on TAVI showed that expectations concerning the health problem (aortic valve stenosis) for which TAVI is held to be a solution, and what the relevant comparators are, are not explicitly questioned during scoping and assumed to be similar to what is claimed by health professionals and / or described in literature. Consequently, TAVI is only compared with the current standard in clinical practice (SAVR) and alternative interventions (e.g., preventative treatment, drug-based treatment etc.) seem not to be considered. The scoping processes conducted for TAVI are also not reported, only their output is part of the final HTA report (e.g., specifications of objectives or terms of reference for the assessments), or a brief description of input collected from stakeholders during scoping is included in the report (e.g. the NIPH report on TAVI includes an appendix on 'user involvement') (Haute Autorité de Santé (HAS), 2020; Health Information and Quality Authority (HIQA), 2019; Himmels et al., 2021; Ontario Health, 2020a).

Interviewees also mentioned that the scope of an assessment is often already predetermined by legal requirements and/or official HTA guidelines for conducting assessments (see Supplementary Table 1 and 3).

Table 3. Illustrative fragments from summaries of interviews.

Theme	Fragments
Scoping	*Not for TAVI for low surgical risk patients, because at the time of the HTA SAVR was considered to be the proper comparator as it was considered the standard of care according to experts in the field. If there would be another relevant comparator, that intervention would already have been tried in the treatment of these patients. And at the time of the HTA, patients at this stage of the disease always received SAVR. We don't question this golden standard in clinical practice. […] Not in the case of TAVI because no other relevant comparator was identified during scoping and this was validated by experts in the field. Additionally, the quantitative and qualitative preferences literature, and engagement with patients, did not identify any other relevant comparators. [Interview #3]* *As part of the prioritization process, we often provide an initial recommendation about what is required for the topic. For some topics, we will conclude that there is insufficient evidence to support an HTA or that the only information needed is on clinical effectiveness. If it is agreed upon that an HTA is needed and possible, it is discussed with the decision-maker what information is needed for them to make a decision. The outcome of this is the terms of reference for the report, and stakeholders are asked to provide input (e.g., do they miss anything?). [Interview #6]*
The use of different types of evidence in assessments of medical devices	*No, it's not a black and white matter. There is some recognition at HTA agencies that real-world data and observational data should be considered in assessments. How I see it is that it renders a methodological inquiry rather than a concern on neutrality and impartiality. The challenge is in integrating these approaches in assessments while simultaneously adhering to the current legal frameworks which are still focused on RCT data. But which types of data are used should depend on the type of questions raised by an assessment.* [Interview #2] *The requirements on evidence for assessing medical devices should not be different from those for assessing drugs. However, for medical devices the availability of RCTs is often limited, but we always use the highest level of evidence that is available for a given outcome. Therefore, observational data and real-world data can be used to assess medical devices when deemed appropriate.* […] *The use of observational and / or real-world data for assessing TAVI was part of the discussion before the methodology and literature search was finalized (it was determined during the scoping phase). If observational studies provide information on the same outcomes and for the same follow-up duration as RCTs, and RCTs are of high quality (no risk of bias), RCTs are preferred because they are higher in the hierarchy of evidence. If RCTS are available, observational studies are considered only if they provide additional information to RCTs (i.e., in terms of types and/or duration of outcomes, e.g., longer-term outcomes) or if observational studies are of comparable quality to RCTs. In the case of TAVI, there were two high-quality RCTs available and no information was missed, i.e., there were no observational studies known that could add any relevant information.* [Interview #3] *What we try to do to address these challenges with medical devices is to make comparisons (e.g., comparing outcomes of interventions using different devices), because that is really important. […] Because, from the perspective of the decision-maker (Ministry of Health) you are focused on the health of the population and the healthcare system, not on a single device. You need information that allows you to compare different technologies to make decisions on that level, to know what you sacrifice if you decide to invest in a particular technology (because resources are limited).* [Interview #5]

Table 3. *Continued.*

Theme	Fragments
Aspects considered in assessments of medical devices	*Quality of life depends on the medical device. We can't have the quality of life evidence for every medical device. In general, the outcomes depend on the device. […] We look at RCTs, and if not available we use observational studies. If they have reported on quality of life we will include the information in the report, but we do not only focus on it. […] I do think that patient experiences and quality of life is important as a reference for reimbursement decisions, but we do not just focus on patient opinions during the assessment and do not use quality of life as a search key word. [Interview #1]* *Sometimes decisions are based on things like political expediency, or some other reasons that we cannot capture as part of the evidence base. For example, in the case of orphan drugs, which are not cost-effective, there may be reasons to reimburse them because of care for a group of people who don't have other options. But an HTA struggles to capture that information because it is very hard to do that objectively, although we can highlight it under patient, social and ethical issues. It is not the role of an HTA agency to get everything that is required for the decision, we have to look at the things we can manage objectively. [Interview #6]* *Although the relevance of ethical analysis is acknowledged, in practice it is mostly not conducted. Important barrier is that the assumption is that it is sufficient that clinicians, health economists, epidemiologists, HTA practitioners, can take ethical aspects into account as part of their analysis. So it is not recognized as a separate domain or analysis step. There is no strong perceived need for an ethicist being explicitly involved in these domains, or a formal integration of an additional ethical analysis. […] It seems to be no one's concrete responsibility, or all stakeholders (HTA practitioners, decision-makers etc.) refer to each other. There are different views about what is the appropriate place to address this, some would say that it is the responsibility for political parties or decision-makers.* [Interview #2]
Stakeholder involvement in assessments of medical devices	*In our country, the HTA report is used for reimbursement decisions. When conducting an assessment, we think about the benefits of a technology for society. This means it is important that there is a link with potential benefits for the patients. […] The patient is the most important stakeholder, but not the only one. The perspective and satisfaction of the clinician is also important. For a good use of medical devices, the clinicians and patients are both needed. Both influence the safety and efficacy of medical devices. […] We have to focus on the issues considered relevant by Ministry of Health, both specific issues as a given medical device or wider as pseudo therapies assessments directed to avoid population use them instead of their treatments. [Interview #4]* *We have been engaging the community and stakeholders in our analysis, but this is hard because people in our country are not used to being involved in these analyses. Therefore, we have been training patients and families about HTA. In addition, the results of an HTA are presented to panels consisting of healthcare professionals that are going to use the device, stakeholders (excluding industry), and the government. These can provide feedback on the results. And a bioethicist and lawyer are usually part of an HTA team, conducting an ethical analysis within the limits of our national law. [Interview #5]* *Therefore, asking patients whether they can recall a particular experience (prompted by anectodical evidence) may lead to confirmation bias. We cannot base conclusions on anectodical evidence. What we can do is saying that there is some evidence that some patients are unhappy with the intervention, but that it is unclear whether that is a general experience. […] In the case of pharmaceuticals, manufacturers are very clever and know how to involve patients to maximize the chances of a good outcome. For medical devices the manufacturers are not that mature yet, and they involve patients to tell them what is important to them. Only patients can tell you what is important them, and patients are the ones you ultimately want to help. But this needs education, to inform patients about how HTA processes works, and which evidence is required. But it can only be for the good of HTA if patients are more involved and have a better understanding of what is required. But we have to be careful that we don't end up with people that are gaming the system, it is important that the evidence is impartial. And it is important that people think about the greater good.* [Interview #6]

Assessment

Use of different types of evidence

Participating agencies predominantly use traditional types of studies (e.g., RCT, meta-analysis, systematic review), see Table 4. Also, the use of qualitative research methods is less than 50 percent and confined to obtaining information about patients' perspectives and experiences, to contextualize quantitative evidence, and it has no role as formal evidence in assessments.

Table 4. Overview of answers provided to survey questions on evidence considerations in assessments of high-risk medical devices.

Question	Answers	Percentage
Which type of studies are primarily considered by your HTA agency when assessing high-risk medical devices? [multiple answers possible] (n=14)	RCT	100%
	Meta-analysis	71%
	Systematic reviews	64%
	Nonrandomized controlled prospective cohort studies	29%
	Primary studies	29%
	Other, please specify: - Comparative study with a control group (*n* = 1) - Other HTA reports (*n* = 1) - Relevant real-world evidence from the healthcare system (if available) (*n* = 1)	21%
Are qualitative research methods (e.g., interviews, focus groups) used by your HTA agency for assessing high-risk medical devices? (n=14)	Yes	43%
	No	57%
For which types of analyses are qualitative research methods considered? [open question] (n=14)	*To assess the perspectives and satisfaction of patients regarding the medical device used* *For patient perspectives and experiences, caregiver perspectives and experiences, implementation considerations, ethical analysis* *Mainly patient and public involvement aspects, e.g., we use available qualitative evidence from literature or primary evidence we collect directly using interviews, focus groups etc.* *Yes, we evaluated medical device re-manufacturing for the health ministry using a multidimensional approach.* *For assessment of patients' perspectives; experts and Qualitative Evidence Synthesis (QES)* *For signaling inappropriate use and for agenda-setting, not for formal assessments*	

Table 4. *Continued.*

Question	Answers	Percentage
What are the considerations with regard to assessing the quality of evidence when conducting an evaluation of high-risk medical devices? [open question] (n=15)	*GRADE (n = 6)* *We consider the internal validity of the studies assessed (i.e., risk of bias) and the applicability to our health system and target population (external validity) in relation with the population (or subgroup of patients with a given baseline characteristics) in which the medical device evaluated is intended to use.* *Because high-risk medical devices sometimes have ethical issues impeding the conduct of double-blind trials, evidence is sometimes from open-label or without comparator trials, this might affect the quality of evidence.* *Similar to other technologies (n = 2).* *Assessment of certainty of study results.* *Study design, population included in the study, comparator, risk of bias, confounding factors.* *PICO relevance, published in peer-reviewed journals, if necessary we use GRADE.*	
Is the quality of evidence interpreted differently for various types of methods (qualitative vs quantitative methods? [open question] (n=15)	*"No."* *"Yes." (n=2)* *"Yes, depending on the research questions and studies being included."* *"If qualitative is carried out through interviews or focus groups, it may be more open-ended, and many different views and opinions may be collected, or the existing evidence results may be summarized through systematic review, which is less likely understand the actual effect size, and the evidence may come from multiple sources, would lower the quality of the evidence. However, if it is quantitative, the effect size can be provided by statistical methods, but it may also be limited by the quality of the data source and affect the quality of the evidence."* *"The certainty and quality of evidence is interpreted according to the specific analysis. There is not the same framework to assess clinical effectiveness and to assess perceived needs from the community because the objectives and the potential outcomes are different."* *"Yes. We do not apply/complete formal QA checklists as we operate a rapid review model. But our researchers are highly experienced and apply quality assessment implicitly, drawing out any key issues."* *N/A; Qualitative research methods are not (formally) considered in an assessment (n = 6)*	

Survey responses and interviews with HTA practitioners show their acknowledgment of challenges involved in collecting data for medical devices, but that they also think the same epistemic principles apply (e.g., evidence hierarchy, risk of bias) and that alternatives like real-world evidence introduce more uncertainty (see Table 3 and 4, and Supplementary Table 3). What is mentioned several times by HTA practitioners is that they only consider *comparative data*, i.e., data that allows you to draw conclusions about the *relative effectiveness* of different health technologies, which is considered important from the viewpoint of the purpose of HTA (to inform decisions on the level of the healthcare system). The main reasons for considering real-world evidence are a) that this could address iterative developments in medical devices (i.e.,

traditional methods for gathering evidence cannot keep up with this pace of development), and b) to address the context dependency of medical devices (i.e., contextual factors in 'real-world' circumstances).

Interviews on TAVI showed (see Table 3 and Supplementary Table 3) that other data types were considered by HTA agencies but not used when assessing safety or comparative clinical effectiveness of medical devices because they were deemed to provide no additional information with respect to available (high-quality) RCT data. The HTA reports on TAVI also show this reliance on RCT data, only one agency (i.e., HIQA) reported findings of registries in their safety assessment but these were only used as an addition to RCT data. The data from registries was presented only narratively and without any explicit critical appraisal of their quality (besides evaluating the relevance and appropriateness of the included patient populations in registries) (Health Information and Quality Authority (HIQA), 2019).

Aspects considered in assessment

Aspects primarily considered in assessments of medical devices are *clinical effectiveness* (100 percent), *safety* (93 percent), *costs and economic implications* (79 percent), and quality of life (71 percent); followed by organizational aspects (64 percent), and legal and ethical issues (both 50 percent); see Supplementary Table 4.

Interviewees express a lack of expertise, time and capacity to consider a broader spectrum of aspects, and that explicit consideration of ethical issues is not always seen as the responsibility of HTA practitioners or is not recognized as requiring explicit attention (see Table 3 and Supplementary Table 3). The inclusion of a broader spectrum of aspects is also limited due to legal frameworks that pre-define a narrower scope for assessments.

For TAVI, Ontario Health assessed a broad range of aspects (clinical effectiveness, safety, cost-effectiveness, budget impact, values and preferences of patients and informal caregivers), and these were integrated in the conclusions and recommendations (Ontario Health, 2020a, 2020b; Smith & Argaez, 2019). Patient preferences were included by reviewing published qualitative and quantitative preferences evidence, and direct engagement of patients with lived experience with TAVI. Ethical issues were not assessed because during scoping it was concluded that there was no need for it. At HIQA, safety, clinical effectiveness, cost-effectiveness, budget impact, and organizational aspects (e.g. impact on healthcare capacity) of TAVI were assessed, whereas ethical issues were only described (with equity as a primary concern) (Health Information and Quality Authority (HIQA), 2019). NIPH and HAS assessed safety,

clinical effectiveness, cost-effectiveness and budget impact of TAVI (Haute Autorité de Santé (HAS), 2020; Himmels et al., 2021).

Stakeholder involvement

Stakeholder involvement during assessment is confined to collecting evidence and reviewing its plausibility, and their role in making methodological decisions is limited, see Table 5. Stakeholders involved in all facets of conducting an assessment are patient organizations, providers of care, policy makers, payers / purchasers, and experts in medicine, health economics, epidemiology, ethics, and law. Patients (not represented by an organization), manufacturers, and informal caregivers are involved in collecting evidence, but almost excluded from making methodological decisions and reviewing evidence.

Table 5. Overview of answers provided to survey questions on stakeholder involvement in assessments of medical devices.

	Involved in collection of evidence		Involved in making methodological decisions		Involved in reviewing plausibility of evidence reports	
Are stakeholders involved in assessments, at which stage and how?	Yes (n=8) (62%) No (n=5) (38%)		Yes (n=3) (23%) No (n=10) (77%)		Yes (n=8) (62%) No (n=5) (38%)	
	Consultation	***Participation***	***Consultation***	***Participation***	***Consultation***	***Participation***
Patient's organization	75%	75%		33%	75%	25%
Providers of care (clinician, nurse, hospital board member etc.)	63%	63%	33%	67%	63%	38%
Patients with the disease but not yet treated	50%	13%			13%	13%
Patients with the disease and already treated with the comparator	50%	25%			13%	13%
Experts in Medicine	50%	63%		33%	63%	50%
Manufacturers	50%	50%			38%	
Patients treated with the new intervention	38%	13%			13%	13%

Table 5. *Contiued.*

	Involved in collection of evidence		Involved in making methodological decisions		Involved in reviewing plausibility of evidence reports	
Experts in (health) economics	38%	38%	33%	33%	38%	25%
Policy makers	38%	50%	33%	67%	50%	50%
Other	38%	13%	33%	33%	13%	25%
Informal caregivers	25%					
Experts in health-care administration	25%	38%			13%	
Experts in Epidemiology	25%	25%	33%	33%	38%	38%
Public / (organized) group of citizens	25%	13%			13%	
Experts in Ethics	13%	25%		33%	25%	25%
Experts in Patient and/or Public involvement	13%	13%			13%	
Experts in Bioengineering	13%				13%	13%
Experts in Psychology	13%	13%			25%	
Experts in Law		13%		33%		25%
Payers / purchasers (health insurer, HMO etc.)		38%	33%	33%	38%	13%
Experts in Sociology					13%	
Experts in Statistics					13%	13%

Interviewees expressed concerns with stakeholder involvement, mentioning potential threats to the impartiality and objectivity of the evidence base, as stakeholders may have vested interests and information provided by them may be skewed to be in favor of certain outcomes. Additionally, interviewees noted that stakeholders have a limited understanding of HTA processes (see Table 3 and Supplementary Table 3). Despite these concerns, interviewees acknowledge the importance of stakeholder involvement, especially for obtaining information on what are relevant outcomes, and to address challenges related to medical devices (e.g., for an appropriate use of medical devices the engagement of both clinicians and patients is needed; manufacturers can provide technical information about different generations of a device).

Regarding TAVI, stakeholder involvement was limited to a literature review of quantitative and qualitative research into patient preferences, direct engagement of patients

(excluding those at low surgical risk) and including a patient representative in the Expert Advisory Group. Their direct contributions involved providing feedback to drafts of HTA reports and sharing their experiences (Health Information and Quality Authority (HIQA), 2019; Himmels et al., 2021; Ontario Health, 2020a).

DISCUSSION

Despite the recognized need for changes in HTA methodology for medical devices, HTA agencies still resort to methods developed for assessing drugs and focus on assessing clinical aspects (safety, effectiveness) and cost-effectiveness using quantitative data. The broadening of who is involved (stakeholder involvement), what is assessed (which aspects of health technology), and which information is considered (e.g., real-world evidence, qualitative research), proposed by VALIDATE and other groups of experts in HTA, is not yet fully seen in current practice at HTA agencies (J. J. Enzing et al., 2021; Tarricone et al., 2017; Gert Jan van der Wilt et al., 2022). This discrepancy aligns with previous observations in surveys and reviews of guidelines (Bluher et al., 2019; Ciani et al., 2015; Fuchs et al., 2017; Ming et al., 2022). A recently published review of full HTA reports on TAVI for patients at low surgical risk, including the reports discussed in this study, also showed their predominant reliance on traditional RCT data and clinical outcome measures (Rumi et al., 2023). What our findings add to these studies is the understanding that, although HTA practitioners recognize the relevance of other types of evidence and methods, they are committed to existing epistemological principles (e.g., evidence hierarchy, risk of bias) that automatically downgrade non-RCT data, effectively excluding it from having impact on recommendations as previously observed in a study on real-world data policies for HTA of drugs (Makady et al., 2017). HTA scholars have also expressed critique on the quality of real-world evidence used in HTAs of high-risk medical devices (Klein et al., 2022).

Certain practical factors may also explain the reluctance to introducing new methods for assessing medical devices. Both in responses to survey questions and during interviews it became clear that HTA practitioners work under time pressure, must pay attention to demands of decision-makers, and need to adhere to existing legal frameworks and HTA guidelines, limiting their ability to experiment with new methodology. Therefore, HTA practitioners need a supportive environment (institutional context) that recognizes the importance of changing methodology for assessing medical devices.

In addition to this role of the environment, our interviews with HTA practitioners highlight some normative considerations also playing a role in sustaining the status quo. HTA practitioners frequently expressed concerns about how uncertainties and

biases associated with other types of evidence and stakeholders might influence the HTA process, potentially conflicting with the responsibility of HTA to guarantee an impartial ('neutral', 'objective') synthesis and interpretation of the available evidence. Therefore, the persistent use of traditional methods and evidence hierarchies, and the exclusion of stakeholders in parts of the process, may not only be the result of demands from decision-makers and official frameworks, but also because it is regarded the best way for ensuring this neutral role of HTA in decision-making. As observed in another interview study, HTA practitioners reliance on certain epistemological ideas may originate from ideas about the intrinsic value of HTA itself (Ducey et al., 2017).

Therefore, the adoption of new methodology for assessing medical devices at HTA agencies requires a discussion within the HTA community about the roles, responsibilities, and goals of HTA, and how to realize them. This includes acknowledging the implicit normative underpinnings of HTA processes and methods. For example, we agree with interviewees that the role and responsibility of HTA is to provide information on the *public value* of health technology, requiring expertise, processes and methods that ensure collected information is not influenced by interests. However, this does not imply that HTA practitioners need to refrain from making value judgments. Increasingly, HTA agencies and scholars acknowledge that conducting assessments requires making value judgments (Charlton et al., 2023). Although this may be a matter of degree, partly depending on the mandate of the HTA practitioner (e.g., working within a decision-making body or at an academic institute), every assessment requires making value-laden decisions about what are *good* methods and outcome measures to consider in evaluating a health technology (Hofmann et al., 2014). Given this recognition of the normativity of HTA, there is room to reflect upon whether current epistemic norms (like the strict adherence to a hierarchy of evidence) are still helpful in fulfilling the role of HTA in decision-making. Methods evolve, offering new ways for obtaining reliable data on effects of health technology, and HTA guidelines already provide some room to consider diverse outcome measures (Kinchin et al., 2023; Subbiah, 2023). Together with the broader HTA community (those using outcomes of HTA or being impacted by it), HTA practitioners may explore how this new methodology may help in assessing medical devices and improve the relevance of HTA (Freitas et al., 2023).

Future research on the impact of changes in HTA methodology on decision-making, and ideas of decision-makers and stakeholders about evidential requirements for different types of technology, could guide this collaborative rethinking of how new technologies, including medical devices, are assessed (Loblova et al., 2020).

Strengths and limitations

Although we managed to collect survey responses and conduct interviews with HTA practitioners working at seventeen different agencies, we cannot verify whether we collected all diversity in used methodology and views of HTA practitioners. Future research should try to include more agencies from different regions and interview multiple practitioners per agency. However, we are assured about the validity of our results by the convergence with findings of previous studies on HTA practice for medical devices and interviews with HTA practitioners about their views on appropriate methodology (Bluher et al., 2019; Boothe, 2021; Ciani et al., 2015; Ducey et al., 2017; Fuchs et al., 2017). By combining surveys and interviews, we have provided an in-depth understanding of *why* certain methodologies are used.

Although we tried to explore websites, published guidelines, and HTA reports of participating agencies, to verify findings, we were sometimes unable to retrieve or understand material because it was not (publicly) available (in English).

CONCLUSIONS

Despite recognizing the need for changes in HTA methodology for medical devices, HTA agencies predominantly use methods developed for assessing drugs. Both practical factors (available capacity, existing legal frameworks and HTA guidelines) and HTA practitioners' commitments to principles of evidence-based medicine make adoption of new methodology difficult. Therefore, the adoption of new methodologies at HTA agencies may require a discussion within the HTA community on the roles, responsibilities, and goals of HTA, and how these can be realized by changes in methodology and institutional context.

ACKNOWLEDGMENTS

We thank all the survey respondents and interviewees for their participation in this study. We also thank Professor Gert Jan van der Wilt for reviewing and providing insightful and valuable comments on an early version of this manuscript.

FUNDING STATEMENT

This research received a non-restricted grant from Edwards Lifesciences. The funding organization did not have any influence on the development of the data collection methods, analysis, and reporting.

COMPETING INTEREST

WO reports an unrestricted research grant from Edwards Lifesciences for the conduct of the study which was paid to Radboud University Medical Center. BB declares no conflict of interest related to this study.

SUPPLEMENTARY MATERIAL

The supplementary material for this chapter can be found online as part of the published article in the *International Journal of Technology Assessment in Health Care:* https://doi.org/10.1017/s0266462324000254.

REFERENCES

Bluher, M., Saunders, S. J., Mittard, V., Torrejon Torres, R., Davis, J. A., & Saunders, R. (2019). Critical Review of European Health-Economic Guidelines for the Health Technology Assessment of Medical Devices. *Frontiers in Medicine*, *6*, 278. https://doi.org/10.3389/fmed.2019.00278

Boothe, K. (2021). (Re)defining legitimacy in Canadian drug assessment policy? Comparing ideas over time. *Health Economics, Policy and Law*, *16*(4), 424 - 439. https://doi.org/10.1017/S1744133121000013

Charlton, V., DiStefano, M., Mitchell, P., Morrell, L., Rand, L., Badano, G., Baker, R., Calnan, M., Chalkidou, K., Culyer, A., Howdon, D., Hughes, D., Lomas, J., Max, C., McCabe, C., O'Mahony, J. F., Paulden, M., Pemberton-Whiteley, Z., Rid, A., . . . Wester, G. (2023). We need to talk about values: a proposed framework for the articulation of normative reasoning in health technology assessment. *Health Economics, Policy and Law*, 1-21. https://doi.org/10.1017/S1744133123000038

Ciani, O., Wilcher, B., Blankart, C. R., Hatz, M., Rupel, V. P., Erker, R. S., Varabyova, Y., & Taylor, R. S. (2015). Health technology assessment of medical devices: a survey of non-European union agencies. *International Journal of Technology Assessment in Health Care*, *31*(3), 154-165. https://doi.org/10.1017/S0266462315000185

Ducey, A., Ross, S., Pott, T., & Thompson, C. (2017). The moral economy of health technology assessment: an empirical qualitative study. *Evidence & Policy*, *13*(1), 7-27. https://doi.org/10.1332/174426415X14443053123024

Enzing, J. J., Knies, S., Boer, B., & Brouwer, W. B. F. (2021). Broadening the application of health technology assessment in the Netherlands: a worthwhile destination but not an easy ride? *Health Economics, Policy and Law*, *16*(4), 440-456. https://doi.org/10.1017/S1744133120000237

Enzing, J. J., Vijgen, S., Knies, S., Boer, B., & Brouwer, W. B. F. (2021). Do economic evaluations of TAVI deal with learning effects, innovation, and context dependency? A review. *Health Policy and Technology*, *10*(1), 111-119. https://doi.org/10.1016/j.hlpt.2020.09.006

Freitas, L., Vieira, A. C. L., Oliveira, M. D., Monteiro, H., & Bana, E. C. C. A. (2023). Which value aspects are relevant for the evaluation of medical devices? Exploring stakeholders' views through a Web-Delphi process. *BMC Health Services Research*, *23*(1), 593. https://doi.org/10.1186/s12913-023-09550-0

Fuchs, S., Olberg, B., Panteli, D., Perleth, M., & Busse, R. (2017). HTA of medical devices: Challenges and ideas for the future from a European perspective. *Health Policy*, *121*(3), 215-229. https://doi.org/10.1016/j.healthpol.2016.08.010

Gagnon, H., Legault, G. A., Bellemare, C. A., Parent, M., Dagenais, P., S, K. B., Tapin, D., Bernier, L., Beland, J. P., Daniel, C. E., & Patenaude, J. (2020). How does HTA addresses current social expectations? An international survey. *International Journal of Technology Assessment in Health Care*, *37*, e9. https://doi.org/10.1017/S0266462320000793

Haute Autorité de Santé (HAS). (2020). *Commission Nationale D'Évaluation des Dispositifs Medicaux et des Technologies de Santé. EDWARDS SAPIEN 3, bioprothèse valvulaire aortique implantée par voie transfémorale (système COMMANDER) Avis de la CNEDiMTS*.

