KU LEUVEN
DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

KU Leuven
Biomedical Sciences Group
Department of Public Health and Primary Care
Leuven Institute for Healthcare Policy

Accounting for future lives in Health Technology Assessment

Making people healthy or making healthy people?

Evelyn Verbeke

Jury:

Supervisor:	Prof. dr. Jeroen Luyten
Co-supervisor:	Prof. dr. Erik Schokkaert
Chair committee:	Prof. dr. Koen Milisen
Chair defence:	Prof. dr. Pascal Borry
Jury members:	Prof. dr. Yvonne Denier
	Prof. dr. Karen Peeraer
	Prof. dr. Sandy Tubeuf (UCLouvain)
	Prof. dr. Pieter van Baal (Erasmus University)
	Prof. dr. Werner Brouwer (Erasmus University)

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Biomedical Sciences

September 2022

© 2022 KU Leuven – Faculty of Medicine
Uitgegeven in eigen beheer, Evelyn Verbeke, Kapucijnenvoer 35, 3000 Leuven

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Thank You

It is not without reason that this thesis starts with an acknowledgement section, as the most important lesson I learned during the past four years, is that the value of present as well as future lives is undoubtedly determined by the people surrounding and supporting them.

Thank you Jeroen, for your guidance, patience and your support when I was, once in a while, drowning in existential questions (both literally and figuratively). Thank you for always being ready to help and to have the courage to keep unravelling my writings. I am very grateful to have had a supervisor that was both demanding and supportive, that wanted to push me to new heights while also always having my back. Thank you for your trust and giving me the opportunity to go to Princeton on a population ethics adventure and to collaborate with so many interesting people.

Thank you to my co-supervisor Erik, who has the magical ability to provide fundamental insights about a complex topic, while at the same time lighting up the (virtual) room with his humour. I am very grateful that I was able to work with you and that I could always find an open door/phone line when I had questions or concerns. Thanks to your welfare economics class, I discovered an interest in a completely different side of economics, which has broadened my view and will definitely have a lasting impact on my future path.

Thank you to Karen Peeraer and Yvonne Denier for providing me with feedback as well as positive encouragement during my research. Thank you to Thomas D'Hooghe, Annick Bogaerts, Tinne Nuyts, and Neeltje Crombag for the nice collaboration. Thank you to the colleagues from the LIHP. It is a pity that we were not able to see each other more often, but it was nice knowing that we were 'in it together' and it was always a

pleasure to see enthusiastic faces in Leuven. I wish you the very best of luck in the future, thank you Elias, Lina, Miet, Mathijs, Jonas Steel and Jonas Brouwers, Erinn, Britt, Simon, Astrid Van Wilder and Astrid Van Muylder, Fien, Kaat, Dorothea and Charlotte.

Thank you to my family. Thank you to my dad for his absolutely, unconditional support, I really could not have done this without him. He managed to remain enthusiastic about my unreadable drafts during four years and has therefore unintendedly become an expert in my research field. Thank you to my dad and Kaat for providing me a warm home when I needed it the most. Thank you to my sister Myriam, one of my best friends, for believing more in me than I did myself and for introducing me to my favourite Covid-19 hobby. Thank you to my mom, Erik and Margo for the walks, bike and train rides and your valuable advice. Thank you Sven and the new family members, Julien (included on the cover page) and Alois for the nice family moments. Thank you Karin and Wim for inspiring me to explore the field of health economics and thank you Max for supporting me during the first part of this journey.

Thank you to my friends, thanks to the Covid-19 pandemic I realised that I am so, so lucky to have you. Thank you to the tata's (Shakske and Sloor), Vercrusse, Anouk and Flaaf. Thank you for always recharging my batteries (special thanks to Sinterklaas, aka Fifty) and for the unlimited support during the 'laatste loodjes' that lasted plus/minus two years. Thank you Signe and Julie, two true friends that I got to know because of this PhD, I am so happy we could share this experience together. Thank you Milaan, Hannes, Shumie, Leeuwi, Pieter, Karen, Ellen and all the Gurls, the Waanzinners, the Patrons and CC de rijdende bloempjes, every walk, talk, run, yoga session, coffee, swim or bike ride really made a difference.
Last but not least, thank *you*, for taking the time to read this dissertation.