Health Information and Quality Authority (HIQA). (2019). *Health Technology Assessment of transcatheter aortic valve implantation (TAVI) in patients with severe symptomatic aortic stenosis at low and intermediate risk of surgical complications*.

Himmels, J. P. W., Flottorp, S., Stoinska-Schneider, A., Kvist, B., & Robberstad, B. (2021). *Transcatheter aortic valve implantation (TAVI) versus surgical aortic valve replacement (SAVR) for patients with severe aortic stenosis and low surgical risk and across surgical risk groups: a health technology assessment. Report 2021*.

Hofmann, B., Cleemput, I., Bond, K., Krones, T., Droste, S., Sacchini, D., & Oortwijn, W. (2014). Revealing and acknowledging value judgments in health technology assessment. *International Journal of Technology Assessment in Health Care*, *30*(6), 579-586. https://doi.org/10.1017/S0266462314000671

Kinchin, I., Walshe, V., Normand, C., Coast, J., Elliott, R., Kroll, T., Kinghorn, P., Thompson, A., Viney, R., Currow, D., & O'Mahony, J. F. (2023). Expanding health technology assessment towards broader value: Ireland as a case study. *International Journal of Technology Assessment in Health Care*, *39*(1), e26. https://doi.org/10.1017/S0266462323000235

Klein, P., Blommestein, H., Al, M., Pongiglione, B., Torbica, A., & Groot, S. (2022). Real-world evidence in health technology assessment of high-risk medical devices: Fit for purpose? *Health Economics*, *31 Suppl 1*, 10-24. https://doi.org/10.1002/hec.4575

Loblova, O., Trayanov, T., Csanadi, M., & Ozieranski, P. (2020). The Emerging Social Science Literature on Health Technology Assessment: A Narrative Review. *Value in Health*, *23*(1), 3-9. https://doi.org/10.1016/j.jval.2019.07.016

Makady, A., Ham, R. T., de Boer, A., Hillege, H., Klungel, O., Goettsch, W., & GetReal Workpackage 1. (2017). Policies for Use of Real-World Data in Health Technology Assessment (HTA): A Comparative Study of Six HTA Agencies. *Value in Health*, *20*(4), 520-532. https://doi.org/10.1016/j.jval.2016.12.003

Ming, J., He, Y., Yang, Y., Hu, M., Zhao, X., Liu, J., Xie, Y., Wei, Y., & Chen, Y. (2022). Health technology assessment of medical devices: current landscape, challenges, and a way forward. *Cost Effectiveness and Resource Allocation*, *20*(1), 54. https://doi.org/10.1186/s12962-022-00389-6

O'Rourke, B., Oortwijn, W., Schuller, T., & International Joint Task, G. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, *36*(3), 187-190. https://doi.org/10.1017/S0266462320000215

Ontario Health. (2020a). Transcatheter Aortic Valve Implantation in Patients With Severe Aortic Valve Stenosis at Low Surgical Risk: A Health Technology Assessment. *Ontario health technology assessment series*, *20*(14), 1-148.

Ontario Health. (2020b). *Transcatheter aortic valve implantation in patients with severe aortic valve stenosis at low surgical risk: recommendation [Internet]*. https://www.hqontario.ca/evidence-to-improve-care/health-technology-assessment/reviews-and-recommendations/transcatheter-aortic-valve-implantation-in-patients-with-severe-aortic-valve-stenosis-and-low-surgical-risk

Oortwijn, W., Husereau, D., Abelson, J., Barasa, E., Bayani, D. D., Santos, V. C., Culyer, A., Facey, K., Grainger, D., Kieslich, K., Ollendorf, D., Pichon-Riviere, A., Sandman, L., Strammiello, V., & Teerawattananon, Y. (2022). Designing and Implementing Deliberative Processes for Health Technology Assessment: A Good Practices Report of a Joint HTAi/ISPOR Task Force. *International Journal of Technology Assessment in Health Care*, *38*(1), e37. https://doi.org/10.1017/S0266462322000198

Oortwijn, W., Jansen, M., & Baltussen, R. (2020). Use of Evidence-Informed Deliberative Processes by Health Technology Assessment Agencies Around the Globe. *International Journal of Health Policy and Management*, *9*(1), 27-33. https://doi.org/10.15171/ijhpm.2019.72

Pomey, M. P., Brouillard, P., Ganache, I., Lambert, L., Boothroyd, L., Collette, C., Bedard, S., Gregoire, A., Pelaez, S., Demers-Payette, O., Goetghebeur, M., de Guise, M., & Roy, D. (2020). Co-construction of health technology assessment recommendations with patients: An example with cardiac defibrillator replacement. *Health Expectations*, *23*(1), 182-192. https://doi.org/10.1111/hex.12989

Rumi, F., Fortunato, A., Antonini, D., Siviero, L., & Cicchetti, A. (2023). Analysis of heterogeneity of the different health technology assessment reports produced on the transcatheter aortic valve implantation in patients with severe aortic valve stenosis at low surgical risk. *Frontiers in Cardiovascular Medicine*, *10*, 1204520. https://doi.org/10.3389/fcvm.2023.1204520

Smith, A., & Argaez, C. (2019). *Transcatheter aortic valve implantation for aortic stenosis: a rapid qualitative review*.

Subbiah, V. (2023). The next generation of evidence-based medicine. *Nature Medicine*, *29*(1), 49-58. https://doi.org/10.1038/s41591-022-02160-z

Tarricone, R., Torbica, A., Drummond, M., & Medtec, H. T. A. P. G. (2017). Key Recommendations from the MedtecHTA Project. *Health Economics*, *26 Suppl 1*, 145-152. https://doi.org/10.1002/hec.3468

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*, *19*(6), 349-357. https://doi.org/https://doi.org/10.1093/intqhc/mzm042

Torbica, A., Tarricone, R., Schreyogg, J., & Drummond, M. (2022). Pushing the boundaries of evaluation, diffusion, and use of medical devices in Europe: Insights from the COMED project. *Health Economics*, *31 Suppl 1*, 1-9. https://doi.org/10.1002/hec.4600

Vahanian, A., Beyersdorf, F., Praz, F., Milojevic, M., Baldus, S., Bauersachs, J., Capodanno, D., Conradi, L., De Bonis, M., De Paulis, R., Delgado, V., Freemantle, N., Gilard, M., Haugaa, K. H., Jeppsson, A., Juni, P., Pierard, L., Prendergast, B. D., Sadaba, J. R., . . . Group, E. E. S. D. (2022). 2021 ESC/EACTS Guidelines for the management of valvular heart disease. *European Heart Journal*, *43*(7), 561-632. https://doi.org/10.1093/eurheartj/ehab395

van der Wilt, G. J., Bloemen, B., Grin, J., Gutierrez-Ibarluzea, I., Sampietro-Colom, L., Refolo, P., Sacchini, D., Hofmann, B., Sandman, L., & Oortwijn, W. (2022). Integrating Empirical Analysis and Normative Inquiry in Health Technology Assessment: The Values in Doing Assessments of Health Technologies Approach. *International Journal of Technology Assessment in Health Care*, *38*(1), e52. https://doi.org/10.1017/s0266462321001768

van der Wilt, G. J., Oortwijn, W., & Consortium, V.-H. (2022). Health technology assessment: A matter of facts and values. *International Journal of Technology Assessment in Health Care*, *38*(1), e53. https://doi.org/10.1017/S0266462322000101


5

Chapter 5

Different operationalizations of the capability approach in evaluating rehabilitation for persons with neuromuscular diseases. A mixed-methods study.

Bloemen, B., Pijpers, E.J., Groothuis, J.T., Cup, E.H.C., van Engelen, B.G.M., Oortwijn, W., & van der Wilt, G. J. (2024).
Submitted for publication.

ABSTRACT

Purpose: The capability approach (CA) is increasingly used in healthcare, but its operationalization to evaluate impact of interventions remains challenging. Therefore, we conducted a mixed-methods analysis to evaluate impact of rehabilitation on capability well-being of persons with facioscapulohumeral muscular dystrophy (FSHD) or myotonic dystrophy type 1 (DM1). We aimed to determine whether different CA operationalizations yield different results and to draw lessons about using the CA for evaluating impact of rehabilitation on capability well-being.

Methods: We compared semi-structed interviews with the ICEpop CAPability measure for Adults (ICECAP-A) and the Canadian Occupational Performance Measure (COPM) in evaluating changes in valuable functionings during rehabilitation. Semi-structured interviews were used to independently categorize participants into having worsened, unchanged, or improved valuable functionings. Electronic health records of participants were examined to confirm their experiences. Quantitative (descriptive statistics, Kruskal-Wallis tests) and qualitative comparisons between the interview-based categorization and changes in ICECAP-A and COPM scores were conducted to identify commonalities and differences.

Results: Although participants with improved valuable functionings also showed higher COPM scores at follow-up, ICECAP-A scores were often similar between baseline and follow-up. Particularly, changes related to work (paid or voluntary) and the need for participants to make choices between valuable functionings due to limited energy were not reflected by changes in ICECAP-A scores.

Conclusion: Only by combining information from quantitative outcome measures and interviews were we able to capture and understand changes in valuable functionings that occurred during rehabilitation for persons with NMD.

INTRODUCTION

The capability approach (CA) defines well-being as the *capability* to achieve *valuable functionings* in life, determined by (access to) resources, personal and environmental factors (Robeyns, 2005). It is increasingly used in healthcare but its operationalization to evaluate impact of interventions remains challenging (Mitchell et al., 2017; Rijke, Meerman, et al., 2023; Till et al., 2021). Challenges include specifying the valuable functionings that should be attainable for persons in a particular context, and establishing whether a person is *able* to be or do something when, in fact, that person is not displaying such beings or doings (Rijke, Meerman, et al., 2023; Ubels et al., 2022).

Measures, such as the ICEpop CAPability measure for Adults (ICECAP-A), are developed to evaluate impact of interventions on capabilities (Al-Janabi et al., 2012). However, critics argue that its fixed list of capabilities overlooks the elements of *choice* and *differences between people* regarding their values (both central features of the CA) (Lopez Barreda et al., 2019; Ubels et al., 2022; van Loon et al., 2018).

Within the Rehabilitation and Capability care for patients with Neuromuscular Diseases (ReCap-NMD) study, we conducted a mixed-methods analysis exploring different operationalizations of the CA in evaluating impact of rehabilitation for persons with facioscapulohumeral muscular dystrophy (FSHD) or myotonic dystrophy type 1 (DM1) (Bloemen et al., 2021; Pijpers et al., 2024). These slowly progressive neuromuscular diseases (NMD), characterized by muscle weakness and fatigue, involve a progressive loss of physical condition ranging from difficulty walking long distances to being unable to perform activities of daily living (e.g., walking, eating). Progression rate and severity vary significantly between persons. Since therapeutic options are limited, standard treatment consists of personalized rehabilitation to maintain or improve functioning (Ashizawa et al., 2018; Deenen et al., 2014; Mul, 2022; Tawil et al., 2015; van Engelen & The OPTIMISTIC Consortium, 2015). Therefore, to evaluate effects and inform clinical decisions, outcome measures are needed that capture rehabilitation impact on well-being (Voet et al., 2024). The CA, with its holistic focus on personal and environmental factors influencing functioning, as well as individual preferences, seems suitable for this purpose (van der Veen et al., 2023).

We compared semi-structured interviews with the ICECAP-A and a standard rehabilitation outcome measure (Canadian Occupational Performance Measure, COPM) to examine their ability to capture changes during rehabilitation in the ability of persons with NMD to realize valuable functionings. Because the COPM measures occupational performance, which is the ability to choose, organize, and satisfactorily

perform meaningful activities one wants, needs, or is expected to perform, it not only measures realization of rehabilitation goals but is also a proxy for measuring abilities in realizing valuable functionings.

METHODS

Study design and setting

A convergent parallel mixed-methods analysis was conducted within the ReCap-NMD study, which implemented and evaluated capability-based rehabilitation ('capability care') (Bloemen et al., 2021). The development of capability care is described elsewhere (Pijpers et al., 2024). Data from participants assigned to the intervention group (capability care) and of those assigned to the control group (care as usual at the Radboudumc Expertise center for neuromuscular diseases) were merged, since our research question was not about the impact of the intervention but about capturing changes in capability (irrespective of treatment allocation) while using different methods, see also Figure 1. Our hypothesis was that changes in capability would occur in both groups, potentially to a greater extent in participants receiving capability care.

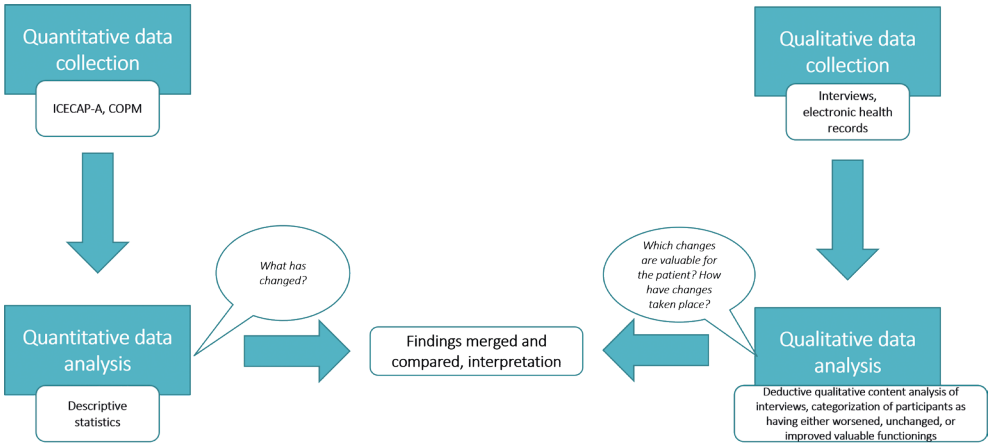


Figure 1. Illustration of the convergent parallel mixed-methods design. Quantitative (ICEpop CAPability measure for Adults, ICECAP-A; Canadian Occupational Performance Measure, COPM) and qualitative data (interviews, reports from participants' electronic health records) were collected 1-2 weeks before the visit of a participant to the department of Rehabilitation (ICECAP-A, COPM) and six months later (ICECAP-A, COPM, interviews, reports from electronic health records). Quantitative data was analyzed by descriptive statistics, qualitative data (interviews) through deductive qualitative content analysis. Results were merged and compared, and illustrative fragments from interviews were used to interpret results.

Participants

Persons diagnosed with FSHD or DM1, meeting pre-defined criteria, were included at the Radboudumc Expertise center for neuromuscular diseases in Nijmegen, The Netherlands, via either the department of Rehabilitation or Neurology. A first group of participants, recruited between November 2020 and July 2021, received multidisciplinary outpatient rehabilitation consisting of an 'analysis and advice' trajectory as usually provided. A second group of participants, recruited between March 2022 and November 2022, received multidisciplinary outpatient rehabilitation by professionals trained in applying the CA. For more details, including in- and exclusion criteria, see our published protocol (Bloemen et al., 2021).

Statistical power analyses was conducted for the primary outcome analysis of the ReCap-NMD study, estimating the power of the COPM to show statistical differences between usual care and capability care; see details in the protocol (Bloemen et al., 2021). Based on that analysis, we aimed to include at least 30 participants per intervention group.

Written informed consent was obtained from all participants before inclusion. The medical ethical reviewing committee CMO Regio Arnhem-Nijmegen granted full ethical approval (NL72794.091.20), and the study was registered at trialregister.nl (NL8946) on October 12, 2020.

Quantitative data collection

Quantitative outcome measures (COPM and ICECAP-A) were collected 1-2 weeks before a participant visited the outpatient clinic of the Rehabilitation department (baseline, T0) and six months later (follow-up, T1). The COPM was administered through semi-structured interviews performed by four independent occupational therapists during online or telephone appointments, and they entered the results manually into the Castor Electronic Data Capture (EDC) system (RRID:SCR_022150). The ICECAP-A was sent digitally using Castor EDC and completed online by participants themselves.

Canadian Occupational Performance Measure (COPM)

The COPM is an individualized instrument for assessing occupational performance in self-care, productivity and leisure (Law et al., 1990). Via semi-structured interviews, participants are asked to identify three to five priorities in their occupational performance and rate their current performance (COPM-P; 1= "not able to do it at all", 10= "able to do it extremely well") and satisfaction (COPM-S; 1= "not satisfied at all", 10= "extremely satisfied") with these occupations on an ordinal 10-point scale. Mean scores for both scales are calculated by dividing sum performance or

satisfaction scores by the number of identified priorities. To evaluate changes over time, participants rerate performance and satisfaction without seeing initial scores. Validity, reliability, and responsiveness of the COPM for evaluating rehabilitation has been demonstrated (Cup et al., 2003; Eyssen et al., 2005; Eyssen et al., 2011; Veenhuizen et al., 2019). We used the validated Dutch version (Eyssen et al., 2011; Van Duijn et al., 1999).

ICEpop CAPability measure for Adults (ICECAP-A)

The ICECAP-A measures five capabilities important to well-being: (1) *stability,* i.e., being able to feel settled and secure in all areas of life; (2) *attachment*, i.e., being able to have love, friendship, and support; (3) *autonomy*, i.e., being able to be independent; (4) *achievement*, i.e., being able to achieve and progress in all aspects of life; and (5) *enjoyment*, i.e., being able to have a lot of enjoyment and pleasure (Al-Janabi et al., 2012). It asks respondents to rate these abilities on a four-point scale ranging from no ability (level 1) to full ability (level 4), with higher scores indicating higher capability well-being. We used the validated Dutch version (Rohrbach et al., 2021; van Hoof et al., 2016).

Qualitative data collection

Interviews

Six months after visiting the Rehabilitation department, participants were invited via email for an interview, including details about its aim and the interviewer (name, position). Purposeful sampling was used to invite participants based on age, sex (male or female), diagnosis (FSHD or DM1), and time of inclusion (since the start of the study, to account for potential learning effects of healthcare professionals in delivering capability care), to achieve a representative distribution of interviewees. Participants receiving usual care were interviewed between May 2021 and September 2021; participants receiving capability care were interviewed between September 2022 and June 2023.

Semi-structured interviews were conducted by the first two authors (BB, EP) and a research assistant until saturation was achieved. The interviewers had a background in human movement sciences (EP), health technology assessment (BB), or occupational therapy (EP, research assistant). Interviewers had not met participants previously. Interviews were in Dutch, conducted from home using video-conferencing applications (Zaurus, Microsoft Teams), were either fully audio- or video-recorded, and took approximately one hour. Intelligent (non-verbatim) transcripts were subsequently created. Summaries of interviews were discussed among interviewers to evaluate whether new insights still arose from interviews or saturation was achieved.

Inspired by elements of the CA, we developed a semi-structured interview guide (see Supplementary 1) with three main questions: 1) *What are the valuable functionings of the participant?*; 2) *Have there been any changes in these valuable functionings in the past 6 months?*; and 3) *Can the changes be attributed to the received rehabilitation*? To define valuable functionings that should be attainable for participants, we supplemented the CA with a theory about *basic human goods* previously used in evaluating impact of interventions on capabilities, see Figure 2 (Alkire, 2002; Rijke, Vermeulen, et al., 2023). This theory identifies seven 'basic human goods' that guide human actions and represent their underlying value, see the interview guide in Supplementary 1 (Finnis, 1980).

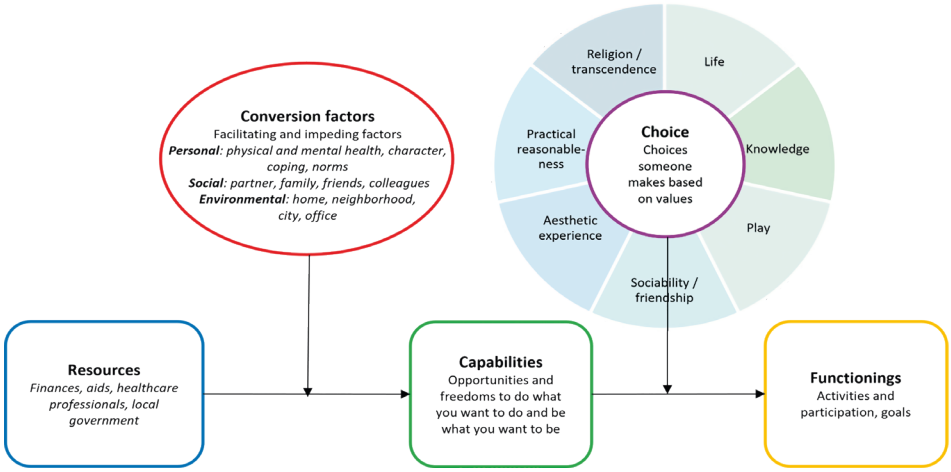


Figure 2. Schematic illustration of central terms from the capability approach supplemented with the theory of basic human goods. Among the *capabilities* that are available to a person, *functionings* are chosen by a person to pursue because they represent a basic human good, e.g., activities to maintain relationships are chosen because of the value of '*Sociability / friendship*'. These basic human goods (*Sociability / friendship, Play, Knowledge, Life, Religion / transcendence, Practical reasonableness, Aesthetic experience*) define the scope of functionings that should be attainable (figure adapted from Robeyns 2005).

Rather than asking participants how rehabilitation affected their capabilities, researchers prompted them to identify any changes in their daily lives since their visit to the department of Rehabilitation. Guided by the seven basic human goods, the interviewer aimed to identify all changes in *valuable functionings* and asked participants about the contribution of rehabilitation, *resources* (e.g., assistive devices), and *conversion factors* (e.g., personal traits, social support).

Electronic health records
To better understand reasons for reported changes, and confirm participant experiences, we examined participants' electronic health records to obtain information about recommendations participants received, and subsequent actions taken (e.g., received therapy, assistive devices etc., if reported).

Data analysis

Quantitative data
Clinical (diagnosis) and socio-demographic characteristics (age, sex) of participants were described. Descriptive statistics were generated for COPM-P, COPM-S, and ICECAP-A scores.

To perform statistical analysis and visualize results, the open-source R Project for Statistical Computing programming language (R version 4.1.3, R Core Team, 2022, RRID:SCR_001905) and RStudio user interface (version 2022.2.1.461, RStudio Team, 2022) were used.

Interview-based categorization of participants
Interviews were analyzed using a deductive qualitative content analysis approach (Mayring, 2000). We used codes derived from the CA, and the code *Experience with rehabilitation care* to code fragments in which a participant provides information about received rehabilitation and its impact. See also the full codebook (Supplementary 2).

Two researchers (BB, EP) independently coded interview transcripts applying the pre-defined codes, using ATLAS.ti version 23 for Windows (RRID:SCR_022920). Codes could overlap, as a fragment could contain information on multiple codes. Fragments identified by one researcher were coded by the other, and vice versa, to check consistency in using the codebook. The researchers discussed until consensus was reached on interpretation and application of the codebook.

Based on coded fragments, two researchers (BB and EP) independently categorized participants as follows:

- **Worsened valuable functionings**: a participant who experienced more problems in the achievement of at least one valuable functioning, without an equivalent substitutive functioning (i.e., belonging to the same category of basic human goods).
- **Unchanged valuable functionings**: a participant who experienced no changes in the achievement of valuable functionings.

- **Improved valuable functionings**: a participant who experienced less problems in the achievement of at least one valuable functioning or realized an equivalent substitutive functioning, or realized a new functioning, while maintaining other valuable functionings.

If a participant decided to stop pursuing a functioning due to restrictions imposed by the disease (e.g., limited energy, physical impairments) this was classified as a worsened functioning irrespective of whether the participant was satisfied with current performance. Researchers (BB and EP) categorized participants independently, and discrepancies were resolved by discussion to obtain consensus.

Comparison of qualitative and quantitative findings

To evaluate agreement, we compared change scores (T1-T0) for the ICECAP-A, COPM-P and COPM-S between the different interview-based groups. Non-parametric Kruskal-Wallis tests were performed to test for significant differences between change scores of the groups, with a significance level of 0.05.

To facilitate comparisons, information on quantitative and qualitative findings, experiences with rehabilitation, and the value of changes for participants were combined in a single table. Differences and commonalities between the interview-based categorization and changes in COPM and ICECAP-A scores were examined, and illustrative interview fragments identified to explain findings. These fragments were translated from English to Dutch by the first author (BB), using DeepL (www.deepl.com), and all personally identifiable information was removed to ensure participants' anonymity.

The consolidated criteria for reporting qualitative research (COREQ) checklist was used to ensure that methods, results, and discussion were reported appropriately (Tong et al., 2007).

RESULTS

Study participants

Out of 101 invited persons for the ReCap-NMD study, 64 participated (response rate 63%). Foremost reasons for declining participation were: timing of rehabilitation appointments did not allow completing measurements in time (n = 10), too time / energy consuming (n = 8), inclusion already completed for their type of NMD (FSHD or DM1) (n = 6), and having no current rehabilitation aims (n = 5). Table 1 displays characteristics of the 26 participants included in current analysis that completed COPM and ICECAP-A instruments as well as the interview. Whereas average

COPM scores were higher at follow-up, average ICECAP-A sum scores were similar at baseline and follow-up.

Table 1. Characteristics of all participants included in current analysis.

Total N		26
Diagnosis	FSHD	16 (62%)
	DM1	10 (38%)
Sex	Male	10 (38%)
	Female	16 (62%)
Age (years)	Mean (Range)	46 (28-68)
ICECAP-A sum score	Mean baseline (SD)	14.8 (2.4)
	Mean follow-up (SD)	15.0 (2.5)
COPM-P	Mean baseline (SD)	5.2 (1.6)
	Mean follow-up (SD)	6.2 (1.2)
COPM-S	Mean baseline (SD)	4.6 (1.6)
	Mean follow-up (SD)	5.9 (1.5)
Follow-up time (months between T1 and T0)	Mean (range)	6.3 (5.6, 9.1)

Abbreviations: FSHD: facioscapulohumeral dystrophy; DM1: myotonic dystrophy type 1; ICECAP-A: ICEpop CAPability measure for Adults; COPM-P: Canadian Occupational Performance Measure Performance score; COPM-S: Canadian Occupational Performance Measure Satisfaction score.

Interview-based categorization of participants

Based on the interviews, participants were categorized as having either worsened (n=7), unchanged (n=7), or improved (n=12) valuable functionings, see Table 2. For six participants (23%) there was disagreement between researchers (BB, EP) on the categorization, which was resolved after one round of discussion. As can be seen in Table 2, the participants with improved valuable functionings had higher COPM scores at follow-up, whereas the ICECAP-A scores did not change on average.

Table 2. Characteristics of participants categorized in terms of changes in valuable functionings (based on interviews).

	Worsened valuable functionings	Unchanged valuable functionings	Improved valuable functionings
Number of participants (%)	7 (27%)	7 (27%)	12 (46%)
Age (years), mean (range)	47 (35-59)	48 (29-56)	44 (28-68)
Male, n (% of total)	1 (10%)	5 (50%)	4 (40%)
Female, n (% of total)	6 (38%)	2 (12%)	8 (50%)
FSHD, n (% of total)	3 (19%)	5 (31%)	8 (50%)
MD1, n (% of total)	4 (40%)	2 (20%)	4 (40%)
COPM-P baseline, mean (SD)	5.10 (1.92)	5.03 (0.74)	5.29 (1.94)
COPM-P follow-up, mean (SD)	5.37 (1.14)	6.03 (0.63)	6.70 (1.31)
COPM-S baseline, mean (SD)	4.25 (2.01)	5.14 (1.05)	4.41 (1.65)
COPM-S follow-up, mean (SD)	4.94 (2.04)	5.89 (0.79)	6.45 (1.28)
ICECAP-A level sum score baseline, mean (SD)	13.71 (1.80)	14.86 (2.79)	15.50 (2.32)
ICECAP-A level sum score follow-up, mean (SD)	13.71 (1.89)	15.14 (2.79)	15.58 (2.50)

Abbreviations: FHSD: Facioscapulohumeral muscular dystrophy; MD1: myotonic dystrophy type 1; COPM-P: Canadian Occupational Performance Measure Performance score; COPM-S: Canadian Occupational Performance Measure Satisfaction score; ICECAP-A: ICEpop CAPability measure for Adults.

Comparison of qualitative and quantitative findings

Quantitative differences of COPM and ICECAP-A change scores between the groups

The ICECAP-A, COPM-P and COPM-S change scores were plotted for the different groups, see Figure 3. Only changes in COPM-P and COPM-S scores varied among the groups. Kruskal-Wallis tests indicated significant differences in COPM-S scores between the different groups, $\chi 2(2) = 7.90$, $p = .020$.

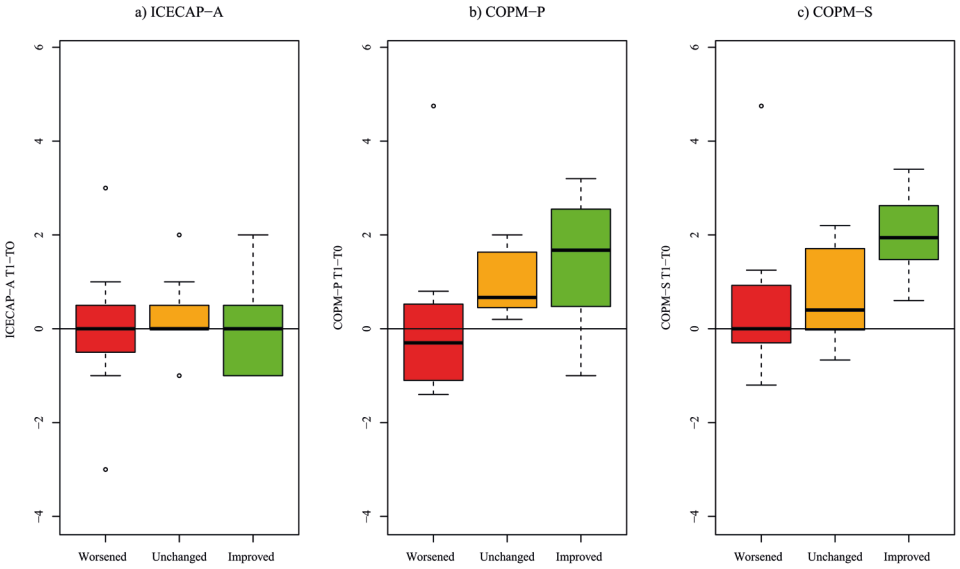


Figure 3. Distribution of change scores (T1-T0) among the different groups of participants (categorized as having *worsened*, *unchanged* or *improved* valuable functionings) for the a) ICEpop CAPability measure for Adults (ICECAP-A) level sum scores, b) Canadian Occupational Performance Measure performance (COPM-P) scores, and c) Canadian Occupational Performance Measure satisfaction (COPM-S) scores.

Comparing interview-based categorization and changes in ICECAP-A scores

Of the seven participants with worsened valuable functionings, only two (29%) had lower sum ICECAP-A scores at T1; among the seven participants with unchanged valuable functionings, four (57%) had also unchanged ICECAP-A scores; and only three participants (25%) with improved valuable functionings had improved ICECAP-A sum scores (see Table 3).

Thirteen participants (50%) mentioned changes in their ability to fulfill paid or voluntary work, which was not always reflected by changes in ICECAP-A scores (Table 3). One participant (#51) even improved on *attachment, achievement,* and *enjoyment*, while being enforced to take a significant step back at work:

> "*I have switched back to my old role as a teacher. That is doable. But sometimes it is also quite tough.* […] *My current position as a teacher is not satisfying.* […] *In my previous position I had much more authority*" (Participant #51)

Participants described the value of work as being able to contribute to society, and a source of meaning and self-esteem (see illustrative interview fragments in Table 4). These valuable aspects of work are not (explicitly) covered by ICECAP-A dimensions.

Twelve participants (46%) struggled with fatigue and/or the progressive nature of NMD, forcing them to stop valuable activities, do them differently or less, or rethink goals, to save energy or anticipate future problems (Table 3). They need energy management strategies to stay vital and fulfill other important activities, enforcing them to make difficult trade-offs between what is (still) valuable to them. Or, in setting priorities, they sometimes anticipated a worsening of their condition (see Table 4). This leads to differences between experienced changes in valuable functionings and (level sum) ICECAP-A scores. For example, one participant (#13) decided to stop running because she anticipated risks of falling due to her worsening physical condition (stability +1, enjoyment – 1):

> *"I stopped running because the risk of stumbling became too high. I thought it is just not smart. Having a brace to walk and simultaneously trying to continue running. That's too dangerous.* […] *Well, it wasn't an advice I received from the rehabilitation team, but they did hint at the risk of falling. And I stumble a lot when I walk, so you can draw that conclusion yourself that it's not so smart.*" (Participant #13)

Another participant (#57), although still able to perform valuable functionings, already made adjustments to anticipate future problems (reflected by lower COPM-P and COPM-S scores, ICECAP-A scores unchanged, see Table 3):

> *"Since December I work 4 times 6 hours, first it was 4 times 7 and I think there are a lot of colleagues, not in the same department, who didn't even know that I work less, because I just did my thing and that worked out fine. But I do feel like I'm kind of at a tipping point now. It does get harder to schedule meetings and things like that. If it were to become even less hours, I think it does become more difficult to continue to fulfill my current position. "* (Participant #57)

One participant (#5) showed no changes in ICECAP-A scores despite improvements in valuable functionings. In the interview she explained that she adapted her daily schedule to save energy for reading and walking (see interview fragment, Table 4). These functionings are valuable to her as a form of leisure and to work on her physical fitness. These changes may be too small (*enjoyment* was already at level 3) or represent a value not reflected by ICECAP-A dimensions.

For some participants, (physical) abilities did not improve but they still realized valuable functionings due to enhanced opportunities. Two participants (#8, #37) were supported by the rehabilitation team in coping with job loss, and opportunities were provided (e.g., help with applying for disability benefits, energy management strate-

gies) to enhance their capabilities in finding valuable alternatives. Another participant (#16) regained the ability to cycle and walk by receiving an ankle foot orthosis. These improvements were not reflected by higher ICECAP-A scores.

Comparing interview-based categorization and changes in COPM sores

Of the seven participants with worsened valuable functionings, three (43%) had lower COPM-S scores and four (57%) had lower COPM-P at T1; none of the seven participants with unchanged valuable functionings had also unchanged COPM scores; among participants with improved valuable functionings, twelve participants (100%) had improved COPM-S scores and ten participants (83%) had improved COPM-P scores (see Table 3).

Three participants, despite having worsened valuable functionings, improved on the COPM-S (see Table 3). One participant (#40) explained this was due to acceptance of reduced performance:

> *"For me it's Monday and Wednesday sports and the other three days I work. And to play sport for a third time in the week is just not possible, because I also must do housekeeping and run some errands. And, if possible, the occasional social contacts. And I try to find my way in this. Does everything work out as I would like it? No. But I learned to let it go".* (Participant #40)

Two participants (#17, #38) showed significant improvements on COPM scales but the identified problems in occupations were daily life activities that did not represent valuable functionings (i.e., they were not mentioned as such during interviews). Another participant (#47) improved on COPM scales due to better performance in the occupations 'energy management' and 'having a conversation / talking' but mentioned no valuable functionings that became easier to perform because of this, and attributed improvements to increased awareness of speaking speed.

Table 3. Findings of quantitative and qualitative analyses, grouped for participants categorized (based on interviews) as having worsened, unchanged, or improved valuable functionings. Shown are, per participant, the problems in occupations prioritized by the participant when conducting the COPM ('COPM occupations'); changes in COPM-P (T1-T0) and COPM-S (T1-T0) scores; the identified changed valuable functionings in the interview; the value of the respective valuable functionings (the *basic human goods* that they represent), i.e., 'no explicit information' means that the underlying value was not explicitly mentioned by the participant); information about the received rehabilitation care (based on information from the interview and the participants' electronic health record).

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
Participants with worsened valuable functionings							
#4	1. Walking (to be able to engage in activities with family, day trips, social activities, work) 2. To be able to stand for a longer period of time (to engage in social activities) 3. Working more / sustain working activities 4. Percussion (be able to drum again) 5. Cooking	-1.4	-0.2	*Stability:* -1 *Attachment:* 0 *Autonomy*: -1 *Achievement:* 0 *Enjoyment*: -1	Reduced ability to work (is in process of working less) Percussion band (stopped due to lack of energy) Yoga (stopped doing it) Activities with friend (less due to lack of energy) Vacation with family (experiences increased limitations and barriers)	Knowledge, transcendence Play, friendship No explicit information No explicit information No explicit information	Referral to physical therapist and speech therapist; slightly improved physical condition due to advice to do other types of exercises with physical therapist.
#13	1. Activity to get out of bed in the afternoon (energy) 2. Improve / maintain grip strength 3. Walking at home / Running outside (reduce stumbling) 4. Be able to get up 5. Use the computer	-0.3	+0.6	*Stability:* +1 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: -1	Running (discontinued because of fear of stumbling) Gardening (not able to sustain activities as long as wished for)	Life, play Aesthetic experience	A brace will be fitted; referral to physical therapist (at own request, wishes to have a therapist with more knowledge about FSHD) – started exercises for shoulders; discussed risk of stumbling with physical therapist (based on that, participant seemed to have decided to stop with running); discussed fatigue with occupational therapist.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#40	1. Relaxing activities / time to do nothing 2. Housekeeping (e.g., cleaning bathroom, cleaning windows) 3. Work: assist clients in activities of daily living 4. Social activities	+0.25	+1.25	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* -1 *Enjoyment*: 0	Housekeeping (would like to do more, but increasingly choses to do less to save energy)	No explicit information	Received advice to use the brace more often, this led to reduced knee pain; Changed diet, lost weight; received help in contacting the occupational health physician at work; received advice to resume daily rhythm / structure, this led to more energy and inner peace.
					Work (moved to a different team after a consult with the occupational health physician, this led to more work satisfaction)	Play, transcendence, friendship	
					Water aerobics (more different types of exercises)	Life, play	
					Mentorship of a family member (forced to quit)	No explicit information	
#43	1. Activities with family (e.g., playing games, baking cookies) 2. Housekeeping (organising) 3. Work (de-escalating: lifting children) 4. Playing flute 5. Running	+0.8	0	*Stability:* +1 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Teacher (switched to a group with younger children to physically cope with all responsibilities)	Transcendence, knowledge	Started running again with help of physical therapist (and support of partner), experienced positive impact of on physical and mental health.
					Running (resumed training)	Life, play	

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#51	1. Receive support in making important life choices (related to work, finances, pregnancy, self-care) 2. Wish to have children 3. Avoid bad work pattern (to many hours and responsibilities, sitting for long periods of time, eat less) 4. Physical exercise for shoulders and neck (to reduce pain and stiffness, improve self-care)	+4.75	+4.75	*Stability:* 0 *Attachment:* +1 *Autonomy*: 0 *Achievement:*+1 *Enjoyment*: +1	Teacher (stepped back from a management role to save energy, but this reduced work satisfaction) Relationship with partner (improved, living together again) Activities with friend (has to decline often due to lack of energy)	Knowledge, transcendence Friendship Friendship, play, life	Referral to specialized physical therapist to treat neck and shoulder problems; due to limited time / travel distance the participant decided to go to another therapist, this had a negative impact (to heavy exercises); talked with occupational therapist about priorities at work, this contributed to decision to take a step back but this reduced work satisfaction.
#57	1. Walking 2. Work (energy balance) 3. Getting on and off the electric bicycle 4. Going on vacation / abroad 5. Doing chores around the house	-1.4	-0.4	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Work as project manager (reduced working hours due to limited energy, difficulties with participating in social activities) Walking with the dog (more difficulties due to risk of falling / stumbling) Vacation, travelling (more difficulties, forced to go to less hilly areas in Italy due to problems with walking)	Knowledge, play, transcendence No explicit information Play, aesthetic experience	Referral to occupational therapist, helped with funding application for a wheelchair; Physical therapy: exercises, waistband, this did not help in reducing back pain, temporarily stopped therapy.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#61	1. Getting up from a (lower) chair / bench / bed 2. Energy management / control fatigue 3. Take a shower (standing upright, getting in and out, washing hair) 4. Standing without help (problems with stability) 5. Walking	-0.8	-1.2	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Reduced physical condition (difficulties with walking, getting up from the bed, taking a shower, swimming exercises) Would like to do more activities with daughter (experiences more difficulties)	No explicit information No explicit information	Referral to occupational therapist; support in funding application for a walk-in shower; appoint made for fitting an aid to get out of bed.
Participants with unchanged valuable functionings							
#6	1. Evening activities with partner 2. Social activities with friends 3. Playing soccer with son 4. Standing and walking for more than one hour	+0.25	-0.25	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: -1	-	-	Recommendations by occupational therapist helped participant to make choices in terms of energy management, takes more breaks (and delegates tasks to partner); on waiting list for the B-fit program; ankle foot orthosis fitted, experienced no effect yet but participant would like to wait to see what happens.
#10	1. Time management (autonomy) / Informal care (reduce efforts, take a break) 2. Walking up and down the stairs 3. Taking the stairs with a laundry basket 4. Going to the cinema with a friend 5. Gardening	+0.65	+0.4	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: +1	-	-	Neurologist referred participant to ophthalmologist to check blepharitis and irritated eyes, this led to less complaints; reduced chest pain due to changes in therapy received by physical therapist (based on advice by Radboudumc), this improved mood.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#12	1. To sit comfortable at desk 2. Walk for more than an hour on hilly terrain 3. Going on vacation (use airplane, public transport)	+0.67	-0.67	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	-	-	Video consult with occupational therapist about ergonomics related to working at home.
#17	1. Improve balance while getting up from a chair 2. Improve balance during walking 3. Improve hand function (writing) 4. Take initiative for daily activities	+2	+1.75	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	-	-	Home visit by occupational therapist, talked about daily structure, and received advice related to daily life activities; hand therapy.
#38	1. Getting on and off the electric bicycle and stopping at a traffic light 2. Walking downstairs with a full basket of laundry 3. Getting up from the ground (after falling or squatting or kneeling) 4. Stirring in a pan 5. Walking in a forest	+1.6	+2.2	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	-	-	Advice given to local physical therapist to start shoulder exercises, but this had a negative impact and was therefore stopped; advice to wear a helmet when cycling; Advice to be mindful while swallowing; Worries about progression of the disease were resolved because tests showed no changes in muscle strength.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#47	1. Walking (on the farm, to the car, together with partner) 2. Energy management (being able to go out for a dinner, sport, parties) 3. Having a conversation / talking	+1.67	+1.67	*Stability:* +1 *Attachment:* 0 *Autonomy*: 0 *Achievement:*+1 *Enjoyment*: 0	-	-	Ankle foot orthosis fitted, no success (not suitable for working environment); Ignored advice to see a physical therapist; participant decided to decline help with energy management; less problems with talking due to slowing down speaking speed / being more conscious.
#64	1. Articulation (speaking at work where language is English, as the day progresses speaking takes more energy and others have difficulty understanding him) 2. Self-care (e.g., cutting toenails) 3. Opening bottle caps with tool 4. Typing (hitting the right keys) 5. Driving a car (driving in the dark, estimating depth and distance to other cars)	+0.2	+0.2	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	-	-	Ignored received advice to have a check to receive a fitness to drive a car certificate, and to fit an orthosis or use aids); local rehabilitation center concluded that only a yearly check is needed.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
Participants with improved valuable functionings							
#5	1. Work: reading and responding to mail, reviewing literature, writing 2. Day out with partner: visiting a museum, other cities, concert 3. Maintaining social contacts 4. Have a walk of 45 min or longer 5. Self-care / organizing care	+2.6	+1.8	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Voluntary work as dietician (found other moments during the day to do some work) Reading (started doing it) Walking (started together with social worker) Day trips (less able to do this due to COVID-19 measures)	No explicit information Life, play Life Life	Fibrapubic catheter instead of catherisation 5x a day, coughing machine – both as a result of referral / advice by rehabilitation team, this makes daily activities more easy (saves energy); rehabilitation helped in making choices with respect to work (delegating tasks) and taking more time to relax; referral to psychologist; based on advice by rehabilitation team the participant searched for a social worker, this helped.
#7	1. Improve hand function to perform activities of daily living 2. Learning an adequate gait pattern to be able to walk longer distances 3. Creating a balanced daily schedule (to return to work) 4. Improving work pace in performing activities of daily living	+1.75	+1.75	*Stability:* -1 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Walking (more easy and more often)	No explicit information	Referral to physical therapist, exercises helped to be able to walk for a longer period of time and without pain; ignored advice received by occupational therapist.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#8	1. Being able to talk about illness with colleagues, standing up for oneself 2. Find a balanced daily rhythm 3. Improve physical condition / fitness 4. Attitude related to work place 5. Hobby: repair vintage cars	+3.2	+3.4	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Search for alternative daytime activities after quitting work (has more time after fully quitting job)	No explicit information	Received counselling by occupational therapist on work and daily schedule, this led to initiation of process to quit work and receive disability benefits.
#9	1. Sailing 2. Walking with the dog in a forest (5 km) 3. Painting the boat (doing chores) 4. Brush the dog	+2.5	+2.08	*Stability:* 0 *Attachment:* 0 *Autonomy*: +1 *Achievement:* 0 *Enjoyment*: +1	Houseboat maintenance, making things for the boat with a sewing machine (was able to do more due to improved energy levels)	No explicit information	Referral to a specialized physical therapist, now has two therapists – one that treats shoulder trigger points, and one for stabilizing shoulders, experienced improvements (more energy, better physical condition and ability to move; improved ability to work at boat); referral to occupational therapist to discuss burn-out (no positive experience); also received advice to see a psychologist, but participant has no interest.
#11	1. Walking > 45 minutes (weekly) 2. Voluntary work at the library and patient organization 3. Reading (sitting) > 30 minutes 4. Skiing (on vacation); 5. Wiping buttocks	+0.35	+1.55	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:*+1 *Enjoyment*: 0	Mountain biking (more frequently and less fatigue due to using an electrical mountainbike) Reading aloud (started as volunteer at elementary school)	Play, friendship; Transcendence, play, friendship	Lung test and sleep test done to exclude sleep apnea etc., tips for sleep position already applied by patient; Dynamic ankle foot orthosis fitted, patient still has to try out whether it helps with walking and korfball; Discussion with occupational therapist about need for paid job vs more time for relaxation, alternative voluntary work, this has played a role in how the participant thinks about these issues and making decisions.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#16	1. Walking with partner and being able to walk longer distances on vacation 2. Cooking (improve grip strength) 3. Reduce back pain to be able to perform housekeeping activities	+1.0	+3.0	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Cycling with partner (resumed activities) Walking with partner (resumed activities with the help of adaptations, Nordic walking sticks, ankle foot orthosis)	Life, friendship; Life, friendship;	Fitted ankle foot orthosis, tips for physical therapist – this resolved ankle complaints, able to walk again (although with some difficulties).
#35	1. Pass a shooting test (work as police officer) 2. Activities with children 3. Taking initiative for activities 4. Talking (with family members, others; to present)	+1.75	+3.25	*Stability:* 0 *Attachment:* 0 *Autonomy*: -1 *Achievement:* 0 *Enjoyment*: 0	Work as police officer (passed shooting test, improved physical condition)	Play, transcendence, friendship, Aesthetic experience, Knowledge	Impact on lifestyle (improved frequency of sport activities together with physical therapist, nutritional advice) and physical condition; ankle foot orthosis fitted, helped to sport more outside (running); hand exercise, helped to pass shooting test at work; participant decided to look for a speech therapist.
#36	1. Holding daughter 2. Keeping balance during longer stretches of walking and while taking the stairs 3. Eating an apple 4. Getting up from a lying position 5. Fitness exercises	+0.6	+0.6	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Walking (on vacation, still able but with more stability problems) In vitro fertilization procedure started (to avoid having a child with FSHD) Work (switched to another client, this improved work satisfaction)	Play, friendship Life Life, play	Physical therapist advised to do more aerobic exercises, experienced no effects.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#37	1. Taking a shower 2. Time for children and partner 3. Walking 1 km 4. Visiting or receiving friends 5. Doing chores around the house	-0.2	+1.4	*Stability:* 0 *Attachment:* 0 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: -1	Started voluntary work at community association Wishes to contribute to FSHD research and the patient association	Frienship, play, transcendence; Knowledge, transcendence, friendship	Started a new rehabilitation program focused on relaxation, this helped to maintain relaxing activities for a longer amount of time; received advice about daily schedule / taking time for breaks; conversations with social workers about dealing with adolescent daughters and how they cope with FSHD, and about how to cope with job loss (and finding alternatives).
#46	1. Assume an interested posture (facial expressions) 2. Running away from danger (traffic, event) 3. Participate in (sport) activities at work 4. Washing hair 5. Walk 10 km (staying fit)	+1.6	2.2	*Stability:* 0 *Attachment:* 0 *Autonomy*: -1 *Achievement:* 0 *Enjoyment*: 0	Work (started on a new job, more work satisfaction) Rowing (started training for a marathon)	Knowledge, play, friendship, life; Play, aesthetic experience, life, friendship	Discussed how to deal with employer / colleagues and telling / not telling about illness, sometimes unable to participate in activities at work; Referral to psychologist to compensate for reduced facial expressions was not continued; received advice about exercises to conduct with own physical therapist, did not experience an effect.
#49	1. Improve physical condition (reduce pain during walking) 2. Improve energy levels (vacation, work for own company) 3. Improve energy balance to have more time for partner 4. Improve energy levels for self-care	+2.75	+2.25	*Stability:* +1 *Attachment:* 0 *Autonomy*: 0 *Achievement:* -1 *Enjoyment*: 0	Work as freelancer (was fired, decided to start working as freelancer) In vitro fertilization procedure (started)	Life, knowledge Life	During the consult with the rehabilitation physician it was discussed what type of FSHD the participant has (lack of clarity) and associated heredity; received information about gene therapy, received advice to do (more) aerobic exercises.

Table 3. *Continued.*

Participant	COPM occupations	COPM-P change	COPM-S change	ICECAP-A change	Changed valuable functionings (What has changed?)	Choice (why is it valuable?)	Received rehabilitation care
#52	1. Sitting straight (working at a desk for more than one hour) 2. Worries /Negative mindset / Uncertainty about future (receive support) 3. Smiling (facial expressions) 4. Improve bladder function	-1	+1.25	*Stability:* +1 *Attachment:*+1 *Autonomy*: 0 *Achievement:* 0 *Enjoyment*: 0	Meditation (started together with friend, helps to relax)	Life, friendship;	Received advice on physical therapy; information about triple chair; received advice to actively use facial muscles to improve speech and facial expressions, actively pay attention to closing the mouth at rest.
					After receiving diagnosis (1.5 years ago) started to reflect upon what matters, decided to make a journey with family and relativized the importance of work	No explicit information	
					Thinks about the possibility of another job in the future ('limited time' due to progressive disease, focus on what is important and satisfying)	No explicit information	
					Helped with project at school of daughter (completed, was satisfying)	Transcendence, friendship	

Table 4. Fragments from interviews that illustrate commonalities and differences between the interview-based categorization of participants and changes in ICECAP-A scores.