Evelyn Verbeke
April, 2022, Antwerp

Abstract

Different types of healthcare programs not only affect people living today, but also people that will come to exist in the future. Examples are fertility treatments, genetic screening, prevention of vertically transmittable diseases, maternal vaccination and disease eradication programs. It remains an open question whether or not these consequences for future lives should be taken into account in Health Technology Assessment (HTA). The way in which future lives are treated in HTA will however significantly influence the attractiveness of the intervention and subsequently affect the priority-setting decisions made by policy makers. The aim of this PhD thesis is to contribute to the development of a standardized framework by exploring the relevance of 'future life effects' (and costs) for cost-effectiveness calculations, to study the different ways in which these consequences can be taken into account and to build bridges between HTA and other disciplines that study this matter.

This thesis starts with a conceptual approach of future life effects in healthcare and explores how these effects are currently taken into account in the evaluation of perinatal mental health interventions and Assisted Reproductive Technologies. Second, the literature on population ethics is studied and alternative ethical views that could be relevant in the HTA context are evaluated. By means of a survey, the third part of the thesis explores how the general public intuitively value future lives when setting healthcare priorities and how their view corresponds to the previously studied theoretical perspectives. Finally, the lessons learned become practically relevant in a cost-utility analysis of Noninvasive Prenatal Testing (NIPT), an intervention of which cost-effectiveness is largely driven by how the consequences for future offspring are taken into account.

Table of contents

Introduction 1

Chapter 1: The consideration of future lives in cost-effectiveness studies 12

Part 1: Cost-effectiveness of mental health interventions during and after pregnancy
- 1.1.1 Introduction 14
- 1.1.2 Methods 17
- 1.1.3 Results 19
- 1.1.4 Discussion 25
- 1.1.5 Conclusion 28

Part 2: Public funding of Assisted Reproductive Technologies
- 1.2.1 Overview of public funding of ARTs (in selected countries) 103
- 1.2.2 Health-economic evaluation of ARTs 105
- 1.2.3 Limitations of cost-effectiveness studies 111
- 1.2.4 Conclusion 121

Chapter 2: Insights from the population ethics literature 145

Part 1: To be or not to be: future lives in economic evaluation
- 2.1.1 Introduction 146
- 2.1.2 Guidelines and practice 146
- 2.1.3 Non-impartial views in population ethics 148
- 2.1.4 Implications and criticisms 168
- 2.1.5 Conclusion 169

Part 2: Future offspring costs in economic evaluation
- 2.2.1 Future costs 173
- 2.2.2 Future offspring costs 174
- 2.2.3 A proposal 184
- 2.2.4 Conclusion 187

Chapter 3: Valuing future lives in healthcare priority-setting: a population survey **192**

3.1 Introduction 194
3.2 Methods 199
3.3 Results 206
3.4 Discussion 216

Chapter 4: Long-term cost-utility analysis of NIPT in Belgium **225**

4.1 Introduction 227
4.2 Methods 229
4.3 Results 242
4.4 Discussion 252

Discussion 262

List of publications related to this thesis 277

Curriculum Vitae 279

Acknowledgements, Personal Contribution and Conflict of Interest Statements 281

List of abbreviations

ART	Assisted Reproductive Technology
BE	Belgium
CBA	Cost-Benefit Analysis
CBRC	Cross-Border Reproductive Care
CE	Cost-Effectiveness
CEA	Cost-Effectiveness Analysis
CUA	Cost-Utility Analysis
DALY	Disability-Adjusted Life Year
DET	Double Embryo Transfer
EPDS	Edinburgh Postnatal Depression Scale
eSET	elective Single Embryo Transfer
ET	Embryo Transfer
FET	Frozen-thawed Embryo Transfer
FL	Flanders
FLE	Future Life Effect
GIFT	Gamete Intrafallopian Transfer
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
ICSI	Intracytoplasmic Sperm Injection
IVF	In Vitro Fertilization
NHS	National Health Service (UK)
NICE	National Institute for Clinical Excellence
NIPT	Non-Invasive Prenatal Testing
NL	The Netherlands
NMB	Net Monetary Benefit
NML	Net Monetary Loss
OHSS	Ovarian Hyperstimulation Syndrome
PAID	Practical Application to Include Future Disease Costs