Finding	Illustrative fragments from interviews
The importance of (voluntary or paid) work for being able to be an active member of society (meaning, self-esteem), not always reflected by changes in ICECAP-A scores	*"Yes, I am trying. I thought, I'm not going to give up if I don't have to. Yes of course, I am really sad sometimes, especially about my work. You feel less part of it. Sometimes I also wonder whether colleagues think so too, and I really try to express that and test whether they think so. And then I do get to hear, I assume that they are honest, that people say that I do have enough qualities that make me very valuable for my job. And I try to avoid being sad, although I am sometimes, but I always try to make something of it."* (Participant #4) (In process of working less due to reduced abilities; *Stability -1; Autonomy -1; Enjoyment -1*) *"The meaning of work? Yes, that is important, it means social contacts and being busy. Clients like it that you're there, and I always say that if they had a nice day I had a nice day too. I just think it's important, just having a nice day together. Yes, that's the goal for me. I had, I think in July or August, my annual interview with the team leader at work and he asks, 'What ambitions do you have left'? I said mainly just to enjoy working. I don't need another position, not a higher position with more responsibilities, I'm already happy if I survive. [...] Yes, I did want to go back to work, but then I wanted to do the work I was trained for. And once I went back to work as a flex worker, I noticed that I got some appreciation from colleagues. Only then did I feel that I didn't really have that anywhere else. I didn't have that at home and I didn't have that with my ex. And, consequently, I only started growing when I started working […] I have now permanent groups at work. For the past two years I felt more like a flex worker than that I was still really involved in something. That was not positive because you are always on a different group which makes you have less grip on your clients. And you also don't have the feeling that you're helping to build a team, and I missed that. I do have that again now. This is something that resulted from that."* (Participant #40) (Moved to a different team after consult with the occupational health physician, this led to more work satisfaction; *Achievement -1*) *"Now there is plenty of work in pedagogy itself, but when I was... I was always busy with applying for jobs. But that never really paid off. [...] It was my dream to be able to continue working in teaching. And to perform at a certain level. Also, eventually, to get paid. I don't think the money is the most important thing but, still, if you are working as a teacher and you don't get paid accordingly, so to speak. I mean, there is just a very big difference between a teaching assistant and a teacher. […] I now work again just 3 days and I now have a group with children that are less strong, and I can handle them on my own."* (Participant #43) (Switched to teaching a group with younger children to physically cope with all responsibilities; *Stability +1*) *"I think it is important for myself that I remain active in a socially responsible way, and that I try to postpone as long as possible the situation that I stay at home and are not active anymore. My goal is to reach the age of retirement while actively working. I hope to be able to do that."* (Participant #57) (Reduced working hours due to limited energy, difficulties with participating in social activities at work; no changes in ICECAP-A scores) *"Because I am a police officer I have much and longstanding contact with people, it gives me the opportunity to help people with the problems they have. This gives me some gratitude and honor" [..] I talked to the physical therapist about my hands, because one of the important things in my job is a shooting test and I noticed that it was more difficult for me to move the trigger of my gun. So, I was recommended some additional exercises for improving that, and informed the physical therapist that my last shooting test had gone well and that was the result of his recommendations."* (Participant #35) (Passed shooting test at work and improved physical condition; *Autonomy -1*) *"I really like my job and I really want the things that I've learned I want to, you know, if I can help other people with that, then well, that makes me really happy too. So, I took another four-day training last week to learn more things then hopefully I can apply that when working for my client."* (Participant #46) (Started on a new job, more work satisfaction; *Autonomy -1*)

Table 4. *Continued.*

Finding	Illustrative fragments from interviews
The necessity of making trade-offs due to reduced energy levels or uncertainty about future progression of the disease, not always reflected by (level sum) ICECAP-A scores	*"A friend of mine asked me the other day about my hobbies. Because it seemed to my friend like I am only busy with things that need to be done. And then I thought 'oh yes'. Well, I like cooking, I still do. And then the other day I thought maybe I could see if I could make some kind of music. Because I really like making music. But with my work being so busy I don't feel the energy for that yet. But yes, it can also be energy-giving. [...] Well, what I can't do anymore is to go for dinner with someone spontaneously. At least not in a crowded restaurant. Because people don't hear or understand me anymore. I have to make too much effort to talk. [...] Yes, and the same goes is the case at work, those meetings. Then I just talk less. And that is not who I am, because I like to take an active role in discussions. But yes, I have to let it go. "* (Participant #4) (*Stability -1; Autonomy -1; Enjoyment -1*) *"For example, when we have a meeting on Friday afternoon to discuss all the classes, I think I can add less than I could, or I can be less professional than I am. [...] Yes, then my energy just runs out. And then I better, [laughs], focus on the things that have the highest urgency so that the other teacher can continue on Monday. Yes, then you can not expect much more from me. But yes, I find that difficult, because you are responsible for a class. "* (Participant #43) (*Stability +1*) *"Yeah, I think I just allow myself a little more rest and space, so there's less pain. That I'm a little bit kinder to myself. If something doesn't work out I think now that it doesn't always have to be perfect, so then I suffer less. I make sure that I have a little bit more moments of rest and that ensures that I have less pain, so to speak. That I don't have an overwhelming pain. [...] Yes, then I just have to make other choices with work and so on. Yes, I often discuss that with the occupational therapist, but I'm not really ready for that yet. Yes, that is still a bit. The struggle. [...] My occupational therapist thinks it's stupid that I still work full-time. [Laughs] But yes, I've already made choices in the sense of… […] I have now switched to my old role as a teacher. That is doable. But sometimes it is also quite tough. […] Yes, I think it would be best to work four half days or so. I think that would be best for my health. That I have time to meet my physical therapist and manage all the appointments for my health. Try to stabilize what is already, what, (...), say already deteriorating but still try to stabilize. […] But yes, I have vacation in 5 weeks. Or sabbatical. And then I want to get some things sorted out. And I want to talk with my occupational therapist again and the physical therapist. About how to get things started. Because I do find that. Yes, it is also the case that you are kind of left to your own devices in that respect, like, yes, this is your condition and uhm, you cannot get better so uhm, yes, you will have to make do with it. [...] I mean, I've been for years, say, super tired. I just thought at first that it was just normal that I just couldn't do anything at night or on weekends. Now I think that maybe, [laughs], it is a result from my disease. But it is confronting to think about that. Because you actually want that everything is good."* (Participant #51) (*Attachment +1; Achievement +1; Enjoyment +1*) *"I am a project manager. That is mainly office work and meetings, and in recent years luckily much more can be done online and at home. I do feel blessed to have that position. I have reduced my working hours over the years, but I still work 4 times 6 hours now per week. And well, the amount of work I can also regulate, so I'm happy with that work and also with my employer. Yes, that fits, that fits very well. [...] So, since December I now work 4 times 6 hours, first it was 4 times 7 and I think there are a lot of colleagues who are not in the same department who didn't even know that I work less, because I just did my thing and that worked out fine. But I do feel like I'm kind of at a tipping point now with those 4 times 6 hours. It does get harder to schedule meetings and things like that. If it were to become even less hours, I think it does become more difficult to continue to fulfill my current position. [...] I, I do notice it when I also, for example, if I work a bit longer and a bit more intensively one day, with a lot of traveling involved, then I notice it at the end of the day. And then I also always try to plan it so that I stay in my bed a little bit longer the next morning. And if I see that coming up in my calendar, then I deliberately block the next morning. That's how it works then. [...] Once a month I go to a cooking club together with a few people from my town. We leave a little before 4 and in the evening around 12 I'm home. And in the meantime, it's running in that kitchen, arranging all kinds of things and that's very intense and burdensome. But I think it's incredibly fun, but I also take some hours off the next morning and I really feel the impact. "* (Participant #57) (No changes in ICECAP-A scores)

Table 4. *Continued.*

Finding	Illustrative fragments from interviews
	"Yes, being able to do things together. With my family and with friends. So that you don't think in advance, well I won't do it, because this just takes too much energy. And sometimes I look forward to it, but then I have to travel again. And if I've been working and I do that, I go out in the evening, yes I do notice that sometimes the next day I'm just really tired. Yes, sometimes I think it is a pity. But on the other hand, you also get energy out of doing it, but the day after I feel the impact, but I also think: yes, but is that a reason not to do it? No, but it is a reason to do less than usual. " (Participant #6) (*Enjoyment -1*)
	"And then I can/you can fight it and sometimes I do. I get angry sometimes too. But that doesn't help either. I just try to focus on how I can organize my activities in a way that it still remains fun. Well that is because I would also like to move to another house and just not have that big garden anymore and not have to worry about someone else. That relieves me a lot". (Participant #10) (*Enjoyment +1*)
	" I do receive increasingly more help from others when working at the farm. I delegate heavier tasks to younger people and my childeren. [...] The work on the farm I actually do together with my son, and he also helps in housekeeping, as well as my wife too, we do it together. [...] Yes, I am becoming increasingly aware that I just have to hand things over and that I shouldn't want to do everything myself. And yes, that is actually going better and better. Yes. [...] I used to have an additional job besides the farm that was also more physically demanding. The job I have now is also a conscious choice, because I know that I have FSHD. This job just takes me physically much less effort, it is more cognitive work rather than physical work. And that's kind of why I made that switch as well." (Participant #47) (*Stability +1; Achievement +1*)
	"The only thing I notice is that I am indeed occasionally reading again. And I really didn't do that for years. I didn't get around to it at all and now I feel like I not only have to make it easier for myself physically, but I also have to slowly start handing things over and save some time to rest for myself. And also to be able to pick up a book once in a while. That is what I am doing now. [...] Yes. I know I have a basic confidence buildup over the years that some things are hard, but I always manage it. And I also always try to think 'okay this or that can't be done anymore, but I'm going to do it that way now or I'm going to let go of that,' like I'm now trying to let go of my volunteer work out of necessity and that I have to start finding ways to make it enjoyable for myself. " (Participant #5) (No changes in ICECAP-A scores)
	"It means for me that I always have to think about how to spend my energy and which people I can meet. For example, I have a friend who is very negative about everything, always. And I do notice that demands more from me than, for example, going to see a friend who is more positive in life. And I do notice that I also need that to keep going. And I can visit that friend, but I don't have to go every week. So, I look at that very carefully. What costs me energy and what saves energy?" (Participant #37) (*Enjoyment -1*)
	"I am very social. But I notice I do increasingly need time to rest. And for me that means sitting on the couch and watching series or movies, I really do like that to just do nothing, not have to do anything. And well, sometimes I just notice that in terms of energy, that's really the only thing I can do. Or well, want to, that's more like it, because well, there we come back to my conviction that I think: if you really want something, then it should be possible. Once, I didn't have a very good day, I was very busy at work and I came home and I went grocery shopping and had super efficient grocery shopping, found everything in one go and got out right away. And then I got outside and the sun was shining. And yeah, then I was just really happy for a moment. It was very random, but, it was then. I don't know exactly why either. Maybe the efficiency or that the sun was shining, or I don't know, but so I can also really enjoy the little things in life. Because, well, I worry about that sometimes quite a bit. I'm very curious about how long I can hold on to my current lifestyle. Then I also think, I have to enjoy now those things that are still possible. Maybe I'm also more aware of that than the average person, that the little things in life can also just add a lot of value." (Participant #46) (*Autonomy* -1)

DISCUSSION

Principal findings

Our mixed-methods analysis shows that capturing changes in valuable functionings during rehabilitation for persons with NMD requires both quantitative and qualitative methods. Only using the ICECAP-A would have suggested that no changes occurred, since the average sum level scores were similar between baseline and follow-up. Combining ICECAP-A results with the COPM would indicate differences, with the COPM showing improvement at follow-up, and that changes in occupational performance are apparently not examples of changes in valuable functionings. Only by combining the results obtained by the ICECAP-A and COPM with interviews we could conclude that changes in valuable functionings have occurred, and these changes have (partly) been measured by the COPM and missed by the ICECAP-A.

The use of interviews to categorize participants into those with worsened, unchanged, or improved valuable functionings was supported by corresponding changes in COPM scores. In cases of differences, occupations identified for the COPM did not reflect valuable functionings (which is a broader category than occupations).

In-depth analysis at participant level, using information from interviews, showed that problems at work (paid or voluntary) not always led to lower ICECAP-A scores. Work is not an explicit dimension on the ICECAP-A, and the value of work for persons with NMD (being an active member of society, self-esteem) might not be covered by ICECAP-A dimensions. Previous studies also highlighted this importance and meaning of work for persons with NMD (Bakker et al., 2017; Dany et al., 2017; Minis et al., 2014).

The ICECAP-A might also miss changes because it explicitly asks participants to rate their abilities *at that moment*, providing only a snapshot of their capabilities over time. Interviews revealed that, due to the progressive nature of NMD and associated fatigue, participants were often enforced to make *trade-offs* in realizing valuable functionings to save energy or anticipate future problems (Bakker et al., 2017; Landfeldt et al., 2019; Schipper et al., 2017; Smith et al., 2014). Unlike the ICECAP-A, both the COPM and interviews allow participants to provide information about changes over time, choices they make between functionings, and about advantages and disadvantages associated with changes. For example, they may explain that they have recently spent more energy on social activities, leading to a higher *attachment* score on the ICECAP-A, but that this negatively influenced their energy available to perform at work, leading to a lower score on the ICECAP-A dimension of *achievement*.

Consequently, (summed) ICECAP-A scores may mask trade-offs and changes over time (Karimi et al., 2016).

Another explanation for changes missed by the ICECAP-A might be that our sample size was insufficient. This, in combination with the lower number of levels of the ICECAP-A (four per dimension) than the COPM (1-10), and short follow-up time (6 months), may be another explanation for the lower sensitivity of the ICECAP-A in our study. However, recruiting larger samples is challenging for rare diseases like DM1 and FSHD. Therefore, we conducted a comprehensive mixed-methods analysis, as recommended by a recent scoping review, to get rich data per participant (Whittal et al., 2021).

Strengths and limitations

Interpreting our results is challenging due to absence of a gold standard for evaluating rehabilitation impact on capabilities, leaving the 'true' effect uncertain. However, the correspondence between improvements in COPM scores and the information from interviews and electronic health records suggests a link between rehabilitation and improved valuable functionings. And, although the COPM was developed for assessing performance and satisfaction with meaningful occupations, it may serve as a proxy for measuring changes in capabilities during rehabilitation since occupational performance is part of one's capabilities (Hammell, 2022).

We operationalized *capabilities* as *valuable functionings*. Although the CA distinguishes between what people are *able to do* ('capabilities') and what they *actually do* ('functionings'), we argue that it is difficult for participants to envision their *potential* abilities. Moreover, people with impaired health status may have a relatively large gap between capabilities and functionings, and therefore achievement of functionings may matter most to them (Al-Janabi, 2018). From information about functionings one can draw inferences about underlying capabilities, i.e., logically, functionings cannot be realized without respective capabilities (Karimi et al., 2016; Rijke, Meerman, et al., 2023). By asking participants about why they value particular functionings and the choices they make, one can derive information about whether one had options or not (the freedom aspect of capabilities) (Fleurbaey, 2006).

Recommendations and future research

When developing instruments for evaluating impact of rehabilitation on capability well-being of persons with NMD, more attention should be given to fatigue (energy management) and (paid or voluntary) work (being able to actively contribute to society). Additionally, flexibility is needed to capture the variety of functionings important to persons with NMD and trade-offs that they make in spending time and

energy on specific valuable functionings. Therefore, to evaluate whether individuals experienced changes in capabilities during rehabilitation it would be recommended to adopt a mixed-methods approach. By combining quantitative and qualitative information about *what* (if anything) has changed in participants' functionings and *how,* and the value of these changes, comprehensive data can be obtained.

Exploring the relation between occupational performance measures and capabilities could offer additional insights. While the COPM allows individualized identification of important occupations, its focus on identifying problems in occupational performance (to identify treatment goals) may miss functionings important to persons with NMD. Embedding it into a capability framework that includes intrinsically valuable functionings (and not only practical and necessary occupations) and *beings* (personal aspirations like being a good parent) could offer ways for evaluating impact of occupations on capability well-being.

CONCLUSION

Based on semi-structured interviews, we categorized participants into having worsened, unchanged, or improved valuable functionings during rehabilitation. Changes in COPM scores aligned with improvements in valuable functionings, whereas ICECAP-A scores did not reflect changes in valuable functionings. The ICECAP-A sometimes missed changes related to (paid or voluntary) work and the need for participants to make trade-offs in energy being spend on different valuable functionings. Only by combining information from the ICECAP-A, COPM, and interviews were we able to capture and understand the changes in capability well-being that occurred in the context of rehabilitation for persons with NMD.

REFERENCES

Al-Janabi, H. (2018). Do capability and functioning differ? A study of U.K. survey responses. *Health Economics*, *27*(3), 465-479. https://doi.org/10.1002/hec.3586

Al-Janabi, H., Flynn, T. N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research*, *21*, 167-176. https://doi.org/10.1007/s11136-011-9927-2

Alkire, S. (2002). *Valuing freedoms: Sen's capability approach and poverty reduction*. Oxford University Press.

Ashizawa, T., Gagnon, C., Groh, W. J., Gutmann, L., Johnson, N. E., Meola, G., Moxley, R., Pandya, S., Rogers, M. T., Simpson, E., Angeard, N., Bassez, G., Berggren, K. N., Bhakta, D., Bozzali, M., Broderick, A., Byrne, J. L. B., Campbell, C., Cup, E., . . . Winblad, S. (2018). Consensus-based care recommendations for adults with myotonic dystrophy type 1. *Neurology Clinical Practice*, *8*(6), 507-520. https://doi.org/10.1212/CPJ.0000000000000531

Bakker, M., Schipper, K., Geurts, A. C., & Abma, T. A. (2017). It's not just physical: a qualitative study regarding the illness experiences of people with facioscapulohumeral muscular dystrophy. *Disability and Rehabilitation*, *39*(10), 978-986. https://doi.org/10.3109/09638288.2016.1172673

Bloemen, B., Pijpers, E., Cup, E., Groothuis, J., van Engelen, B., & van der Wilt, G. J. (2021). Care for capabilities: Implementing the capability approach in rehabilitation of patients with neuromuscular diseases. Study protocol of the controlled before-after ReCap-NMD study. *PLoS One*, *16*(12), e0261475. https://doi.org/10.1371/journal.pone.0261475

Cup, E. H., Scholte op Reimer, W. J., Thijssen, M. C., & van Kuyk-Minis, M. A. (2003). Reliability and validity of the Canadian Occupational Performance Measure in stroke patients. *Clinical Rehabilitation*, *17*(4), 402-409. https://doi.org/10.1191/0269215503cr635oa

Dany, A., Rapin, A., Réveillère, C., Calmus, A., Tiffreau, V., Morrone, I., Novella, J. L., Jolly, D., & Boyer, F. C. (2017). Exploring quality of life in people with slowly-progressive neuromuscular disease. *Disability and Rehabilitation*, *39*(13), 1262-1270. https://doi.org/10.1080/09638288.2016.1191552

Deenen, J. C., Arnts, H., van der Maarel, S. M., Padberg, G. W., Verschuuren, J. J., Bakker, E., Weinreich, S. S., Verbeek, A. L., & van Engelen, B. G. (2014). Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology*, *83*(12), 1056-1059. https://doi.org/10.1212/wnl.0000000000000797

Eyssen, I. C., Beelen, A., Dedding, C., Cardol, M., & Dekker, J. (2005). The reproducibility of the Canadian Occupational Performance Measure. *Clinical Rehabilitation*, *19*(8), 888-894. https://doi.org/10.1191/0269215505cr883oa

Eyssen, I. C., Steultjens, M. P., Oud, T. A., Bolt, E. M., Maasdam, A., & Dekker, J. (2011). Responsiveness of the Canadian occupational performance measure. *Journal of Rehabilitation Research and Development*, *48*(5), 517-528. https://doi.org/10.1682/jrrd.2010.06.0110

Finnis, J. (1980). *Natural law and natural rights* (Vol. 31). Oxford University Press.

Fleurbaey, M. (2006). Capabilities, Functionings and Refined Functionings. *Journal of Human Development*, *7*(3), 299-310. https://doi.org/10.1080/14649880600815875

Hammell, K. W. (2022). Securing occupational rights by addressing capabilities: A professional obligation. *Scandinivian Journal of Occupational Therapy*, *29*(1), 1-12. https://doi.org/10.1080/11038128.2021.1895308

Karimi, M., Brazier, J., & Basarir, H. (2016). The Capability Approach: A Critical Review of Its Application in Health Economics. *Value in Health*, *19*(6), 795-799. https://doi.org/10.1016/j.jval.2016.05.006

Landfeldt, E., Edstrom, J., Jimenez-Moreno, C., van Engelen, B. G. M., Kirschner, J., & Lochmuller, H. (2019). Health-Related Quality of Life in Patients with Adult-Onset Myotonic Dystrophy Type 1: A Systematic Review. *Patient*, *12*, 365-373. https://doi.org/10.1007/s40271-019-00357-y

Law, M., Baptiste, S., McColl, M., Opzoomer, A., Polatajko, H., & Pollock, N. (1990). The Canadian occupational performance measure: an outcome measure for occupational therapy. *Canadian Journal of Occupational Therapy*, *57*(2), 82-87. https://doi.org/10.1177/000841749005700207

Lopez Barreda, R., Robertson-Preidler, J., & Bedregal Garcia, P. (2019). Health assessment and the capability approach. *Global Bioethics*, *30*(1), 19-27. https://doi.org/10.1080/11287462.2019.1673028

Mayring, P. (2000). Qualitative content analysis. *Qualitative Social Research*, *1*(2). http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204

Minis, M. A., Satink, T., Kinebanian, A., Engels, J. A., Heerkens, Y. F., van Engelen, B. G., & Nijhuis-van der Sanden, M. W. (2014). How persons with a neuromuscular disease perceive employment participation: a qualitative study. *Journal of Occupational Rehabilitation*, *24*, 52-67. https://doi.org/10.1007/s10926-013-9447-8

Mitchell, P. M., Roberts, T. E., Barton, P. M., & Coast, J. (2017). Applications of the Capability Approach in the Health Field: A Literature Review. *Social indicators research*, *133*, 345-371. https://doi.org/10.1007/s11205-016-1356-8

Mul, K. (2022). Facioscapulohumeral Muscular Dystrophy. *Continuum (Minneap Minn)*, *28*(6), 1735-1751. https://doi.org/10.1212/con.0000000000001155

Pijpers, E. J., Bloemen, B., Cup, E. H. C., Groothuis, J. T., Oortwijn, W. J., van Engelen, B. G. M., & van der Wilt, G. J. (2024). The capability approach in rehabilitation: developing capability care. *Disability and Rehabilitation*, 1-13. https://doi.org/10.1080/09638288.2024.2342494

Rijke, W. J., Meerman, J., Bloemen, B., Venkatapuram, S., Van der Klink, J., & Van der Wilt, G. J. (2023). Strategies for Researching Programs' Impact on Capability: A Scoping Review. *Journal of Human Development and Capabilities*, *24*(3), 1-23. https://doi.org/10.1080/19452829.2023.2209027

Rijke, W. J., Vermeulen, A. M., Willeboer, C., Knoors, H. E. T., Langereis, M. C., & van der Wilt, G. J. (2023). Capability of Children with Hearing Devices: A Mixed Methods Study. *The Journal of Deaf Studies and Deaf Education*, *28*(4), 363-372. https://doi.org/10.1093/deafed/enad010

Robeyns, I. (2005). The Capability Approach: a theoretical survey. *Journal of Human Development*, *6*(1), 93-117. https://doi.org/10.1080/146498805200034266

Rohrbach, P. J., Dingemans, A. E., Essers, B. A., Van Furth, E. F., Spinhoven, P., Groothuis-Oudshoorn, C. G. M., Van Til, J. A., & Van den Akker-Van Marle, M. E. (2021). The ICECAP-A instrument for capabilities: assessment of construct validity and test-retest reliability in a general Dutch population. *Quality of Life Research*, *31*, 687 - 696. https://doi.org/10.1007/s11136-021-02980-5

Schipper, K., Bakker, M., & Abma, T. (2017). Fatigue in facioscapulohumeral muscular dystrophy: a qualitative study of people's experiences. *Disability and Rehabilitation*, *39*(18), 1840-1846. https://doi.org/10.1080/09638288.2016.1212109

Smith, A. E., McMullen, K., Jensen, M. P., Carter, G. T., & Molton, I. R. (2014). Symptom burden in persons with myotonic and facioscapulohumeral muscular dystrophy. *American Journal of Physical Medicine and Rehabilitation*, *93*(5), 387-395. https://doi.org/10.1097/PHM.0000000000000032

Tawil, R., Kissel, J. T., Heatwole, C., Pandya, S., Gronseth, G., & Benatar, M. (2015). Evidence-based guideline summary: Evaluation, diagnosis, and management of facioscapulohumeral muscular dystrophy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the Practice Issues Review Panel of the American

Association of Neuromuscular & Electrodiagnostic Medicine. *Neurology*, *85*(4), 357-364. https://doi.org/10.1212/wnl.0000000000001783

Till, M., Abu-Omar, K., Ferschl, S., Reimers, A. K., & Gelius, P. (2021). Measuring capabilities in health and physical activity promotion: a systematic review. *BMC Public Health*, *21*, 353. https://doi.org/10.1186/s12889-020-10151-3

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*, *19*(6), 349-357. https://doi.org/https://doi.org/10.1093/intqhc/mzm042

Ubels, J., Hernandez-Villafuerte, K., & Schlander, M. (2022). The Value of Freedom: A Review of the Current Developments and Conceptual Issues in the Measurement of Capability. *Journal of Human Development and Capabilities*, *23*(3), 327-353. https://doi.org/10.1080/19452829.2022.2053506

van der Veen, S., Evans, N., Huisman, M., Welch Saleeby, P., & Widdershoven, G. (2023). Toward a paradigm shift in healthcare: using the International Classification of Functioning, Disability and Health (ICF) and the capability approach (CA) jointly in theory and practice. *Disability and Rehabilitation*, *45*(14), 2382-2389. https://doi.org/10.1080/09638288.2022.2089737

Van Duijn, H. M., Niezen, A. A., Cardol, M., Corder-Bartels, I. E., & Verkerk, G. J. Q. (1999). *Canadees meetinstrument voor handelingen en vaardigheden [Canadian Occupational Performance Measure (COPM) - Dutch version]*.

van Engelen, B., & The OPTIMISTIC Consortium. (2015). Cognitive behaviour therapy plus aerobic exercise training to increase activity in patients with myotonic dystrophy type 1 (DM1) compared to usual care (OPTIMISTIC): study protocol for randomised controlled trial. *Trials*, *16*, 224. https://doi.org/10.1186/s13063-015-0737-7

van Hoof, M., Jeuring, S. F. G., Jonkers, D. M. A. E., Masclee, A. A. M., Pierik, M. J., Stokroos, R. J., & Joore, M. A. (2016). De Nederlandse vertaling en indruksvalidatie van de ICECAP-A: meten van kwaliteit van leven volgens de capability-benadering. *Tijdschrift voor gezondheidswetenschappen*, *94*, 313-320. https://doi.org/10.1007/s12508-016-0109-z

van Loon, M. S., van Leeuwen, K. M., Ostelo, R., Bosmans, J. E., & Widdershoven, G. A. M. (2018). Quality of life in a broader perspective: Does ASCOT reflect the capability approach? *Quality of Life Research*, *27*, 1181-1189. https://doi.org/10.1007/s11136-017-1756-5

Veenhuizen, Y., Cup, E. H. C., Jonker, M. A., Voet, N. B. M., van Keulen, B. J., Maas, D. M., Heeren, A., Groothuis, J. T., van Engelen, B. G. M., & Geurts, A. C. H. (2019). Self-management program improves participation in patients with neuromuscular disease: A randomized controlled trial. *Neurology*, *93*(18), e1720-e1731. https://doi.org/10.1212/wnl.0000000000008393

Voet, N., Pater, R., Garmendia, J., Sistiaga, A., Labayru, G., Gallais, B., de Groot, I., Muslemani, S., Gagnon, C., & Graham, C. (2024). Patient-Reported Outcome Measures in Neuromuscular Diseases: A Scoping Review. *Journal of Neuromuscular Diseases*, *11*(3), 567 - 577. https://doi.org/10.3233/JND-240003

Whittal, A., Meregaglia, M., & Nicod, E. (2021). The Use of Patient-Reported Outcome Measures in Rare Diseases and Implications for Health Technology Assessment. *Patient*, *14*, 485-503. https://doi.org/10.1007/s40271-020-00493-w

STATEMENTS & DECLARATIONS

Funding

This work was supported by the Prinses Beatrix Spierfonds (the Dutch neuromuscular foundation) under Grant W.OK18-08. The funder had no role in the study design, writing of the paper, and the decision to submit for publication.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

Conceptualization: Bart Bloemen, Eirlys Pijpers, Wija Oortwijn, Jan Groothuis, Edith Cup, Baziel van Engelen, Gert Jan van der Wilt.

Data Curation: Bart Bloemen and Eirlys Pijpers.

Formal Analysis: Bart Bloemen and Eirlys Pijpers.

Funding Acquisition: Baziel van Engelen, Gert Jan van der Wilt, Jan Groothuis, Edith Cup.

Investigation: Bart Bloemen and Eirlys Pijpers.

Methodology: Bart Bloemen, Eirlys Pijpers, Wija Oortwijn, Jan Groothuis, Edith Cup, Baziel van Engelen, Gert Jan van der Wilt.

Project Administration: Bart Bloemen and Eirlys Pijpers.

Resources: Baziel van Engelen, Jan Groothuis.

Software: Bart Bloemen.

Supervision: Wija Oortwijn, Gert Jan van der Wilt, Edith Cup, Jan Groothuis, Baziel van Engelen.

Visualization: Bart Bloemen and Eirlys Pijpers.

Writing – Original draft: Bart Bloemen

Writing – Review & Editing: Bart Bloemen, Eirlys Pijpers, Wija Oortwijn, Jan Groothuis, Edith Cup, Baziel van Engelen, Gert Jan van der Wilt.

Ethics approval

The study protocol was granted full ethical approval by the medical ethical reviewing committee CMO Regio Arnhem-Nijmegen (NL72794.091.20) and registered at trialregister.nl (NL8946) on October 12, 2020.

Consent to participate

Written informed consent was obtained from all participants before inclusion.

Consent to publish

Consent to publish was included in the written informed consent obtained from all participants before inclusion.

Data availability

Because FSHD and DM1 are rare diseases, the de-identification of data will not guarantee that participants cannot be identified by a combination of diagnosis, sex, and age. Therefore, our data can only be shared in a way inaccessible to the general public. The data generated and analyzed for this study is stored in a Data Acquisition Collection (DAC) in the Radboud Data Repository (https://doi.org/10.34973/jbjb-jp23) to which access can only be obtained after being invited by the first author (BB) and signing a contract. Supporting information (interview guide, codebook, R script used for analyzing data) can be found in an open access Data Sharing Collection (DSC) in the Radboud Data Repository (https://doi.org/10.34973/41aw-zg68).

Code availability

The R script used to conduct the quantitative analyses described in this manuscript can be found in an open access Data Sharing Collection (DSC) in the Radboud Data Repository (https://doi.org/10.34973/41aw-zg68).

ACKNOWLEDGEMENTS

We would like to thank our research assistants Nina de Bakker, Kim Willems, Jana Zajec, and Vanja Rutten, and our intern Rosanne Krakers for supporting with data collection and analysis. We would like to thank statistician Marianne Jonker for support with quantitative data analysis. We would also like to thank Ilse Karnebeek, nurse practitioner, for her involvement in participant recruitment. Finally, we would like to thank all stakeholders involved, especially the persons with NMD who participated in the study, and the members of the rehabilitation team that provided care (both usual and capability care) to the participants.

SUPPLEMENTARY MATERIAL

Supplementary 1. Interview guide ReCap-NMD study

Background

Aim of the interview

The aim of the interview is to map the valuable functionings of the interviewee, whether these have changed during the past 6 months, and his or her thoughts on the contribution (facilitating or impeding) of rehabilitation on these valuable functionings.

Interview method

The interview is semi-structured, which means that the interview protocol is flexible. There are main questions, each with examples of questions, remarks, and suggestions that the researcher can use to obtain an answer to the main question. It is not necessary to ask these questions literally, they serve as a guide for the interview to make sure that the interview provides relevant information for the research questions. Depending on the course of the interview other questions are possible, or the order of the questions can be changed. The topics of the interview will be based on what the interviewee tells, and the researcher will ask questions for clarification or explanation, or to raise additional points.

The aim of the interview is to identify valuable changes since the start of rehabilitation, and to determine what the contribution of rehabilitation (according to the interviewee) has been to these changes. This approach is based on the work of Sabina Alkire (Alkire 2002, *Valuing Freedoms*) who used John Finnis' seven valuable dimensions of life. These dimensions can be used as a starting point for a conversation about changes in capabilities.

Interview guide

Preparation

Beforehand, the researcher conducting the interview reads the notes in the electronic health record relating to the rehabilitation consultations, to get a general understanding of the interviewee's situation. Additionally, the researcher looks at the scores (T0 and T1 if available) of the Canadian Occupational Performance Measure (COPM).

Introduction of the interview

At the start of the interview, the researcher gives a summary of the aim of the research and the interview. The researcher explains his or her role, being independent from the healthcare team. Consent for recording the interview is checked, and after this consent the recording is started.

Question 1: What are the interviewees valuable functionings in the following dimensions?

Dimension (*basic human good*)	Example questions
1) Life: Every aspect of life which is necessary to sustain life and feel comfortable or at ease with oneself. It includes physical and mental health and freedom from injury and suffering. Examples: *self-care, eating and drinking, feeling secure in one's (living) environment, obtaining an income.*	- How would you describe your health? - How does your health impact on your daily functioning? - Where and how do you live? Do you feel comfortable in your living environment? - Do you have an income?
2) Knowledge: Activities pursued for collecting formal and informal knowledge. Examples: *following education/courses, reading the newspaper, watching the news, searching for information on the internet, watching documentaries.*	- Do you follow formal education, or a course or class? - Which topics have your interest to learn more about?
3) Play: Activities that have intrinsic value, they are enjoyed for its own sake and/or help to relax. Examples: *sports, games, crafts, playing music, playing with (grand)children.*	- Which activities do you like to do in your leisure time? - Do you have hobbies or activities that bring you positive energy? - Do you perform sports?
4) Sociability / friendship: Having valuable human relationships and participating in social activities. Examples: *partner, children, family, friends, colleagues, neighbours, contacts in a club or association, peers.*	- How are your social contacts? - How many and what type of social contact do you have? - How is the contact with neighbours, family, friends, acquaintances, colleagues? - What is your living situation? Do you live alone or with others?
5) Aesthetic experience: Experiencing beauty, natural or man-made, by the spectator or creator. Examples: *enjoying nature during a walk/bike ride/road trip, visiting a museum, a city walk (enjoying architecture), enjoying your own creations (painting, drawing, pottery).*	- What are the things you enjoy in life? (Examples: music, art, nature, movie / tv-series)
6) Practical reasonableness: Being able to use knowledge and skills to choose one's actions and lifestyle. Examples: *goalsetting, making plans, making important decisions.*	- How do you make decisions on what to do or not to do in your life? - Is it easy for you to make decisions? Do you take a lot of time to make decisions? - Are you able to make decisions on your own? - Do you discuss your decisions with others?
7) Religion/transcendence: Experiencing meaning in life. It includes religion, but also being part of a community or club, and contributing to society for example by working (paid or voluntary). It is about the experience of being part of something larger or belonging to a community. Examples: *visiting church/mosque/synagogue, spiritual activities, (voluntary) work (belonging to an organisation, contributing to society), membership of a club/association.*	- Are you religious or spiritual? - What is your vision of life? - Does your religion or vision of life influence your experience of living with a neuromuscular disease? - Are you a member of an association or club? - Do you volunteer?

Remarks and recommendations

- For answering question 1, asking how the interviewee would describe his or her health ("*if you were asked to describe your health, how would you say you are doing* currently") is a good starting point, followed by questions on the other dimensions.
- For the different dimensions, a starting point is to ask the interviewees for activities they do or that give them energy. From there, questions such as 'what is the value of this activity for you' or 'why is this activity important to you' can lead to information on the underlying value (dimension).
- Activities or parts of life can belong to multiple dimensions.

Question 2: Since the visit to the department of rehabilitation, have there been any major changes in the valuable functionings of the interviewee in the following dimensions?

I. Life
II. Knowledge
III. Play
IV. Sociability / friendship
V. Aesthetic experience
VI. Practical reasonableness
VII. Religion / transcendence

For this main question the focus is on gathering information on changes that have happened since the visit to the department of rehabilitation; the focus is on changes in the different dimensions that have been discussed during main question 1. It is possible that changes have already been discussed during question 1. The aim of question 2 is to specify when and how these changes have happened, and whether the interviewee experiences these changes as a positive or negative change.

Example questions:

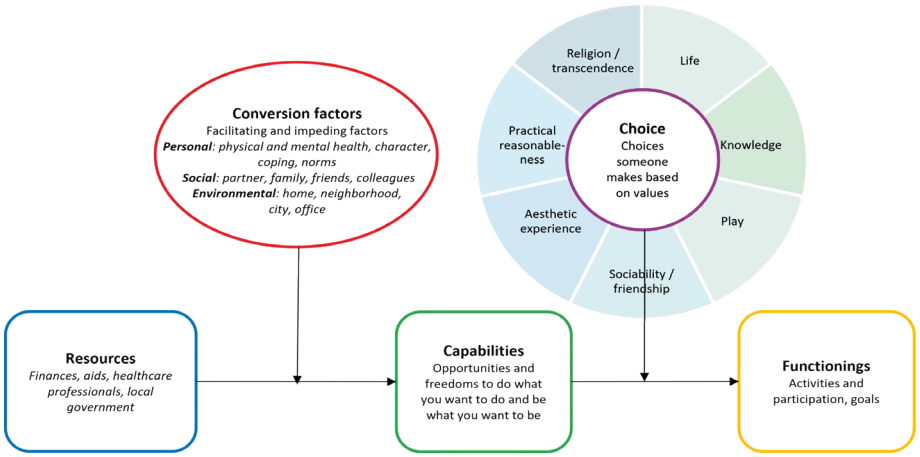
- What are important memories for you from the past six months? Why are these memories important?
- Do you feel that your health remains stable? Can you elaborate on this?
- Would you say that your life has improved over the past few months? Can you elaborate on this?

Remarks and recommendations

- If the interviewee has trouble reminding the visit to the department of rehabilitation, you can help by providing information described in the electronic health record (i.e., topics that have been discussed according to the healthcare professional). This is not preferred however, as the intention is to collect information from the interviewee point of view.

Question 3: According to the interviewee, can any of the changes be attributed to rehabilitation?

For this main question, the aim is to gather information on what has positively or negatively contributed (facilitating and impeding factors) to the changes identified in main question 2. More specifically, we also want to know to what extent rehabilitation has contributed to the realisation of the interviewee's goals. Therefore, we need to know what caused the changes. We make use of the elements of the capability approach:



The functionings (valuable activities) that an interviewee realises can be influenced by the *resources* that an interviewee has access to, personal characteristics, social and environmental factors (*conversion factors*), and the choices a interviewee makes. The questions aim to gather information on these causes.

Example questions:

- What has helped you to perform valuable activities?
- What has helped you to make progress in the discussed dimensions?
- What are, according to you, the causes of these changes?
- Do you receive any formal or informal help? How have you organised this? Have there been any changes in this help over the past few months?

- Do you use assistive devices? Have there been any changes in the use of assistive devices?
- What are your personal characteristics and how do you use these personal qualities? What are your weaknesses, and how do you cope with this? Have there been any changes?
- What was the role of rehabilitation?
- What was the advice from the rehabilitation team? Have there been any actions from the rehabilitation team?
- Have you been referred for further treatment (e.g., physiotherapy, occupational therapy, other/local rehabilitation team, psychology, social work, city council)? What was the referral?
- What was the result of this referral and further treatment?
- Do you encounter any problems at this moment?
- Have the healthcare professionals at the department of rehabilitation discussed the topics that were of value to you?
- Are there any other topics that you would have liked to discuss?
- Has anything been missed by the rehabilitation team?

Remarks and recommendations

· If the interviewee does not mention one of the topics/advice/referrals that are described in the medical file (e.g. referral to a physiotherapist), the researcher could mention this.

Question 4: Do you have any idea in which group (usual/capability) you participated and what makes you think so?

Participants are blinded for treatment group for research purposes. **This blinding needs to be maintained until all participants have completed the study!** However, as a check on whether the blinding has been maintained and on the participants experience, we ask whether the participant has any idea of the treatment group to which they are allocated. To introduce this question, we give a short summary of the research and the two groups. We also indicate the timeline when the participant will be informed about treatment allocation (after completion of the study).

End of interview

At the end of the interview, the researcher explains about the planning of the other measurements (questionnaires and COPM) and explains that the interviewee will receive information about the results of the study. Permission is asked to contact the interviewee again if there is any uncertainty about the answers the interviewee has provided during the interview.

Supplementary 2. Codebook ReCap-NMD interviews

Instructions for using this codebook

- A fragment can only be assigned multiple codes from different categories (e.g., a resource can also be an impact of rehabilitation at the same time, so receiving both the code 'Resource' and 'Experience with rehabilitation care')
- A fragment only receives a code when it provides new information

The coding is used to answer two research questions:

1) Has something changed in the valuable functionings of the participant?
- We assume that functionings or goals mentioned by the participant are also valuable for the participant.
- We want to identify valuable functionings that are *changed* since the participant visit the department of rehabilitation; changes can be:
  - Changes in already existing functionings (reduced or improved permance)
  - Existing functionings of which performance is maintained by the help of an intervention received during rehabilitation (recommendations / advice, assistive device, therapy, etc.)
  - A new alternative functioning representing a similar value as a previous functioning
  - A new alternative functioning representing a new value

2) Why has something changed in the valuable functionings of the participant?
- Factors that influence the effect of rehabilitation, or are the effect itself, can also be *resources* and *conversion factors*
- Resources and conversion factors are only coded when they are related to changed valuable functionings
- If nothing has changed in valuable functionings there are no codes applicable

Codebook

Code (category)	Description	Comments, clarifications
Resource	Resources or access to services that help or hinder the participant in achieving valuable functionings (E.g., finances, unemployment benefits, assistive devices, access to care, access to information).	If the participant already has access to a local health provider (e.g., primary care, physical therapist) this also is a resource.
Personal conversion factor	Personal traits, skills, qualities, and physical and mental state of the participant that enables or hinders the participant in realizing valuable functioning (E.g., personal traits, coping, symptoms - pain, loss of strength, energy; skills, intelligence).	This includes the performance of activities of daily living because this provides information about physical and mental state of the participant. These activities of daily living do not represent valuable functionings unless a daily activity is a goal in itself (has intrinsic value for the patient). E.g., suppose a patient is very concerned about independence, personal care activities could be considered a valuable functioning.
Social conversion factor	Support or resistance from the social environment of the participant that contributes to, or make it more difficult to achieve, valuable functionings (E.g., partner, family, friends, colleagues, etc.; social norms).	
Environmental conversion factor	Helping or hindering factors in the environment of the participant that make it easier or more difficult to achieve valuable functionings, this can include cultural and social norms, facilities for people with disabilities (E.g., adapted home, inaccessible street, COVID-19 measures, etc.).	
Functioning	Changes in valuable activities or goals that the participant can perform again or better, has retained (with help from rehabilitation), would like to perform, has had to give up, or has been replaced by another functioning (to realize the same value), or a new activity or goal. Changes include better / less performance, thinking differently about its value.	We assume that something is valuable if a participant mentioned it multiple times; these are activities or goals that are important for their own sake (have intrinsic value). Goals can also be wished for the future, e.g., continuing to live independently, becoming a parent etc. Fragments in which the participant literally says something about whether something valuable has changed (in general) are also assigned this code.

Choice	Information about choices that the participant makes, both enforced by the situation (consequences of health condition, etc.) or made freely. It concerns fragments where a participant also provides information about *why* he or she is (not) doing or aspiring something; reasons a person gives for considering certain functionings or goals important. These reasons can represent the following underlying values (basic human goods): *life; knowledge; play; Sociability / friendship; aesthetic experience; practical reasonableness (ability to make choices); religion / transcendence*.	It concerns reasons someone gives for regarding something valuable/important as well as regarding something unvaluable/unimportant.
Experience with rehabilitation care	How did the participant experience the visit to the rehabilitation department of Radboudumc (*What was discussed during the consultations with the healthcare professionals? What did participant like/dislike about the care provided?);* What was the advice received from the healthcare professionals (Lifestyle advice? Additional check-ups? Referrals to health care providers, social workers, home care, occupational physician)? Assistive device?) Were subsequent actions (based on the advice received) taken? What happened during rehabilitation? Did the participant experience any effects of rehabilitation?	This concerns information that the participant provides about the experience of the visit to Radboudumc, and information about the follow-up, including steps taken by participant himself or a local healthcare provider.


6

Chapter 6

General discussion

This thesis aimed to explore the normativity of HTA, focusing on the entanglement of norms and evidence. Conceptual and empirical studies were conducted to understand the normativity of HTA, to make it visible, and explore its influence on HTA practice and conclusions of assessments. This final chapter summarizes and integrates the main findings, discusses their implications for HTA practice, and provides recommendations for the integration of normative analysis in HTA.

MAIN FINDINGS

HTA is guided by normative commitments and these can be explicated by unveiling decisions made by HTA practitioners

As argued in **Chapter 2**, HTA is inherently normative. Its procedures and methods pre-suppose norms regarding what makes a health technology desirable (*moral* normativity), which effects are conceivable (*ontological* normativity), and how to obtain reliable information about health technology (*epistemological* normativity). Participating in the practice of HTA *commits* one to these norms. This commitment does not have to be an active explicit endorsement, but in conducting assessments one must make decisions in which certain norms ought to be followed or deviations justified, the latter requiring explication of norms.

Chapter 3 illustrated that this normativity extends to assessing causal claims regarding effects of health technology. Analyzing a published HTA report on Non-Invasive Prenatal Testing (NIPT), it was found that assessing such claims involves defining the (un)desirable effects, which requires normative judgments about its nature. For example, assessing the claim that NIPT is going to enhance reproductive autonomy requires an idea about what reproductive autonomy *is* (ontological commitment) which, given that it is regarded a relevant outcome by stakeholders involved, should simultaneously explain what its features are *that makes it desirable* (moral commitment). This leads to an idea about *which evidence is required* (epistemological commitment) for the assessment (e.g., data on prospective parents' preferences if reproductive autonomy is understood as satisfying their preferences).

By explicating the argumentation used in conducting assessments, the inevitability of normative commitments in HTA can be made visible. This explication broadens the debate on normativity in HTA to include epistemological and ontological norms, and their entanglement with moral norms, which has received less attention in literature until now.

Epistemic normative commitments shape methodology used for assessing medical devices

In **Chapter 4,** we explored how normative commitments shape the procedures and methods used by HTA agencies for assessing medical devices. Using an online survey and in-depth interviews with HTA practitioners, we showed that current methods and procedures for assessing medical devices are still shaped by epistemic norms developed for assessing drugs. The adoption of new methodology (e.g., real-world data, other study designs) for assessing medical devices does not only raise practical concerns (e.g., limited capacity to adopt new methodology, existing regulations that

specify evidence requirements for assessments) but also questions commitments held by HTA practitioners to epistemological norms (i.e., principles of evidence-based medicine).

By guiding methodological decisions in outcome measurement, normative commitments influence conclusions of assessments

In Chapter 5, we report on the results of a mixed-methods analysis comparing different ways to evaluate the impact of rehabilitation for persons with neuromuscular disease (NMD). The capability approach is used to develop alternative outcome measures that could be used in HTA to assess the effects of health technology on quality of life. One of these measures, the ICEpop CAPability measure for Adults (ICECAP-A), is already accepted by some HTA agencies and we compared its results with interviews and a standard rehabilitation outcome measure. Only by combining the ICECAP-A results with information from interviews and the COPM were we able to conclude that valuable changes in the lives of participants, such as improved energy balance and better performance at (paid or unpaid) work, occurred. This shows that different approaches towards evaluating impact, starting from different epistemological (qualitative or quantitative research methods) and moral (utilitarian or capability concept of quality of life) commitments, could lead to different conclusions concerning whether something, and what, has changed in the lives of participants after receiving rehabilitation care.

Norms and evidence in HTA are entangled (if you like it or not)

The term Health Technology *Assessment* suggests that it determines the value of health technology by applying a set of pre-defined norms to available evidence. However, this thesis demonstrates an entanglement between norms and evidence in HTA through two mechanisms: (1) norms influence the types of evidence considered in HTA and (2) norms influence the evidence generation process.

Norms influence the types of evidence considered in HTA

Evidence on the consequences of health technology is not just *out there*, it is actively generated, collected, and interpreted with a purpose. Similarly, when collecting ingredients at a grocery store it is your idea of what you are going to cook (a recipe) that guides your choices. In HTA, it is an idea about what makes a health technology more desirable than its alternatives, and how that can be established in an accepted way, that guides the collection and use of evidence (the 'ingredients') (van der Wilt et al., 2017).

When conducting an assessment, HTA practitioners focus on those consequences of a health technology that *matter to "us"*. It is the identification of ***what*** *makes a*

health technology valuable that enables to pursue the central question that HTA aims to answer: ***how*** *valuable* is this respective technology? (Sen, 1993). Moral commitments guide HTA practitioners in assessing desirable properties of health technology (Oortwijn et al., 2022). E.g., *clinical effectiveness* is considered important because of *our commitment to doing good* (the moral principle of 'beneficence'), and assessing clinical effectiveness requires identification of those outcomes that we regard as beneficial (e.g., what contributes to the well-being of a patient?).

HTA is also guided by the idea that decisions in healthcare should be based on the "best available evidence", a requirement that is mostly operationalized by the principles of evidence-based medicine (EBM) (Moors & Peine, 2016). This commitment emphasizes using specific types of information, assigning the highest weight to quantitative data from randomized controlled trials.

Finally, ontological commitments regarding background theories about mechanisms of disease and health, working mechanisms of health technology, and the organization of healthcare influence the questions and outcomes considered in assessments. For example, if the boundaries of 'health' are defined narrowly, some potential benefits of a health technology, such as improvements in mental well-being or social functioning, might be excluded from consideration in an assessment.

These commitments influence the scope of assessments through norms described in laws and regulations governing HTA processes, often without explicit reference to underlying commitments. They can also become part of HTA practice via informal standards that later become codified in HTA guidelines (Charlton et al., 2023). This implicit nature renders these commitments invisible, but they can come to the surface in specific situations, as illustrated by the cases studied in this thesis:

- In the case of NIPT (Chapters 2, 3), the moral commitment of HTA to maximization of health-related quality of life (measured as quality-adjusted life years, QALYs) becomes problematic. For NIPT, it is unclear who is the beneficiary: the parents or the unborn child? Any decision to consider either the QALYs of the parents or the unborn child results in framing a particular use of this technology as cost-effective (Kibel & Vanstone, 2017). Additionally, there is no consensus on whether maximizing QALYs is the goal of NIPT and, as an alternative, its ability to enhance reproductive autonomy has been assessed.
- In the case of medical devices (Chapter 4), it is sometimes unfeasible to conduct randomized controlled trials, which requires us to re-consider epistemological commitments of HTA. Ignoring or downgrading other types of evidence, like real-world evidence, could render the value of medical devices uncertain. How-

ever, changing evidential standards risks overestimating their value. This shows that epistemic norms can have (unintended) moral consequences, affecting how easily the value of certain technologies can be demonstrated, potentially impacting subsequent recommendations and decisions.

A recently published comparison of assessments of comparative clinical effectiveness of drugs, conducted by two different HTA bodies, also showed that they disagreed on which evidence was suitable to consider in an assessment, contributing to significant disagreements between their assessments (DiStefano et al., 2024).

Norms guiding the evidence generation process

Normative commitments also guide the *generation of* evidence considered in HTA, so not only guiding the selection of which types of evidence to consider, but also the actual generation of evidence itself (e.g., measurement).

One important example is how evidence on the impact of health interventions on patients' quality of life is generated. This is often done using standardized preference-based measures that have two components: (i) a *survey* that asks respondents to rate their experienced problems or functioning in a number of domains (e.g., self-care, pain, daily activities) held to be important to life and potentially affected by health; (ii) a *valuation algorithm* that assigns weights to these domains to convert responses into a number between 0 and 1 expressing the value of the states of being described by that survey (Hausman, 2010; Krabbe, 2016) These weights are estimated based on elicited preferences of the general population or patients themselves.

The development of such measures invokes normative commitments concerning which domains of quality of life should be included in the survey (moral commitments); whether quality of life is something that is subjectively experienced, or is the satisfaction of preferences, or an objective phenomenon (ontological commitments); and who should be surveyed to get reliable information (epistemological commitments) about quality of life (e.g., are patients that have experienced a particular condition more able to judge what it is to live in that condition or are people without the condition less likely to be biased?).

Normative judgments involved in assessing quality of life are already recognized and discussed (Rand & Kesselheim, 2021; Schroeder, 2016). However, what received most attention are value-laden aspects related to the *use* of quality of life measures, whereas values *embedded* in these measures by how they are measured are not discussed as extensively (Schroeder, 2016, 2019). This is an important distinction because value-laden aspects related to the use of measures can be addressed during decision-making

(e.g., applying different thresholds to the use of the incremental cost-effectiveness ratio), whereas value-laden aspects of measurement (e.g., which domains of quality of life or whose preferences to consider) require changes in the evidence generation itself and specific expertise, such as the ability to recognize and understand the influence of value judgments embedded by outcome measures (Schroeder, 2019).

In Chapter 5 we showed that different ways of assessing impact of rehabilitation on quality of life can lead to different conclusions. Using a general instrument (e.g., ICECAP-A) risks missing some effects of rehabilitation because these are not adequately captured by the (high-level) domains of the instrument. Using interviews, we also observed that patients make trade-offs in prioritizing different domains of life to save energy for those aspects that are most valuable to them, and these priorities may differ from the preferences obtained from a general population.

Mostly, HTA practitioners are only responsible for *synthesizing* and *interpreting* available evidence, not for the *evidence generation* itself. However, normative choices embedded in available evidence can significantly influence this interpretation, and judgments based on that evidence require an understanding of how it is generated. Therefore, HTA practitioners must be aware of underlying normative choices to critically appraise the available evidence and understand its strengths and limitations.

IMPLICATIONS FOR HTA PRACTICE

Integrating normative analysis and stakeholder perspectives in HTA is a prerequisite for addressing *normative uncertainty*

Despite established standards for doing HTA, the complexities involved in conducting assessments often require case-based judgments by HTA practitioners and new norms may (implicitly) become established by habit (Charlton et al., 2023). Integrating normative analysis into HTA practice can make these norms explicit, open to scrutiny, and ensure consistency with guiding principles (Charlton & DiStefano, 2024; Charlton et al., 2023). Given HTA's impact on public decision-making, and the diversity of views that may exist in society about the value of health technology, the active participation of stakeholders in HTA is also required to enhance legitimacy of the norms guiding HTA (Baltussen et al., 2017; Oortwijn et al., 2022).