PGD	Pre-implantation Genetic Diagnosis
PGT/PGS	Preimplantation Genetic Testing/ Screening
PICO	Patients-Intervention-Comparator-Outcome
PTSD	Post-Traumatic Stress Disorder
QALE	Quality-Adjusted Life Expectancy
QALY	Quality-Adjusted Life Year
QoL	Quality of Life
WTP	Willingness-to-Pay

Introduction

The value of future pigeons

The Covid-19 pandemic was undoubtedly the most impactful event during my research period that illustrated how healthcare priority-setting often requires challenging and complex trade-offs. However, I thought it would be nice to start this thesis on a more positive note, by discussing some important news facts with regard to the Belgian pigeon population during this same period. Did you know that Belgium has the most pigeon fanciers per amount of inhabitants in the world and that in February 2020, the most expensive pigeon was sold for 1,6 million in Belgium?(1) This lucrative sector however also has its downsides as many pigeons get lost during training, which has led to the overpopulation of rural pigeons. These pigeons not only cause damage to buildings and agriculture, but they can also transmit diseases and pose public health risks. As a consequence, rural pigeons used to be captured and gasified until this practice was criticized by animal welfare organisations. In response, the city of Brussels has recently prohibited the killing of pigeons and many different Belgian cities such as Leuven and Mechelen started using contraceptives that are mixed in pigeon's food. Through investments of around €5000-€8000 per year, the pigeon population has been shown to decrease by 20% in one year and up to 80% in four years.(2,3)

These seemingly random facts reveal some fundamental value judgements made by policy-makers, about the value of (future) pigeons. First, pigeons can be valued according to different perspectives (pigeon fanciers vs a societal perspective) that clearly lead to different value estimations. Second, the investment in contraceptives shows that policymakers are willing to pay money to decrease the future pigeon population. The benefits of having less rural pigeons are considered to

outweigh the costs of the contraceptives program and the investment therefore offers sufficient value-for-money. In theory, it can be expected that the program that offers the highest decrease in pigeon population for the lowest cost, will be preferred by decision-makers. However, this would imply that a cheap or cruel elimination of the entire rural pigeon population would be considered the most cost-effective investment. In reality, this will not be the case as many other dimensions are also taken into account. For example, on the one hand, a too high mortality rate would lead to a natural imbalance, causing the attraction of exotic bird species that might lead to more problems. On the other hand, the refusal to gasify pigeons shows that we do not morally accept all possible ways in which the pigeon population can be decreased.

Unsurprisingly, this equation becomes far more complex when policymakers have to evaluate the cost-effectiveness of public programs that affect the wellbeing and the size of the present and future human population. Investments in for example climate change, traffic risk reductions or resource preservation not only affect people living today but also the number and wellbeing levels of people that will live in the future. When evaluating the cost-effectiveness of these programs, should an increase in population size be included as a benefit or as a loss? Theoretically, both options seem to imply problematic conclusions. If the coming into existence of future lives is taken into account as a benefit, this would imply that 'more' people are always considered better, even if they might experience a very low quality of life. On the other hand, attributing negative value to additional future lives would in its extreme require an elimination of the human race as every future live that is 'avoided' is considered a 'welfare gain'. In the field of population ethics, philosophers have studied how these different kind of repugnant conclusions might be avoided. At first sight, Jan Narveson proposed an intuitive perspective, arguing that, 'we are in favour of making people happy, but neutral about making happy people'. In this thesis, I explore the same type of dilemma's that arise when we try to define how future

lives should be taken into account in a specific *healthcare priority-setting context*. I cautiously try to open the Pandora's box of fundamental questions that arise and look for solutions in the population ethics literature. Is Narveson right and should healthcare budgets only be used to make people healthy but not to make healthy people?