Acknowledging the role of normative commitments in HTA leads to an additional reason for integrating normative analysis and stakeholder participation: to address ***normative uncertainty***. When making decisions, for example about which health technology to implement, we are often *uncertain* about which of the available options will bring about the desirable outcome (e.g., maximizing health gains). This *empirical*

uncertainty results from a lack of crucial information about the actual consequences of different options (MacAskill, 2014; Ongaro & Andreoletti, 2022). The philosophy behind informed decision-making assumes that decision-makers should aim to reduce this uncertainty as much as possible to improve outcomes of their decisions.

However, the uncertainty surrounding decision-making does not stop with empirical uncertainty. Whereas empirical uncertainty leaves us unsure about which health technology is the best option due to a lack of information about their consequences, ***normative uncertainty*** leaves us unsure about *how to evaluate* these consequences (MacAskill, 2014; Ongaro & Andreoletti, 2022). In the example of NIPT, we are not only uncertain about its consequences for the quality of life of prospective parents and the unborn child, but also about how to value these consequences, how to calculate QALYs in this context, and whether quality of life maximization covers the value of NIPT (see Chapters 2 and 3).

In HTA, normative uncertainty arises due to *competing views in society* about what makes a health technology (un)desirable, *unclarity about how established norms* (e.g., costs per QALY) *apply* to situations created by health technology (e.g., NIPT), or due to *conflicts between norms* (e.g., improving *safety* may require measures that reduce *cost-effectiveness* by making the use of a technology more time consuming). Normative uncertainty challenges the idea that HTA's normativity can be addressed by making it explicit and enable public scrutiny retrospectively, or delegating normative choices to others, because in some situations the conduct of an assessment cannot proceed without committing oneself to controversial normative presumptions, and empirical inquiry cannot be isolated from these normative commitments. In these situations, the HTA practitioner may be uncertain about which norms should guide the assessment. In resolving this uncertainty, by making decisions on how to conduct the assessment, the HTA practitioner commits to the relevance of particular outcome measures (moral commitment), the reliability of certain types of information (epistemological commitment), and the inclusion of factors expected to determine the outcomes of health technology (ontological commitment).

Because of the entanglement between norms and evidence, questions about empirical uncertainty cannot be addressed in isolation from normative uncertainty. A health technology is effective in realizing *certain outcomes*, works in *some ways* and is therefore acceptable for *certain stakeholders*. Only collecting available evidence on a particular implementation of a health technology, without recognizing alternative ways in which that technology could be implemented, assumes that this implementation is the most acceptable one. It ignores that the gathering of evidence is an active process in which choices concerning the (potential) use of a health technology are already

made, and that statements about what works (is effective) shape healthcare practice and decision-making (Ongaro & Andreoletti, 2022; Wehrens & de Graaff, 2024).

As discussed in Chapter 2, making normative commitments subject to normative analysis (conducted together with stakeholders) could resolve that issue. The principle of a *sensitivity analysis*, already applied in the conduct of HTA to address empirical uncertainty, could be extended to normative analysis. Just as a sensitivity analysis is used to empirically observe (and quantify) the influence of uncertainty in evidence and its analysis by varying key parameters or assumptions and record the impact on conclusions, normative presumptions could be varied to evaluate whether these lead to different findings of an assessment. For example, conducting cost-effectiveness analyses starting with and without assuming equivalent value of QALYs (i.e., irrespective of characteristics of patients) may provide empirical data on the sensitivity of outcomes to these different normative presumptions (Luyten & van Hoek, 2021). Such analysis improves the robustness of outcomes in cases when it can be shown that similar conclusions are reached starting from different normative presumptions (e.g., the technology is considered not cost-effective either from the perspective of assuming equivalent value of QALYs or the alternative perspective of assuming different value of QALYs), and / or provides information to decision-makers about how the outcomes of an assessment depend on underlying normative judgments.

The advantage of such approach is that it brings normative analysis, stakeholder perspectives, and empirical inquiry on an equal footing, and it draws an analogy between established ways to address empirical uncertainty and how to mitigate normative uncertainty. By acknowledging that HTA practitioners can be uncertain about how to conduct an assessment because of reasons that are normative in nature, and that both their expert knowledge and the experiential knowledge of stakeholders relies on normative reasoning, the potential contribution of stakeholder perspectives can also be motivated based on epistemic grounds (and not solely to enhance democratic legitimacy) (Lehoux et al., 2009).

Should HTA practitioners avoid being normative?

Proposals to integrate normative analysis and stakeholder perspectives in HTA practice have already been made (Baltussen et al., 2017; EUnetHTA, 2016; Oortwijn et al., 2022; Refolo et al., 2020; Saarni et al., 2022). However, the integration of normative analysis and stakeholder participation at HTA agencies has been challenging and limited (Bellemare et al., 2018; Wale et al., 2021). This raises questions about why integration has been difficult and whether future attempts would encounter similar problems.

As already mentioned in Chapter 1, a fundamental challenge is the tension between integrating normative analysis and stakeholder perspectives, often perceived as being *subjective*, and HTA's epistemological commitment to providing *objective* information. It is difficult for HTA practitioners working at HTA agencies to fully acknowledge their role in making and evaluating normative judgments because this conflicts with their designated role in the HTA process. They are expected to avoid making normative judgments about how a health technology should be used, both to leave the decision on normative matters to those that have the appropriate authority, and to protect the *objectivity* of their assessments by preventing personal beliefs and preferences from influencing their work; see also Figure 1 (Ducey et al., 2017; Sandman & Heintz, 2014; Syrett, 2016).

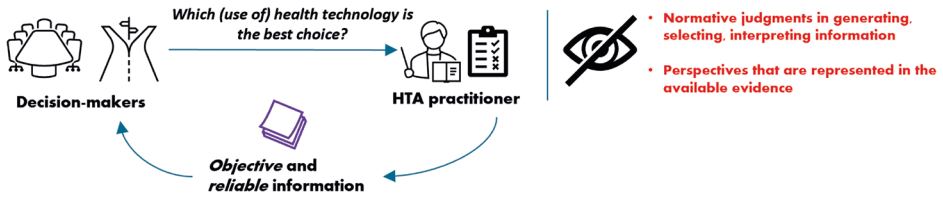


Figure 1. Currently, in most HTA processes those with the authority to make decisions on behalf of society (the *decision-maker*s, i.e., may include stakeholders) may ask *HTA practitioners* (those with the expertise to evaluate information on properties and consequences of health technology) to assess alternative health technologies on their (potential) ability to realize pre-defined goals (decision criteria, e.g., safety, clinical effectiveness, cost-effectiveness). The HTA practitioner is expected to make judgments about the *relevance* and *reliability* of information for supporting judgments about the merits of the alternatives in realizing these goals, while refraining from making any judgments about the appropriateness of the pre-defined goals or the selection of candidate technologies included in the assessment. The assumptions and decisions needed to interpret the available information, and the evidence base itself, are assumed to be impartial and / or any normative presumptions to be recognized and carefully balanced in other parts of the decision-making process, and often remain implicit and hidden from view.

In this role perception an implicit connection is being made between the *reliability* and *objectivity* of information by viewing the *neutrality* of the HTA practitioner as a necessary pre-condition for producing reliable information. 'Objectivity' captures this widespread idea that trust in scientific information is the result from both the reliability of the information provided and the person who collects and interprets the information (Rolin, 2020):

> "*When we call X objective, we endorse it: we say that we rely on X, and that others should do so too. But the word 'objective' is reserved for a specific type of reliance: it is based on the belief that important epistemic risks arising from our* ***imperfections as epistemic agents*** *have been effectively averted*" (Koskinen, 2020).

It are these *imperfections of us as epistemic agents* that we worry about when using the term 'objective', which contrasts with 'subjective', i.e., the worry that individual biases and preferences impede inferring reliable conclusions. Accordingly, normative judgments involved in producing and interpreting information seem to threaten the objectivity and reliability of HTA and open ways for vested interests to influence the decision-making process.

Therefore, epistemological commitments of HTA to principles of evidence-based medicine, emphasizing quantitative and objective information, are not only *epistemological* but also *moral* commitments (Ducey et al., 2017). Considering subjective information in assessments is at odds with the basic idea of HTA to focus on objectively describable dimensions of value and rigorously obtain empirical evidence on what produces improvements in those dimensions (Richardson, 2016). For example, the German HTA agency (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen*, IQWiG), motivates its adherence to principles of evidence-based medicine as follows:

> *"Evidence-based medicine (EBM) is an important basis of the Institute's work. It denotes health care that* ***is not based solely on personal opinions and conventions****, but on proof (evidence). This proof* ***should be obtained*** *using the most* ***objective*** *scientific methods possible and provide* ***reliable*** *results"* [5].

Unsurprisingly, calls to integrate normative analysis and stakeholder perspectives in assessments are sometimes seen as a threat to HTA's intrinsic value and have encountered many difficulties due to conflicts with prevailing epistemic norms. Although HTA practitioners agree that these efforts could be valuable, they do not know how to include certain types of relevant information (e.g., stakeholder perspectives, information from interviews with patients, information on values) that they should consider as less reliable according to their epistemological guidelines (Gunn et al., 2021; Gunn et al., 2023; Moes et al., 2020; Steffensen et al., 2022).

We agree that HTA practitioners should maintain a *neutral attitude*, refraining as much as possible from making normative judgments based on personal preferences. HTA informs public policies, which should be justified by serving publicly articulated purposes that transcend individual interests (Richardson, 2016). However, we challenge whether a neutral attitude is sufficient to ensure objective and reliable results. The entanglement of norms and evidence means that evidence might already reflect specific interests. Staying agnostic about these normative issues risks reinforcing vested interests and obscures

5 See the website of IQWiG: https://www.iqwig.de/en/about-us/methods/evidence-based-medicine/ (accessed on June 4, 2024)

the normative aspects of evidence to decision-makers. Objectivity is better achieved by actively engaging with diverse value perspectives, allowing the influence of these perspectives to be explored, making the evidence base more inclusive, and prevent any single perspective from dominating results (Koskinen, 2022).

Normative analysis *should* also be central to HTA to ensure policy is more responsive to the ways in which health technology can reshape our values. Health technologies are not neutral instruments, they are proposed solutions for solving problems, based on assumptions about which health problems matter and how to address them (Giacomini et al., 2013). Its use requires certain acts from stakeholders (e.g., monitoring devices that can be used at home demand that a patient takes certain measures). As technologies fulfill their purposes, they may also shift our values. For instance, the omnipresence of health checks, such as screening programs and diagnostic tests, might increase support for the idea that health is manageable by individual actions and change our ideas about individual responsibility for health (Stol et al., 2016).

Therefore, health technology is another way of *doing ethics*, which is not accounted for in standard models of HTA that assume that pre-defined criteria can be used to evaluate health technology, unaffected by an influence of health technology on our morality (Smits et al., 2022). To address this, HTA should make explicit, scrutinize, and list alternative ways in which a health problem could be solved, identifying different policy options (either technologies or other relevant interventions) (van der Wilt et al., 2022). This requires a certain openness to different conceptions of value, and imagining diverse uses for a health technology, rather than adhering to the normative assumptions fixed in available evidence (e.g., studies that only tested a particular implementation of a health technology and evaluated it on selected outcomes, preferences elicited before the new health technology arose) (Richardson, 2016).

Assumptions about the nature and desirability of health technology underpin any assessment of its value. To evaluate, one needs a preliminary understanding of what to look for, where and when to find it, and how to interpret the appropriateness of results. At the same time, the results of any assessment are also supposed to increase, and possibly modify, our understanding of the value of a health technology. Thus, our ideas about value both shape and emerge from the assessment process. This raises a paradox: how can an assessment, guided by values, also be the source of these values?

To illustrate the possibility of this reciprocal relation between values and assessment, we may look at how we develop a sensitivity for certain tastes. Imagine that you, for the first time in your life, are drinking a cup of coffee. You will probably notice that it is warm, that it has a certain color, and you may perceive some general indistinctive flavors (you

may not even like it). Despite that it did not really thrill you, you continue drinking coffee and by trying out different types of coffee you gradually acquire a taste for it. Not only do you like coffee now, you are also able to perceive all kinds of aromas that are displayed by different types of coffee. You can distinguish between a *Kopi Luwak* and a *Monsooned Malabar* coffee. Not only is your *understanding* of coffee increased, you also have a different *judgment* about its sensory qualities (its value).

Recognizing this reciprocal relation between values and assessment blurs the line between establishing facts and evaluating. It points towards a *transformative* view on assessment: the collection of information in HTA is not meant to assemble the bare facts about health technology, but to know whether a particular technology can be regarded valuable. An assessment is *part of the valuing of a health technology*, not necessarily in a judgmental kind, but to find out whether, when, and how the technology could be used, and whether policy measures are needed to realize its value.

There is no need to be afraid of normativity in HTA: redefining the role of HTA practitioners

Besides answering the *why* question, it also important to address the *how* question concerning the integration of normative analysis into HTA.

As illustrated in Figure 2, the assessment process should start with a conversation between decision-makers, HTA practitioners, and stakeholders to specify the scope (research questions), objectives, and epistemic criteria that the assessment needs to satisfy. The idea is that an alignment between the goal of the assessment (i.e., the decision that it needs to inform) and the epistemic qualities of the knowledge being produced by the assessment is sought by specifying the epistemic goals, preferences and constraints that guide the assessment. Reasoning from a broader societal aim to an epistemic task that aligns with that aim reveals what being responsive to that aim means in terms of epistemic characteristics of an assessment. The epistemic task can then be fulfilled by the HTA practitioner that can make assessment-related decisions (e.g., which outcome measures to include, how to rate the certainty of evidence, set thresholds, weigh different types of evidence etc.) by reference to the assigned task (Parker, 2024). This process should be constrained by allowing only room for assessments that respect basic requirements for adequate science and include multiple value perspectives (in cases of controversial topics, a sensitivity analysis could be conducted as describe above). This makes room for establishing a shared problem space in which different perspectives contribute to determining what is being assessed and what this means for interpreting the value of health technology (Gunn et al., 2023).

The *VALues in Doing Assessments of health TEchnologies* (VALIDATE) approach offers a way of operationalizing this in HTA (Oortwijn et al., 2022; van der Wilt et al., 2022). This approach helps HTA practitioners to, together with decision-makers and stakeholders, explicate the *type of policy problem* for which an HTA needs to be conducted. Based on the method of *reconstructing interpretive frames*, using diverse methods (e.g., interviews, focus groups, review of grey and scientific literature, government documents etc.) the HTA practitioner tries to identify and explicate the different views that exist in society on a particular health problem. These views consist of assumptions about the nature of the problem (*background theory*), what ought to be pursued (*ethical commitment*), which specific situation(s) is regarded problematic (*problem definition*), and what are appropriate solutions (*judgment of solution*).

For example, increasing waiting lists in mental healthcare might be seen as an urgent problem (*problem definition*) due to mental conditions interfering with daily life activities and causing suffering (*background theory*), with alleviating this suffering regarded a collective responsibility (*ethical commitment*), leading to the judgment that digital technologies could be valuable by reducing waiting lists (*judgment of solution*) (van der Wilt et al., 2022). However, another perspective may argue that mental conditions are over-diagnosed (*problem definition*); that mental conditions are complex and individual responses to life situations, determined by context and social relations (*background theory*); that persons with these conditions should be listened to (*ethical commitment*); and that digital health technology should facilitate a dialogue between these persons, their environment and healthcare professionals (*judgment of solution*).

These different views on the central problem faced by mental healthcare result in different ideas about the potential use of digital health technology, which has consequences for how its value should be assessed. From one view, these technologies should be assessed on their ability to reduce waiting lists (make delivery of care more efficient), whereas from the other view it should be assessed on their ability to stimulate a dialogue between patients and their environment. Therefore, reconstructing these views before conducting an assessment helps structuring the assessment. If stakeholders differ in their ethical commitments, they may also disagree about which questions and types of evidence are relevant. An HTA that does not acknowledge these differences in normative presumptions risks being uninformative because it does not answer the questions that stakeholders may have, and its results could be challenged on normative grounds (Moret-Hartman et al., 2007; van der Wilt et al., 2022).

The VALIDATE approach is an example of how normative commitments could be explicated and addressed in the HTA process. However, its implementation could be challenged by HTA practitioners' commitments to neutrality and objectivity, adher-

ing to principles of evidence-based medicine and pre-defined assessment criteria, as we have shown in Chapter 4 in the case of new methodologies for assessing medical devices. Epistemological analyses that show how the explicit consideration of different normative presumptions can lead to objective results, like our analysis of mixed claims in Chapter 3, would be helpful in reconciling the commitments of HTA practitioners with the ideas underlying an approach like VALIDATE.

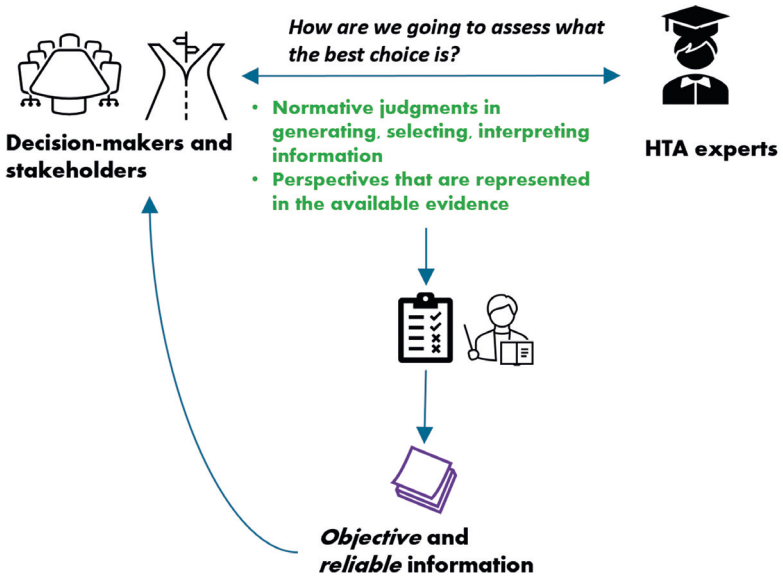


Figure 2. Instead of assuming that the normative commitments underlying the interpretation and generation of information will be recognized and balanced, the assessment process should start with a conversion between *decision-makers*, *stakeholders*, and *HTA experts* that aims to establish the scope and normative presumptions guiding the assessment, and a critical scrutiny of the available evidence to identify gaps in terms of perspectives (concerns, values) not being represented. The outcome of this is a protocol for the conduct of the actual assessment by *HTA practitioners*, potentially including the need for a sensitivity analysis evaluating the influence of different contested normative presumptions on the conclusions of the assessment.

FUTURE RESEARCH DIRECTIONS

In this thesis we have made use of concepts like 'normative', 'ethical, 'value judgment', for which there is no consensus about their exact meaning and there are different interpretations (Bellemare et al., 2018; Charlton et al., 2023). The challenge is not solely to provide definitions, that can be done, but that there are no strict boundaries of these concepts. What is considered a value judgment, especially when contrasted with a factual (or epistemic) judgment, may be contested because these terms invoke different ideas about the appropriate tasks and responsibilities of those

involved in HTA processes. Therefore, as we have done in this chapter, discussions about normativity in HTA should consider these connections with, and implications for, the expectations and perceptions of HTA practitioners concerning their role in evidence-informed decision-making and could be further informed by research on (changes in) different role perceptions (Bauer & Kastenhofer, 2019). It also relates to ideas about the required expertise of HTA practitioners. If they are expected to be involved in conducting normative analysis, some basic knowledge of ethical theories may be required, and ethicists could be embedded in the HTA process to support in explicating and evaluating normative arguments (Refolo et al., 2020).

Because NIPT, medical devices, and rehabilitation are examples of morally challenging technologies (NIPT) and types of interventions not yet commonly assessed by HTA agencies (medical devices, rehabilitation), it may be that we have identified issues that are less salient in other areas of HTA. How often situations of normative uncertainty arise in HTA, and how extensive it is, may be subject to further empirical inquiry. Still, there are already general discussions within the HTA community that expose normative uncertainty:

- Should the scope of HTA be broadened? In literature, HTA practitioners discuss whether HTA should *expand its scope* to consider broader aspects of value beyond safety, effectiveness, and cost-effectiveness, and non-health benefits of health technology (Daniels et al., 2015; Kinchin et al., 2023). Another parallel discussion is on whether, and how, the use of HTA should be broadened to non-pharmaceutical interventions (which has been its traditional focus) (Enzing et al., 2021).
- Should non-RCT data be considered more extensively in HTA? In light of the development of technologies that can not easily be evaluated within the study design of an RCT (e.g., medical devices), situations in which a study population is too small to obtain a sufficient amount of data (e.g., rare diseases, personalized healthcare), and conflicts between experiences of patients and study results, the 'gold standard' of evidence in HTA (RCTs) is increasingly contested. Alternative or supplementary types of information (e.g., real-world data, patient-based evidence, qualitative data) are proposed but raise difficult questions about how to judge their reliability and how they could contribute to the practice of HTA (Gunn et al., 2023; Makady et al., 2017; Moes et al., 2020; Stafinski et al., 2022; Staniszewska & Soderholm Werko, 2021; Steffensen et al., 2022; Szabo et al., 2024; Wehrens & de Graaff, 2024).

However, we acknowledge that an extensive normative analysis would not always be necessary. Especially given the limited capacity and time available for HTA, and that conducting extensive HTAs also has costs, it would be helpful to develop ways

for identifying cases where a formal normative analysis would not be required. This could be done in conjunction with efforts to develop *rapid* or *adaptive* HTA, approaches towards HTA that, by using rapid review methodology or re-use of already published evidence and HTA reports, try to reduce the time needed to conduct HTA (Nemzoff et al., 2023). Besides existing criteria to trigger such rapid approach (e.g., urgency, certainty, low budget impact), the lack of normative uncertainty could be an additional trigger for rapid HTA. Learning from situations in which normative certainty did arose could help in identifying factors where it is to expected, e.g., situations in which existing norms do not easily apply (e.g., digital health technologies that have features not described by current guidelines) or when there is high dissent in society about the appropriate use of a technology (e.g., to which extent genome modification should be used). More guidance should also be developed on how HTA practitioners can make appropriate trade-offs between available time and capacity and different desirable features of a normative analysis.

REFERENCES

Baltussen, R., Jansen, M. P. M., Bijlmakers, L., Grutters, J., Kluytmans, A., Reuzel, R. P., Tummers, M., & der Wilt, G. J. V. (2017). Value Assessment Frameworks for HTA Agencies: The Organization of Evidence-Informed Deliberative Processes. *Value in Health*, *20*(2), 256-260. https://doi.org/10.1016/j.jval.2016.11.019

Bauer, A., & Kastenhofer, K. (2019). Policy advice in technology assessment: Shifting roles, principles and boundaries. *Technological Forecasting and Social Change*, *139*, 32-41. https://doi.org/10.1016/j.techfore.2018.06.023

Bellemare, C. A., Dagenais, P., S, K. B., Beland, J. P., Bernier, L., Daniel, C. E., Gagnon, H., Legault, G. A., Parent, M., & Patenaude, J. (2018). Ethics in Health Technology Assessment: A Systematic Review. *International Journal of Technology Assessment in Health Care*, *34*(5), 447-457. https://doi.org/10.1017/S0266462318000508

Charlton, V., & DiStefano, M. (2024). An empirical ethics study of the coherence of NICE technology appraisal policy and its implications for moral justification. *BMC Medical Ethics*, *25*(1), 28. https://doi.org/10.1186/s12910-024-01016-0

Charlton, V., DiStefano, M., Mitchell, P., Morrell, L., Rand, L., Badano, G., Baker, R., Calnan, M., Chalkidou, K., Culyer, A., Howdon, D., Hughes, D., Lomas, J., Max, C., McCabe, C., O'Mahony, J. F., Paulden, M., Pemberton-Whiteley, Z., Rid, A., . . . Wester, G. (2023). We need to talk about values: a proposed framework for the articulation of normative reasoning in health technology assessment. *Health Economics, Policy and Law*, *19*(2), 153 - 173. https://doi.org/10.1017/S1744133123000038

Daniels, N., Porteny, T., & Urritia, J. (2015). Expanded HTA: Enhancing Fairness and Legitimacy. *International Journal of Health Policy and Management*, *5*(1), 1-3. https://doi.org/10.15171/ijhpm.2015.187

DiStefano, M. J., Pearson, S. D., Rind, D. M., & Zemplenyi, A. (2024). How Do the Institute for Clinical and Economic Review's Assessments of Comparative Effectiveness Compare With the German Federal Joint Committee's Assessments of Added Benefit? A Qualitative Study. *Value in Health*, *27*(8), 1066-1072. https://doi.org/10.1016/j.jval.2024.04.015

Ducey, A., Ross, S., Pott, T., & Thompson, C. (2017). The moral economy of health technology assessment: an empirical qualitative study. *Evidence & Policy*, *13*(1), 7-27. https://doi.org/10.1332/174426415X14443053123024

Enzing, J. J., Knies, S., Boer, B., & Brouwer, W. B. F. (2021). Broadening the application of health technology assessment in the Netherlands: a worthwhile destination but not an easy ride? *Health Economics, Policy and Law*, *16*(4), 440-456. https://doi.org/10.1017/S1744133120000237

EUnetHTA. (2016). *HTA Core Model version 3.0*.

Giacomini, M., Winsor, S., & Abelson, J. (2013). Ethics in Health Technology Assessment: Understanding Health Technologies as Policies. *Healthcare Management Forum*, *26*(2), 72-76. https://doi.org/10.1016/j.hcmf.2013.04.004

Gunn, C. J., Bertelsen, N., Regeer, B. J., & Schuitmaker-Warnaar, T. J. (2021). Valuing patient engagement: Reflexive learning in evidence generation practices for health technology assessment. *Social science & medicine*, *280*, 114048. https://doi.org/10.1016/j.socscimed.2021.114048

Gunn, C. J., Regeer, B. J., & Zuiderent-Jerak, T. (2023). A HTA of what? Reframing through including patient perspectives in health technology assessment processes. *International Journal of Technology Assessment in Health Care*, *39*(1), e27. https://doi.org/10.1017/S0266462323000132

Hausman, D. M. (2010). Valuing health: a new proposal. *Health Economics*, *19*(3), 280-296. https://doi.org/10.1002/hec.1474

Kibel, M., & Vanstone, M. (2017). Reconciling ethical and economic conceptions of value in health policy using the capabilities approach: A qualitative investigation of Non-Invasive Prenatal Testing. *Social science & medicine*, *195*, 97-104. https://doi.org/10.1016/j.socscimed.2017.11.024

Kinchin, I., Walshe, V., Normand, C., Coast, J., Elliott, R., Kroll, T., Kinghorn, P., Thompson, A., Viney, R., Currow, D., & O'Mahony, J. F. (2023). Expanding health technology assessment towards broader value: Ireland as a case study. *International Journal of Technology Assessment in Health Care*, *39*(1), e26. https://doi.org/10.1017/S0266462323000235

Koskinen, I. (2020). Defending a Risk Account of Scientific Objectivity. *The British Journal for the Philosophy of Science*, *71*(4), 1187-1207. https://doi.org/10.1093/bjps/axy053

Koskinen, I. (2022). Participation and Objectivity. *Philosophy of Science*, *90*(2), 413 - 432. https://doi.org/10.1017/psa.2022.77

Krabbe, P. F. M. (2016). *The measurement of health and health status: concepts, methods and applications from a multidisciplinary perspective.* Elsevier / Academic Press.

Lehoux, P., Daudelin, G., Demers-Payette, O., & Boivin, A. (2009). Fostering deliberations about health innovation: what do we want to know from publics? *Social science & medicine*, *68*(11), 2002-2009. https://doi.org/10.1016/j.socscimed.2009.03.017

Luyten, J., & van Hoek, A. J. (2021). Integrating Alternative Social Value Judgments Into Cost-Effectiveness Analysis of Vaccines: An Application to Varicella-Zoster Virus Vaccination. *Value in Health*, *24*(1), 41-49. https://doi.org/10.1016/j.jval.2020.07.011

MacAskill, W. (2014). *Normative uncertainty* [PhD thesis, University of Oxford].

Makady, A., Ham, R. T., de Boer, A., Hillege, H., Klungel, O., Goettsch, W., & GetReal Workpackage 1. (2017). Policies for Use of Real-World Data in Health Technology Assessment (HTA): A Comparative Study of Six HTA Agencies. *Value in Health*, *20*(4), 520-532. https://doi.org/10.1016/j.jval.2016.12.003

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2020). Questions regarding 'epistemic injustice' in knowledge-intensive policymaking: Two examples from Dutch health insurance policy. *Social science & medicine*, *245*, 112674. https://doi.org/10.1016/j.socscimed.2019.112674

Moors, E., & Peine, A. (2016). Valuing Diagnostic Innovations: Towards Responsible Health Technology Assessment. In M. Boenink, H. van Lente, & E. Moors (Eds.), *Emerging Technologies for Diagnosing Alzheimer's Disease: Innovating with Care* (pp. 245-261). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-54097-3_13

Moret-Hartman, M., van der Wilt, G. J., & Grin, J. (2007). Health technology assessment and ill-structured problems: a case study concerning the drug mebeverine. *International Journal of Technology Assessment in Health Care*, *23*(3), 316-323. https://doi.org/10.1017/s0266462307070481

Nemzoff, C., Shah, H. A., Heupink, L. F., Regan, L., Ghosh, S., Pincombe, M., Guzman, J., Sweeney, S., Ruiz, F., & Vassall, A. (2023). Adaptive Health Technology Assessment: A Scoping Review of Methods. *Value in Health*, *26*(10), 1549-1557. https://doi.org/10.1016/j.jval.2023.05.017

Ongaro, M., & Andreoletti, M. (2022). Non-Empirical Uncertainties in Evidence-Based Decision Making. *Perspectives on Science*, *30*(2), 305-320. https://doi.org/10.1162/posc_a_00416

Oortwijn, W., Sampietro-Colom, L., & (Eds.). (2022). *The VALIDATE handbook. An approach on the integration of values in doing assessments of health technologies* (W. Oortwijn & L. Sampietro-Colom, Eds.). Radboud University Press. https://doi.org/10.54195/CKHB1659

Parker, W. S. (2024). The Epistemic Projection Approach to Values in Science. *Philosophy of Science*, *91*(1), 18-36. https://doi.org/10.1017/psa.2023.107

Rand, L. Z., & Kesselheim, A. S. (2021). Controversy Over Using Quality-Adjusted Life-Years In Cost-Effectiveness Analyses: A Systematic Literature Review. *Health Affairs (Millwood)*, *40*(9), 1402-1410. https://doi.org/10.1377/hlthaff.2021.00343

Refolo, P., Bond, K., Bloemen, B., Autti-Ramo, I., Hofmann, B., Mischke, C., Mueller, D., Nabukenya, S., Oortwijn, W., Sandman, L., Stanak, M., Steele, D., van der Wilt, G. J., & Sacchini, D. (2020). Core competencies for ethics experts in health technology assessment. *International Journal of Technology Assessment in Health Care*, *36*(6), 534-539. https://doi.org/10.1017/S0266462320001968

Richardson, H. S. (2016). Intelligence and Transparency in Health Technology Assessment. *International Journal of Technology Assessment in Health Care*, *32*(1-2), 3-9. https://doi.org/10.1017/S0266462316000106

Rolin, K. H. (2020). Objectivity, trust and social responsibility. *Synthese*, *199*, 513 - 533. https://doi.org/10.1007/s11229-020-02669-1

Saarni, S. I., Uusitalo, S., & Autti-Ramo, I. (2022). The role of ethical analysis in conducting a health technology assessment of medical treatments for gender dysphoria. *International Journal of Technology Assessment in Health Care*, *38*(1), e82. https://doi.org/10.1017/S0266462322003257

Sandman, L., & Heintz, E. (2014). Assessment vs. appraisal of ethical aspects of health technology assessment: can the distinction be upheld? *GMS Health Technology Assessment*, *10*. https://doi.org/https://doi.org/10.3205/hta000121

Schroeder, S. A. (2016). Value Choices in Summary Measures of Population Health. *Public Health Ethics*, *10*(2), 176 - 187. https://doi.org/10.1093/phe/phw032

Schroeder, S. A. (2019). Which values should be built into economic measures? *Economics and Philosophy*, *35*(3), 521-536. https://doi.org/10.1017/s0266267118000317

Sen, A. (1993). Capability and well-being. In M. C. Nussbaum & A. K. Sen (Eds.), *The Quality of Life* (pp. 31-49). Clarendon Press, Oxford England. https://doi.org/10.1093/0198287976.001.0001

Smits, M., Ludden, G. D. S., Verbeek, P.-P., & van Goor, H. (2022). Responsible design and assessment of a SARS-CoV virtual reality rehabilitation programme: guidance ethics in context. *Journal of Responsible Innovation*, *9*(3), 344-370. https://doi.org/10.1080/23299460.2022.2076986

Stafinski, T., Glennie, J., Young, A., & Menon, D. (2022). HTA decision-making for drugs for rare diseases: comparison of processes across countries. *Orphanet Journal of Rare Diseases*, *17*, 258. https://doi.org/10.1186/s13023-022-02397-4

Staniszewska, S., & Soderholm Werko, S. (2021). Mind the evidence gap: the use of patient-based evidence to create "complete HTA" in the twenty-first century. *International Journal of Technology Assessment in Health Care*, *37*, e46. https://doi.org/10.1017/S026646232100012X

Steffensen, M. B., Matzen, C. L., & Wadmann, S. (2022). Patient participation in priority setting: Co-existing participant roles. *Social science & medicine*, *294*, 114713. https://doi.org/10.1016/j.socscimed.2022.114713

Stol, Y. H., Schermer, M. H. N., & Asscher, E. C. A. (2016). Omnipresent Health Checks May Result in Over-responsibilization. *Public Health Ethics*, *10*(1), 35 - 48. https://doi.org/10.1093/phe/phw034

Syrett, K. (2016). Expanded HTA, Legitimacy and Independence Comment on "Expanded HTA: Enhancing Fairness and Legitimacy". *International Journal of Health Policy and Management*, *5*(9), 565-567. https://doi.org/10.15171/ijhpm.2016.75

Szabo, S. M., Hawkins, N. S., & Germeni, E. (2024). The extent and quality of qualitative evidence included in health technology assessments: a review of submissions to NICE and CADTH. *International Journal of Technology Assessment in Health Care*, *40*(1), e6. https://doi.org/10.1017/S0266462323002829

van der Wilt, G. J., Bloemen, B., Grin, J., Gutierrez-Ibarluzea, I., Sampietro-Colom, L., Refolo, P., Sacchini, D., Hofmann, B., Sandman, L., & Oortwijn, W. (2022). Integrating Empirical Analysis and Normative Inquiry in Health Technology Assessment: The Values in Doing Assessments of Health Technologies Approach. *International Journal of Technology Assessment in Health Care*, *38*(1), e52. https://doi.org/10.1017/s0266462321001768

van der Wilt, G. J., Gerhardus, A., & Oortwijn, W. (2017). Toward Integration in the Context of Health Technology Assessment: The Need for Evaluative Frameworks. *International Journal of Technology Assessment in Health Care*, *33*(5), 586-590. https://doi.org/10.1017/S0266462317004354

Wale, J. L., Thomas, S., Hamerlijnck, D., & Hollander, R. (2021). Patients and public are important stakeholders in health technology assessment but the level of involvement is low - a call to action. *Research Involvement and Engagement*, *7*(1), 1. https://doi.org/10.1186/s40900-020-00248-9

Wehrens, R., & de Graaff, B. (2024). Working with epistemic uncertainties: Emerging entanglements within conditional reimbursement practices. *Health Policy and Technology*, *13*(2). https://doi.org/10.1016/j.hlpt.2024.100850

Appendix

SUMMARY

Health technology assessment (HTA) is a multidisciplinary process that, using explicit methods, seeks to determine the value of health technology. The purpose of HTA is to inform stakeholders about all possible consequences of health technology to make informed choices about (de)implementation. Increasingly, countries are using the outcomes of HTA to make and justify decisions, for example on the reimbursement of health technology as part of a benefits package.

HTA practitioners recognize that their practice is inherently *normative*; after all, HTA outcomes imply conclusions or recommendations about how we *should* use health technology. This normativity also concerns doing an HTA itself. Determining the value of health technology requires a normative framework for identifying *relevant* information and interpreting it in terms of its implications for the choices to be made. For example, determining the cost-effectiveness of a health technology requires making a statement about what are desirable outcomes (the effects) and, in some countries, setting a threshold to conclude when a health technology can be considered cost-effective.

Despite this normativity being increasingly recognized by its practitioners, the way HTA is institutionalized still often (implicitly) assumes a separation between those who are responsible for conducting an *assessment*, collecting and interpreting the available information on possible consequences of health technology, and those who are responsible for the *appraisal*, formulating recommendations and/or making choices regarding the (de)implementation of the health technology in question. This suggests that making normative statements about the value of health technology can be excluded from the assessment by making them at a different stage in the HTA process.

This practice leads to the remarkable situation that despite recognition of the normativity of HTA, and calls to address it, HTA practitioners must simultaneously avoid this normativity to fulfil their assigned role in the decision-making process. HTA practitioners are expected to remain neutral, their personal views and interests should not interfere with the collection of relevant information and reliable interpretation of the information. This legitimizes their contribution to the process and helps create an objective basis for decisions that serve the public interest.

Although it is understandable that HTA practitioners are expected to avoid normative judgements as much as possible, the question is whether this is possible in practice. In addition to the role already mentioned for normative frameworks in identifying

relevant information, philosophers of science and social scientists have pointed out several ways in which norms and information become entangled in collecting and interpreting evidence.

It is this entanglement of norms and empirical information, and its implications for HTA, that is the focus of this thesis. The research questions to be addressed are:

- How can the normativity of HTA be understood and made visible?
- What is the influence of this normativity on the procedures and methods used in HTA?
- What is the influence of this normativity on conclusions of assessments?

Understanding and making visible the normativity of HTA

In **Chapter 2,** we describe how the normativity of HTA can be understood as the result of *normative commitments*, obligations to follow certain norms, that HTA practitioners make by participating in the practice of HTA. Based on examples from the literature and an analysis of a case study, an assessment of the non-invasive prenatal test (NIPT), we showed that while conducting an assessment there are a variety of decisions to be made that bind the HTA practitioner to *moral* (regarding what makes a health technology *desirable*), *ontological* (regarding what effects of health technology are *conceivable*), and *epistemological* (regarding how *reliable* information about the effects of health technology can be obtained) norms.

In **Chapter 3**, we analyse an HTA report on NIPT, showing how the effects of NIPT have been assessed in practice by evaluating *mixed claims*. These mixed claims connect (implicit) value judgements about desirable effects (e.g., that NIPT should increase reproductive autonomy) with empirical information (that reproductive autonomy could be measured by surveying preferences of people with a desire to have children). Ignoring the normative nature of such claims risks hiding their normativity, presenting results of an assessment as self-evident and not in need of any moral justification. Therefore, while conducting an assessment, it is important to identify and make explicit the implicit value assumptions and evaluate their impact on the conclusions of an assessment. This can include evaluating the extent to which conclusions are independent of normative assumptions, which gives more insight into the robustness of findings.

The influence of normative commitments on methods and procedures in HTA

In **Chapter 4**, we explore the role that normative commitments play in the adoption of new methods for assessing medical devices. Using an online survey, we identified the procedures and methods currently used by HTA practitioners to assess medical

devices. Interviews with HTA practitioners, with a focus on the case of transcatheter aortic valve implantation (TAVI), provided insight into their views on appropriate methods and role of HTA in medical device assessment. The results show that medical device assessments are mainly based on epistemic principles developed for assessment of drugs, and that (in addition to practical factors) commitments to the principles of evidence-based medicine hamper the adoption of new methods. This could lead to delayed or incomplete assessments of the value of medical devices.

The influence of normativity on conclusions of an assessment

In **Chapter 5**, we examine whether, and how, the *capability approach* can be used in evaluating the impact of rehabilitation for persons with neuromuscular diseases. Evaluating effects of health technology on health-related quality of life, measured with a generic questionnaire, is a common part of HTA. However, there is normative debate among experts about this approach, with the discussion focusing on what aspects are considered important when measuring quality of life. The capability approach states that it is important to look at the opportunities that people have to do or be what is of value to them. This approach was translated into a measurement instrument, the ICEpop CAPability measure for Adults (ICECAP-A), that can be used in HTA. We administered this ICECAP-A to persons with neuromuscular diseases before and after rehabilitation. The results were compared with a validated instrument used in rehabilitation studies to measure effects (Canadian Occupational Performance Measure, COPM), and interviews with participants about which valuable changes in their functioning they experienced during rehabilitation. Only by combining the ICECAP-A results with information from interviews and the COPM were we able to conclude that changes, such as improved energy balance and better performance at (paid or unpaid) work, occurred. This shows that the normative choice for a measurement method can influence the conclusions of an assessment.

Discussion and implications

In **Chapter 6**, we discuss the conclusions and implications of our results. We conclude that the entanglement of norms and information in HTA results from two mechanisms: (i) moral, epistemological, and ontological normative commitments influence which types of evidence are considered in an assessment and (ii) norms play a role in generating evidence about the consequences of health technology, particularly in evaluating the impact of health technology on quality of life. We discuss how this inevitable normativity of HTA need not be seen as a threat to its reliability and legitimacy based on an understanding of '*objectivity*' that allows room for value perspectives. We suggest how integration of normative analysis and stakeholder participation into HTA can help in realizing this form of objectivity.

SAMENVATTING

Health technology assessment (HTA) is een multidisciplinair proces dat, met gebruikmaking van expliciete methoden, de waarde van gezondheidstechnologie probeert vast te stellen. Het doel van HTA is om betrokkenen te informeren over alle mogelijke consequenties van gezondheidstechnologie om weloverwogen keuzes te maken over (de)implementatie. In toenemende mate maken landen gebruik van de uitkomsten van HTA om besluiten te nemen, en te rechtvaardigen, bijvoorbeeld over de vergoeding van gezondheidstechnologie via het verzekerde pakket.

HTA-beoefenaars erkennen dat hun praktijk inherent *normatief* is; de uitkomsten van HTA impliceren immers conclusies of aanbevelingen over hoe we gezondheidstechnologie *zouden moeten gebruiken*. Maar de normativiteit betreft ook het doen van een HTA zelf. Het bepalen van de waarde van gezondheidstechnologie vraagt om een normatief kader voor het identificeren van *relevante* informatie en het interpreteren van deze informatie in termen van haar betekenis voor de keuzes die gemaakt moeten worden. Bijvoorbeeld, het bepalen van de kosteneffectiviteit van een gezondheidstechnologie vraagt om een uitspraak over wat wenselijke uitkomsten zijn (de effecten), en in bepaalde landen om het vaststellen van een drempelwaarde om te concluderen wanneer een gezondheidstechnologie als kosteneffectief kan worden beschouwd.

Ondanks dat deze normativiteit steeds breder erkend wordt door haar beoefenaars gaat de manier waarop HTA geïnstitutionaliseerd is nog vaak (impliciet) uit van een scheiding tussen diegene die verantwoordelijk zijn voor een *assessment*, het verzamelen en interpreteren van de beschikbare informatie over mogelijke consequenties van gezondheidstechnologie, en degene die verantwoordelijk zijn voor de *appraisal*, het formuleren van aanbevelingen en/of het maken van keuzes ten aanzien van de (de)implementatie van de betreffende gezondheidstechnologie. Dit suggereert dat het doen van normatieve uitspraken over de waarde van gezondheidstechnologie buiten de assessment gehouden kan worden door ze te laten plaatsvinden in een andere fase in het proces.

Deze praktijk leidt tot de opmerkelijke situatie dat HTA-beoefenaars ondanks erkenning van de normativiteit van HTA, en oproepen om deze te adresseren, ze deze normativiteit tegelijkertijd moeten vermijden om hun toegewezen rol in het proces te kunnen vervullen. Er wordt van HTA-beoefenaars verwacht dat ze neutraal blijven, hun persoonlijke opvattingen en belangen mogen het verzamelen van relevante informatie en een betrouwbare interpretatie van de informatie niet in de weg staan. Dit

legitimeert hun bijdrage aan het proces en helpt om een objectieve basis te creëren voor besluiten die het publieke belang dienen.

Ook al is het begrijpelijk dat van HTA-beoefenaars verwacht wordt om normatieve oordelen zoveel mogelijk te vermijden, de vraag is of dit in de praktijk mogelijk is. Naast de al genoemde rol voor normatieve kaders in het identificeren van relevante informatie, hebben wetenschapsfilosofen en sociale wetenschappers gewezen op verscheidene manieren waarop normen en informatie met elkaar verstrengeld zijn in het verzamelen en interpreteren van *bewijsvoering*.

Het is deze verstrengeling van normen en empirische informatie, en haar implicaties voor HTA, die in dit proefschrift centraal staan. De onderzoeksvragen hierbij zijn:

- Hoe kan de normativiteit van HTA het best begrepen worden en zichtbaar worden gemaakt?
- Wat is de invloed van deze normativiteit op de methoden en procedures die gebruikt worden in HTA?
- Hoe beïnvloedt deze normativiteit de conclusies van een assessment?

Het begrijpen en zichtbaar maken van de normativiteit van HTA

In **hoofdstuk 2** beschrijven we hoe de normativiteit van HTA kan worden begrepen als het resultaat van *normatieve commitments*, verplichtingen om bepaalde normen te volgen, die HTA-beoefenaars aangaan door deel te nemen aan de praktijk van HTA. Op basis van voorbeelden uit de literatuur en een analyse van een casus, een assessment van de niet-invasieve prenatale test (NIPT), laten we zien dat tijdens het uitvoeren van een assessment er allerlei beslissingen moeten worden gemaakt die de HTA-beoefenaar verbinden aan *morele* (betreffende wat een gezondheidstechnologie *wenselijk* maakt), *ontologische* (betreffende welke effecten van gezondheidstechnologie *denkbaar* zijn), en *epistemologische* (betreffende hoe *betrouwbare informatie* over de effecten van gezondheidstechnologie verkregen kan worden) normen.

In hoofdstuk 3 analyseren we een HTA-rapport over NIPT, waarbij we laten zien hoe de effecten van NIPT in de praktijk beoordeeld zijn door het evalueren van *mixed claims*. Deze *mixed claims* verbinden (impliciete) waardeoordelen over wenselijke effecten (bijvoorbeeld dat NIPT reproductieve autonomie zou moeten vergroten) met empirische informatie (dat reproductieve autonomie gemeten zou kunnen worden door voorkeuren van mensen met een kinderwens in kaart te brengen). Het negeren van het normatieve karakter van dergelijke claims riskeert dat deze normativiteit verborgen blijft, waarbij resultaten van een assessment worden gepresenteerd als vanzelfsprekend en niet behoeftig aan enige morele rechtvaardiging. Daarom is

het belangrijk om tijdens het uitvoeren van een assessment de impliciete waarde veronderstellingen te identificeren en expliciet te maken, en te evalueren wat hun impact is op de conclusies van een assessment. Hierbij kan worden geëvalueerd in welke mate de conclusies onafhankelijk zijn van normatieve veronderstellingen, wat meer inzicht geeft in de robuustheid van bevindingen.

De invloed van normatieve commitments op methoden en procedures in HTA

In **hoofdstuk 4** verkennen we de rol van normatieve commitments bij de adoptie van nieuwe methoden voor het beoordelen van medische hulpmiddelen. Met behulp van een online enquête hebben we de procedures en methoden in kaart gebracht die momenteel worden gebruikt door HTA-beoefenaars bij de beoordeling van medische hulpmiddelen. Interviews met HTA-beoefenaars, met een focus op de casus percutane aortaklepimplantatie (*Transcatheter Aortic Valve Implantation*, TAVI), gaven inzicht in hun opvattingen over gepaste methoden en rol van HTA bij de beoordeling van medische hulpmiddelen. De resultaten tonen aan dat beoordelingen van medische hulpmiddelen voornamelijk gebaseerd zijn op epistemische principes ontwikkeld voor de beoordeling van geneesmiddelen, en dat (naast praktische factoren) commitments aan de principes van evidence-based medicine de adoptie van nieuwe methoden bemoeilijken. Dit zou kunnen leiden tot een vertraagde of onvolledige waardebepaling van medische hulpmiddelen.

De invloed van normativiteit op de conclusies van een assessment

In **hoofdstuk 5** hebben we onderzocht of, en hoe, de *capability benadering* gebruikt kan worden bij het evalueren van de impact van revalidatie voor personen met neuromusculaire aandoeningen. Het evalueren van effecten van gezondheidstechnologie op *gezondheid-gerelateerde kwaliteit van leven*, gemeten met een generieke vragenlijst, is een gangbaar onderdeel van HTA. Er is onder experts echter normatieve discussie over deze benadering, waarbij de discussie focust op *welke aspecten* belangrijk worden gevonden bij het meten van kwaliteit van leven. De capability benadering stelt dat het belangrijk is om te kijken naar de *mogelijkheden* die mensen hebben om datgene te doen of zijn wat voor hen waardevol is. Deze benadering is vertaald naar een meetinstrument, de *ICEpop CAPability measure for Adults* (ICECAP-A), die gebruikt kan worden in HTA. Wij hebben deze ICECAP-A afgenomen bij personen met neuromusculaire aandoeningen voor en na revalidatie, en de resultaten vergeleken met een gevalideerd instrument dat in revalidatiestudies gebruikt wordt om effecten te meten (*Canadian Occupational Performance Measure*, COPM), en interviews met deelnemers over wat zij aan waardevolle veranderingen in hun functioneren hebben ervaren tijdens de revalidatie. Alleen door de ICECAP-A resultaten te combineren met informatie uit interviews en COPM hebben we kunnen concluderen dat er veranderingen, zoals verbeterde energiebalans en betere uitvoering van (betaalde of

onbetaalde) arbeid, hebben plaatsgevonden. Dit laat zien dat de normatieve keuze voor een meetmethode invloed kan hebben op de conclusies van een assessment.

Discussie en implicaties

In **hoofdstuk 6** bespreken we de conclusies en implicaties van onze resultaten. We concluderen dat de verstrengeling van normen en informatie in HTA het gevolg is van twee mechanismen: (i) morele, epistemologische, en ontologische normatieve commitments beïnvloeden keuzes voor welke bewijsvoering wordt meegenomen in een assessment; en (ii) normen spelen een rol in het genereren van bewijsvoering over de consequenties van gezondheidstechnologie, in het bijzonder bij het in kaart brengen van de impact van gezondheidstechnologie op kwaliteit van leven. We bespreken hoe deze onvermijdelijke normativiteit van HTA niet gezien hoeft te worden als een bedreiging voor haar betrouwbaarheid en legitimiteit op basis van een begrip van 'objectiviteit' dat ruimte laat voor waarde perspectieven. Wij doen een voorstel voor hoe integratie van normatieve analyse en stakeholder participatie in HTA kan helpen om deze vorm van objectiviteit te benaderen.

RESEARCH DATA MANAGEMENT

Ethics and privacy

The study described in Chapter 4 was based on the results of research involving human participants. Written informed consent was obtained from participants that responded to the survey, and oral consent was obtained from interviewees, to collect and process their data for this research project.

The study described in Chapter 5 was based on the results of medical-scientific research involving human participants, subject to the Medical Research Involving Human Subjects Act (WMO) and was conducted in accordance with the ICH-GCP guidelines (Good Clinical Practice). The recognized Medical Ethics Review Committee 'CMO Regio Arnhem-Nijmegen' has given approval to conduct this study (file number: NL72794.091.20). Written informed consent was obtained from participants to collect and process their data for this research project.

Pseudonymized data were stored and analyzed on the department server, only accessible by project members working at Radboudumc. The pseudonymization key was stored separately from the research data.

Data collection and storage

Data for Chapter 4 was collected by researchers. The online survey tool CheckMarket was used for sending out secured questionnaires, and Microsoft Teams and an Olympus voice recorder was used to conduct and record interviews. Survey data was analyzed using CheckMarket. Interviews were analyzed using summaries, validated by interviewees, and Atlas.Ti.

Data for Chapter 5 was collected by researchers and research assistants, and directly manually entered into the Castor Electronic Data Capture (EDC) system. Castor EDC was used for secured online questionnaires. Quantitative data was analyzed using R version 4.1.3, whereas qualitative data was transcribed verbatim and analyzed using Atlas.Ti. Paper (hardcopy) data is stored in cabinets on the department and can only be accessed by people with authorization to enter the department.

Data sharing according to the FAIR principles

All study results are or will be published open access. Meta-data, supporting information, and aggregated data are published with restricted access in Data Sharing Collections (DSCs) in the Radboud Data Repository (RDR), see the details in the table below. The raw data collected for Chapter 5 is stored in a Data Acquisition Collection (DAC) in the RDR to which access can only be obtained after being invited by the

first author (BB) and signing a contract All data will remain available for at least 15 years after termination of the studies.

Chapter	Data Sharing Collection (DSC) / Data Acquisition Collection (DAC)	DSC License
4	*DSC collection: https://doi.org/10.34973/s07v-9e02*	*RUMC-RA-DUA-1.0*
5	*DSC collection: https://doi.org/10.34973/41aw-zg68* *DAC collection: https://doi.org/10.34973/jbjb-jp23*	*CC-BY-NC-SA-4.0* *Not applicable (closed access)*

LIST OF PUBLICATIONS

Journal articles included in this thesis

Bloemen, B., Oortwijn, W. & van der Wilt, G.J (2024). Understanding the Normativity of Health Technology Assessment: Ontological, Moral, and Epistemological Commitments. *Health Care Analysis*. https://doi.org/10.1007/s10728-024-00487-x

Bloemen, B., & Oortwijn, W. (2024). Assessing medical devices: a qualitative study from the VALIDATE perspective. *International Journal of Technology Assessment in Health Care*, *40*(1), e29. https://doi.org/10.1017/S0266462324000254

Bloemen, B., Jansen, M., Rijke, W., Oortwijn, W., & van der Wilt, G. J. (2021). Mixed claims in Health Technology Assessment: The case of Non-Invasive Prenatal Testing. *Social science & medicine*, *270*. https://doi.org/10.1016/j.socscimed.2021.113689

Journal articles not included in this thesis

Pijpers, E. J., **Bloemen, B**., Cup, E. H. C., Groothuis, J. T., Oortwijn, W. J., van Engelen, B. G. M., & van der Wilt, G. J. (2024). The capability approach in rehabilitation: developing capability care. *Disability and Rehabilitation*, 1-13. https://doi.org/10.1080/09638288.2024.2342494

Karazi, W., Coppers, J., Maas, D., Cup, E., **Bloemen, B**., Voet, N., Groothuis, J. T., Pinos, T., Marti Seves, R., Quinlivan, R., Lokken, N., Vissing, J., Bhai, S., Wakelin, A., Reason, S., & Voermans, N. C. (2024). Toward an understanding of GSD5 (McArdle disease): How do individuals learn to live with the metabolic defect in daily Life. *Journal of Neuromuscular Diseases*, *11*(1), 103-116. https://doi.org/10.3233/JND-230027

Rijke, W. J., Meerman, J., **Bloemen, B**., Venkatapuram, S., Van der Klink, J., & Van der Wilt, G. J. (2023). Strategies for Researching Programs' Impact on Capability: A Scoping Review. *Journal of Human Development and Capabilities*, *24*(3), 401-423. https://doi.org/10.1080/19452829.2023.2209027

Refolo, P., Sacchini, D., **Bloemen, B**., Grin, J., Gutierrez-Ibarluzea, I., Hofmann, B., Oortwijn, W., Raimondi, C., Sampietro-Colom, L., Sandman, L., van der Wilt, G. J., & Spagnolo, A. G. (2023). On the normativity of evidence - Lessons from philosophy of science and the "VALIDATE" project. *European Review for Medical and Pharmacological Sciences*, *27*(23), 11202-11210. https://doi.org/10.26355/eurrev_202312_34560

Refolo, P., **Bloemen, B**., Corsano, B., Grin, J., Gutierrez-Ibarluzea, I., Hofmann, B., Oortwijn, W., Sampietro-Colom, L., Sandman, L., van der Wilt, G. J., & Sacchini, D. (2022). Prioritization of COVID-19 vaccination. The added value of the "VALIDATE" approach. *Health Policy*, *126*(8), 770-776. https://doi.org/10.1016/j.healthpol.2022.05.005

Van der Wilt, G. J., **Bloemen, B**., Grin, J., Gutierrez-Ibarluzea, I., Sampietro-Colom, L., Refolo, P., Sacchini, D., Hofmann, B., Sandman, L., & Oortwijn, W. (2022). Integrating Empirical Analysis and Normative Inquiry in Health Technology Assessment: The Values in Doing Assessments of Health Technologies Approach. *International Journal of Technology Assessment in Health Care*, *38*(1), e52. https://doi.org/10.1017/s0266462321001768

Grin, J**., Bloemen, B**., Gutierrez-Ibarluzea, I., Hofmann, B., Oortwijn, W., Refolo, P., Sacchini, D., Sampietro-Colom, L., Sandman, L., & van der Wilt, G. J. (2022). Learning and practicing more value-reflective, problem-setting health technology assessment: experiences and lessons from the VALIDATE project. *International Journal of Technology Assessment in Health Care*, *38*(1), e63. https://doi.org/10.1017/S0266462322000204

Bloemen, B., Pijpers, E., Cup, E., Groothuis, J., van Engelen, B., & van der Wilt, G. J. (2021). Care for capabilities: Implementing the capability approach in rehabilitation of patients with neuromuscular diseases. Study protocol of the controlled before-after ReCap-NMD study. *PLoS One*, *16*(12), e0261475. https://doi.org/10.1371/journal.pone.0261475

Refolo, P., Bond, K., **Bloemen, B**., Autti-Ramo, I., Hofmann, B., Mischke, C., Mueller, D., Nabukenya, S., Oortwijn, W., Sandman, L., Stanak, M., Steele, D., van der Wilt, G. J., & Sacchini, D. (2020). Core competencies for ethics experts in health technology assessment. *International Journal of Technology Assessment in Health Care*, *36*(6), 534-539. https://doi.org/10.1017/S0266462320001968

Engel, J., Blanchet, L., **Bloemen, B**., Van den Heuvel, L. P., Engelke, U. H. F., Wevers, R. A., & Buydens, L. M. C. (2015). Regularized MANOVA (rMANOVA) in untargeted metabolomics. *Analytica chimica acta*, *899*, 1-12. https://doi.org/10.1016/j.aca.2015.06.042

Book chapters

Bloemen, B., van der Wilt, G.J. (2022). Chapter 1. First things first. In: Oortwijn, W., & Sampietro-Colom, L. (Eds.). *The VALIDATE handbook: an approach on the*

integration of values in doing assessments of health technologies (pp. 13 – 52). Radboud University Press. https://doi.org/10.54195/CKHB1659

Bloemen, B., van der Wilt, G.J. (2022). Chapter 7. A philosophical summary of the VALIDATE approach. In: Oortwijn, W., & Sampietro-Colom, L. (Eds.). *The VALIDATE handbook: an approach on the integration of values in doing assessments of health technologies* (pp. 143 – 158). Radboud University Press. https://doi.org/10.54195/CKHB1659

PORTFOLIO

Department: **IQ Health**
PhD period: **06/06/2017 – 30/08/2024**
PhD Supervisors: **Prof. dr. G.J. van der Wilt, Prof. dr. B.G.M. van Engelen**
PhD Co-supervisor: **dr. W.J. Oortwijn**

Training activities	**Hours**
Courses	
- Donders Graduate School Day (2020, 2021)	14
- Donders Scientific Integrity Course (2021)	7
- Qualitative Research Methods and Analysis (2020)	84
- Perfecting your academic writing skills (2020)	42
- Donders Graduate School introduction day (2020)	7
- Basiscursus Regelgeving en Organisatie voor Klinisch onderzoekers (BROK) (2020)	42
- Basiskwalificatie Onderwijs (BKO) (2018 – 2019)	60
Seminars	
- Hosting and presenting a webinar about ethics at HTA agencies (2024)	9
- Oral presentation in webinar of the Human Development and Capability Association Health Disability Theme group (2021)	9
Conferences	
- Health Technology Assessment international (HTAi) annual meeting, Seville (including oral presentation in co-hosted workshop, member of a panel) (2024)	40
- Health Technology Assessment international (HTAi) annual meeting, Utrecht (including oral presentations in workshop and panel sessions, co-host of workshop) (2022)	40
- Health Technology Assessment international (HTAi) annual meeting, online (including oral presentations in workshop and panel sessions, co-host of workshop) (2021)	28
- Health Technology Assessment international (HTAi) annual meeting, Cologne (including oral presentation and member of a panel)	40
- CaRe days, Den Bosch (including oral presentation) (2018)	20

Teaching activities	
Lecturing	
- Coordinator of course 8RPH 'Personalized Healthcare Research' in Bachelor program Biomedical Sciences (2018 – 2024)	40
- Coordinator of course MMSS 'Science and Society' in Master program Molecular Mechanisms of Disease (2017 – 2024)	175
- Coordinator of course BMS07 'Science, communication and society' in Master program Biomedical Sciences (2017 – 2024)	175
- Guest lecture about ethics in Health Technology Assessment in post-academic course 'Ethiek in de zorg' (2020 – 2023)	18
- Guest lectures in online Health Technology Assessment (HTA) Training Program for HTA users, doers, and trainers in Ukraine (2023)	16
- Guest lecture in 'InScience' program of HAN University of Applied Sciences, biology and chemistry students, discussing the societal and ethical implications of cryo-preservation after watching the documentary 'Hope Frozen' on that topic (2022)	5
- Coordinator of course BMS15 'Big Data' in Master program Biomedical Sciences (2017 – 2019)	80

Supervision of internships / other	
- Supervision of Biomedical Sciences master literature thesis of 4 weeks (2022)	4
- Supervision of Biomedical Sciences master research internship of 24 weeks (2022)	60
- Supervision of Medicine master research internship of 12 weeks (2021)	30
- Coaching a Biomedical Sciences master research internship of 24 weeks (2021)	60
- Supervision of Biomedical Sciences master literature thesis of 4 weeks (2017)	4
Other activities	
Co-chair of the Health Technology Assessment international (HTAi) interest group on ethics in HTA (2022 - 2024)	30
Total	**1139**

CURRICULUM VITAE

Bart Bloemen werd geboren op 30 Oktober 1987 in Overloon. Na het behalen van zijn havodiploma aan Scholengemeenschap Stevensbeek, begon hij in 2005 aan de studie Bio-informatica aan de HAN te Nijmegen. Na het behalen van zijn bachelor diploma Bio-informatica in 2009 vervolgde hij zijn studie met de master Molecular Life Sciences aan de Radboud Universiteit. In 2009 startte hij aldaar ook met het verkorte bachelor programma Wijsbegeerte, waarvoor hij het diploma behaalde in 2011. Zijn afstudeerscriptie ging over wetenschapsfilosofische aspecten van ontwikkelingen in de moleculaire biologie. De master Molecular Life Sciences rondde hij af in 2013. In zijn onderzoekstages bij de afdelingen Analytische Chemie en Biochemie richtte hij zich op het ontwikkelen van algoritmen voor het analyseren van *metabolomics* data en confocale fluorescentie microscopie data.

In 2013 solliciteerde hij op een promotieplaats bij de Diagnostic Image Analysis Group (DIAG) van de afdeling Radiologie en Nucleaire Geneeskunde aan het Radboudumc. Het promotieproject had als doel het ontwikkelen en evalueren van *machine learning* algoritmen voor het diagnosticeren van netvliesaandoeningen op basis van analyses van *optical coherence tomography* beelden. Tijdens dit jaar kwam hij tot de conclusie dat dit project toch niet goed paste bij zijn interesses en besloot het traject te stoppen.

In 2015 solliciteerde hij op een functie als docent in de Health Technology Assessment groep van de afdeling Health Evidence van het Radboudumc. In 2017 startte hij daar met een promotieonderzoek naar de normativiteit van HTA onder begeleiding van Prof. dr. G.J. van der Wilt, Prof. dr. B.G.M. van Engelen en Dr. W.J. Oortwijn, waarvan de resultaten zijn beschreven in dit proefschrift. Daarnaast was hij betrokken bij het EU-project '*VALues In Doing Assessments of health TEchnologies*' (VALIDATE), coördineerde hij meerdere cursussen in de Bachelor- en Masteropleidingen van de medische faculteit, begeleidde studenten bij hun onderzoeksstage, en had een nevenfunctie als co-chair van de *interest group on ethics* van Health Technology Assessment international (HTAi; de internationale beroepsorganisatie van HTA).

DANKWOORD

Promoveren doe je niet alleen, en gelukkig zijn er veel mensen geweest die het de moeite waard hebben gemaakt en me op verschillende manieren hebben gesteund. Deze personen wil ik hier graag bedanken.

Allereerst wil ik mijn promotoren en copromotor bedanken.

Gert Jan. Heel erg bedankt voor al je vertrouwen in mij, en de steun tijdens al die jaren! Ik solliciteerde in 2015 op eigenlijk een kleine en tijdelijke functie, op een voor mij lastig moment in mijn carrière, maar jouw vertrouwen heeft mij geholpen te groeien en zo werd het uiteindelijk een heel promotietraject. Ik ken weinig mensen met zoveel interesses en brede kennis, samenwerken was dan ook erg inspirerend. Hopelijk doet dit proefschrift een beetje recht aan jouw visie op HTA, die nog veel meer aandacht en navolging zou verdienen. Maar ook je humor en persoonlijkheid maakte het een genoegen om jouw promovendus te zijn.

Baziel. Een tweede promotor, wat een luxe! Ik heb veel van je geleerd, je weet ook snel en precies te benoemen wat iemands talenten zijn en waar dat tot zijn recht komt. Je geeft daarbij ook vaak, zoals je het zelf noemt, 'ongevraagd advies'. Dank voor de vele, soms filosofische, gesprekken als 'Brabanders onder elkaar'.

Wija, mijn copromotor. Onze tweewekelijkse meeting op vrijdag ging vaak net zoveel over persoonlijke dingen en de weekend plannen dan over de inhoud. Na met je kennis te hebben gemaakt binnen het VALIDATE project, werd je gelukkig ook collega binnen het Radboudumc en bleek in een 'Happy Healthy HEV' sessie dat je wel mijn copromotor wilde zijn. Daar was ik heel blij mee en dat bleek een goede keus! Dank voor al je hulp, hopelijk kunnen we in de toekomst in enige vorm blijven samenwerken op het gebied van ethiek en HTA. Die dansvloer op HTAi congressen moet natuurlijk wel levendig blijven!

Graag wil ik ook de leden van de manuscriptcommissie bestaande uit **Prof. dr. Marcel Olde Rikkert**, **Prof. dr. Silvia Evers**, en **Dr. Lotte Krabbenborg** bedanken voor de tijd die zij geïnvesteerd hebben in het beoordelen van mijn proefschrift.

Ik heb tijdens mijn promotie ook veel tijd doorgebracht op de afdeling revalidatiegeneeskunde, waar ik als onderzoeker betrokken was bij de *Rehabilitation and Capability care for patients with Neuromuscular diseases* (ReCap-NMD) studie. Ik wil graag een aantal collega's uit dat team bedanken.

Eirlys. Ik was zelfs nog aanwezig bij jouw sollicitatiegesprek, en ik ben blij dat je destijds ons team kwam versterken. Leuk om samen te werken en onze ervaringen rondom het promoveren te delen! Het was ook erg prettig om taken te kunnen verdelen, het includeren van patiënten, data verzamelen, en alle administratie vergt toch veel werk. Maar er was ook altijd genoeg te bespreken op persoonlijk vlak, soms tijdens lange video calls. Hopelijk kunnen we, maar daar ga ik vanuit, over een jaar ook jouw promotie vieren.

Jan, Edith. Dank voor al jullie tijd en aandacht voor mijn werk binnen de ReCap-NMD studie, wat soms voor jullie misschien wat 'vage HTA' was. Jullie klinische blik was voor mij ook zeer leerzaam en waardevol! Ik wil ook de onderzoeksassistenten **Nina** en **Jana** heel erg bedanken voor alle ondersteuning in het verzamelen van de data, en het afnemen van interviews. Verpleegkundig specialist **Ilse** heel erg bedankt voor al het werk bij de screening en inclusie van patiënten, en het team van revalidatie bedankt voor alle inspanningen en de mogelijkheid om aanwezig te zijn bij gesprekken met patiënten.

I also had the privilege to work together with international colleagues. First, the members of the VALIDATE consortium: **Pietro, Dario, Laura, Carla, Iñaki, Lars, Bjørn, John,** it has been a pleasure to work with you on this EU project. Our project meetings in Amsterdam, Barcelona, and Rome sometimes felt more like a holiday trip instead of work. Although the project is completed, we keep seeing each other at HTAi events and conferences, and hopefully we can work together in the future. And thank you **Dario, Ken, Pietro, Costanza**, for the collaboration in the HTAi interest group on ethics, and the opportunity to be the co-chair of this group.

Dan de collega's van de afdeling Health Evidence, in het bijzonder de HTA-sectie, heel erg bedankt voor de gezellige dagjes uit, koffiepauzes, wijnproeverijen, en andere mooie momenten. **Hans**, leuk om een aantal maanden je kamergenoot te zijn geweest, je relativerende humor kon ik zeer waarderen en het was erg gezellig. **Leon**, ook wij waren een tijd kamergenoot, dank voor alle gesprekken.

In het bijzonder wil ik hier **Wouter** en **Jan** noemen, lotgenoten in de 'capability club'. Onze besprekingen van artikelen en boeken voor de scoping review, samen met Gert Jan, zal ik niet snel vergeten. Jan, dank voor je rust en betrokkenheid, en dat je nog steeds bereid bent om eens bij te praten over onze lopende projecten. Wouter, ik had zelfs de eer om je paranimf te zijn, dank voor alle leuke gesprekken onder het genot van een kop koffie of biertje! Je bent nu zelfs buurtgenoot in Lindenholt, dus die kop koffie of biertje is nu wel heel dichtbij.

Anneke. Buiten je ondersteuning op het secretariaat was het toch ook altijd wel heel gezellig om bij je binnen te lopen en bij te praten, of om samen een wandeling te maken. Daarnaast zijn we ook nog een keer gaan tennissen, met natuurlijk een drankje na afloop. We praten nog steeds zo nu en dan bij, en die afspraak om een keer in Grave op bezoek te komen staat nog steeds.

Ik heb naast mijn onderzoekstaken ook veel in het onderwijs gedaan. In dat verband wil ik de samenwerking met **Paul** nog even benoemen. Heel erg bedankt voor de altijd prettige samenwerking en vele gezellige bezoekjes aan je mooie woning (zelfs een thuisbioscoop!). Leuk dat je ook hebt geholpen met mijn lekenpraatje, hopelijk ben ik daardoor toch wat begrijpelijker geworden voor een breed publiek.

Mijn *Mirror Sessions* bondgenoten: **Mira**, **Pleuntje**, **Ivan**, **Kas**. Het was een super waardevolle toevoeging om samen deze sessies te organiseren, en ervaringen te delen als filosofische indringers in 'harde' wetenschap.

Richelle, ik heb je tempo met bier drinken, en ook met hardlopen, nooit bij kunnen houden maar het was in ieder geval gezellig! Met je begroeting, 'Baarrtje', kon ik al horen dat je in de buurt was. Leuk om eens in de tijd bij te praten samen met Daniëlle en Rene.

Daniëlle. Dank voor alle gezelligheid, biertjes (vaak met Rene en Richelle), wandelingen, BBQs, koffietjes, en kerstkaarten! Je humor en nieuwsgierigheid maken het altijd weer gezellig om af te spreken. Je hebt tien jaar geleden al voorgedaan hoe je een proefschrift moet verdedigen, ik hoop dat ik nu eindelijk iets met dat voorbeeld gedaan heb. En we hebben allebei plannen om Nijmegen een keer te verlaten, maar die borrels moeten toch zeker blijven komen.

Martien. Vele leuke koffiepauzes, plaagstootjes over PSV en Ajax, en biertjes in de Aesculaaf verder heb ik nu het genoegen om met je te kunnen borrelen in Tilburg. Zo kan ik met eigen ogen zien of die stad nu echt even gezellig is als Nijmegen. In ieder geval veel Brabantse gezelligheid, en je kan nu voor jou ook nog nieuwe delen van de stad leren kennen.

Tim en **Robert.** Leuk om samen met jullie eens in de tijd herinneringen op te halen uit de Bio-informatica tijd. De 'legendarische' eerste filmavond, met mijn hele keuken onder het pizzadeeg, heeft niet geleid tot een hele vaste traditie, maar gelukkig wel zo nu en dan een borrel of een 'ruimtelijke ervaring' in Amsterdam.

Ik wil ook een aantal vrienden van Phylisha bedanken die ik inmiddels met veel plezier heb leren kennen. **Nicole**, ik denk dat je een studie filosofie gemist hebt, gezien alle interessante vragen die je stelt. En we komen zeker nog een keer met je nieuwe kitten knuffelen! **Freek**, jouw droge en sarcastische humor kan ik wel waarderen! Excuses voor het je helpen aan een ijsmachine. **Alessandra**, jouw enthousiasme, in het bijzonder ook over de voorbereidingen op de bruiloft (je hebt zelfs de jurk al mogen zien!), werkt zeer aanstekelijk! De etentjes met ons drieën zijn altijd heel erg gezellig, en ik denk dat we in de toekomst nog eens naar jouw oratie gaan luisteren.

Danny en **Sabine.** Ik heb jullie leren kennen als vrienden van Twan, en dat is inmiddels uitgegroeid tot een vaste traditie van spelavonden, vaak tot in de kleine uurtjes, waarbij de nieuwste ProefTuin van Hertog Jan ook weer geproefd wordt. En natuurlijk de jaarlijkse 4Daagse. Danny, je vermogen tot het verzinnen van allerlei bijnamen blijft ook hilarisch, evenals je droge humor en vaak goede humeur!

Rene, paranimf. Gelukkig heeft mijn ooit wat vreemde verzoek om hulp om een raam open te krijgen, en de indruk dat ik maar een 'stagiair' was, je er niet van weerhouden om kennis te maken en al die jaren te blijven afspreken. Ik kijk met veel plezier terug op (filosofische) lunchwandelingen, avondjes bioscoop of borrelen, thuis afspreken en Iwan bestellen (Corona…), barbecueën, jaarwisselingen, vakanties in Zuid-Limburg en Zwolle! Je relativerende kijk op het leven, en in het bijzonder het 'HTA wereldje', humor, en gedeelde interesses (zelfs de voorkeur voor 'Cheese onion chips') werken altijd weer aanstekelijk. Heel erg bedankt voor alle gezelligheid! Inmiddels woon je al niet meer in Nijmegen, maar er zijn genoeg andere steden (Den Bosch) om te borrelen, en hopelijk kunnen we dat, ondanks steeds drukkere agenda's, blijven voortzetten!

Rui, paranimf. We zijn ooit, toen we nog samen op de middelbare school in Stevensbeek zaten, samen naar een open dag van Bio-informatica geweest. Samen ook aan deze studie aan de HAN begonnen, wat een zeer gezellige tijd was! Vervolgens zijn we ook allebei verder gaan studeren aan de uni, alleen ben jij nog wel op het 'rechte pad' van bio-informatica doorgegaan. We hebben ondertussen ook al wat reizen samen gemaakt, maar liefst twee keer Curaçao en ik heb in Mallorca mogen meemaken wat voor een enorme trainingsschema's jij kunt volhouden! Leuk dat je nu mijn paranimf bent, zo voelt de cirkel rond nadat we ooit samen zijn gaan studeren.

Jared. Kleine broertje maar toch ook grote broer van Phylisha. Jouw talenkennis, natuurkundig inzicht, en vaardigheden om bordspellen te winnen is indrukwekkend, ik probeer er wat van op te pikken maar het zal nog steeds niet altijd lukken om jouw puzzels (fles wijn in kistje) op te lossen. Gelukkig kunnen we ook het plezier van het

drinken van speciaalbiertjes delen, en probeer ik soms te nippen aan een whisky of een pittig sausje.

Theo en Hanny. Helaas kunnen mijn eigen grootouders mijn promotie niet meer meemaken, maar met jullie heb ik toch een beetje het gevoel dat ik nog een opa en oma heb. Jullie verhalen over vroeger, met al jullie reizen en ervaringen in het buitenland, zijn erg vermakelijk, en het is altijd weer gezellig om bij jullie video's of fotoalbums uit de oude doos te bekijken. De politiek is ook altijd wel gespreksonderwerp, ik denk dat ik wel mag zeggen dat ik wat linkser in het spectrum zit, maar de discussies zijn alleen maar leuk en leerzaam!

Lieve **Marcel** en **Mariëtte.** Heel erg bedankt voor het warme welkom in jullie familie, ik voelde me meteen thuis en had me geen betere schoonouders kunnen bedenken! Jullie gaan momenteel door een hele moeilijke tijd, heel veel waardering voor hoe jullie ondanks alles toch positief blijven en nog steeds met interesse naar mijn promotietraject bleven vragen. Ik heb al heel veel van jullie geleerd, niet alleen door vele inhoudelijke gesprekken, maar vooral door jullie liefdevolle aandacht voor elkaar en iedereen om jullie heen te mogen ervaren.

Twan, lief broertje. We delen veel dezelfde interesses, en hebben ook vele wandelingen samen gemaakt. Natuurlijk de 4Daagse (voor mij alleen de eerste twee dagen), en zelfs een alternatieve vierdaagse in Marbella. Altijd leuk om samen een biertje te drinken, of het nu bij een wedstrijd van Ajax of een spelavond is. **Fieke en Jos.** Lieve zus en schoonbroer, ook met jullie is het altijd gezellig en ik bewonder jullie doorzettingsvermogen, die was ook weer te zien bij de 4Daagse!

Lieve **pap** en **mam.** Ik goj dit 'n bietje ien 't Lóns probiere, wej kunne ten slotte twie proate. Hiël moi dè òllie d'r altied zien, en mej bej alles steunen. Hiël veul dank vör alles! Misschien dat jullie je soms afvroegen wat doet onze zoon daar allemaal bij het Radboudumc, mijn studierichtingen waren ook al niet de meest makkelijk uitlegbare. Maar jullie bleven altijd luisteren naar alle verhalen en in alles steunen, ook praktisch met alle verhuizingen, klusjes, en de tuin. Ik ben dan ook *biëstig vriëd* met jullie als ouders!

Phylisha, lieverd. Er zijn voor mij veel redenen om met plezier terug te denken aan mijn promotietraject, maar dat ik jou heb leren kennen is de belangrijkste! Wat voorzichtig begon met een idee om eens een keer te gaan badmintonnen, en samen wat eten bij Plek, heeft na mooie vakanties in Parijs, Toscane, en Sicilië, samenwonen, en heel veel leuke herinneringen, geleid tot onze verloving in een sprookjesachtig park in Sevilla. Hopelijk kunnen we dat dit jaar bekronen met een mooie bruiloft. Hou van je!

Radboudumc
university medical center

IN DEI NOMINE FELICITER

Radboud University