Healthcare priority-setting and Health Technology Assessment

The past decades, the share of healthcare expenditure in most developed countries has steadily increased while government budgets remain limited. In Belgium, this share increased from 8% in the year 2000 up to 10,6% in 2019.(4) On the one hand, technological progress offers new, but often also more expensive, treatment opportunities. On the other hand, healthcare demand increases because of increasing life expectancy and higher health standards.(5) As a result of this increasing pressure on healthcare budgets, priority-setting has become inevitable. The focus of policymakers has therefore shifted from merely assessing the clinical effectiveness of healthcare programs to also evaluating their cost-effectiveness. At the beginning of the nineties, the first national advisory bodies (e.g. in Australia and Canada) started requiring cost-effectiveness assessments of new pharmaceuticals. Throughout the years, many countries followed and economic evaluations have become a common requirement for different types of health technologies (including medical equipment and public health programs).(6,7)

Whether a program offers sufficient value-for-money and is hence cost-effective, is assessed by comparing the costs and health effects generated by two or more alternative programs. These comparisons can however give rise to thorny ethical, social or political issues. Deciding which costs and effects should be taken into account in these comparisons will determine cost-effectiveness and subsequently priority-setting. For example, whether or not the health of a patient's caregivers or family members is taken into account might determine whether a treatment that

requires more or less intensive follow up is prioritized. The adopted methodology will therefore have important consequences, as resources spent on one intervention cannot be spent on another one. A transparent and systematic methodology is therefore required to ensure that priorities are set in a consistent and transparent manner. This is why *Health Technology Assessment* (HTA) guidelines have been developed to formalize the process of evaluating health technologies in a multidisciplinary way.(6,7)

As HTA is a relatively young field and often requires complex value judgements, some methodological questions still remain. One of the main points of ongoing discussions relate to the boundaries that define which costs and consequences should be included and the degree to which equity considerations should be taken into account. The consideration of for example future costs, family spillover effects or the distributional consequences of healthcare have received increasing attention in the past decades.(8–13) This thesis however aims to draw attention to a different fundamental question that has been typically overlooked in these discussions, but to which current guidelines do not provide an unambiguous answer: how should the effects of healthcare on future lives be considered in HTA?

Future life effects of healthcare

A large number of health interventions not only affects the patient undergoing treatment, but can also affect the patient's future children and even subsequent generations. This occurs when healthcare has an impact on the future quality-adjusted life expectancy of a patient's (prospective) offspring. Maternal vaccination programs can for example protect the unborn child against future diseases such as influenza or whooping cough.(14,15) Disease eradication programs and the prevention of mother-to-child transmission of infectious diseases such as HIV, or genetic diseases such as Cystic Fibrosis, can generate lifelong health gains for future children and their descendants.(16–18) More

generally, evidence increasingly shows that many different health factors such as lifestyle, drug use or mental health determine the health of people's future offspring. Interventions such as diabetes programs, smoking cessation programs, cancer treatments or mental health programs before and during pregnancy, in both male and female patients, are therefore all likely to have an impact on future offspring's health.(19–26) In addition, healthcare can also affect the number of people that will live in the future. Examples are fertility treatments, the prevention of miscarriages or stillbirths, abortion and contraceptives. Also the previously mentioned interventions related to various parental health determinants will typically affect patients' fertility. More indirectly, all lifesaving programs (e.g. heart surgery) can affect future population size, as the (fertile) people whose life is saved can have children in the future who could otherwise not have existed.

Guidelines are however unclear how these *future life effects* should be treated in HTA. This question is not only relevant because of the large number of interventions that generate future life effects, but also because of their potentially long-term impact on subsequent generations. Preventing the transmission of a genetic disease on a patient's future children, grandchildren and next generations can have increasing consequences in the future. Similarly, healthcare programs that change fertility will not only impact the number of offspring in the first generation but also the size of all successive generations. Hence, small changes in current health priority setting might lead to large future changes in the population dynamic, the epidemiology of particular (genetic) diseases and the health of future generations.(4) Whether or not these consequences are included in HTA is likely to have a substantial impact on cost-effectiveness and subsequently on budget allocation decisions. Furthermore, as health technologies are rapidly evolving in different fields such as genetic screening, assisted reproduction techniques and immunotherapy, the way in which future life effects are taken into account will play a decisive role in assessing whether the

higher costs related to these innovations can be justified by sufficient increases in generated health benefits.

The reason why guidance is currently lacking is that there is no obvious answer to the question how future lives should be valued in a priority-setting context in which both the number and wellbeing levels of future people are affected. As I will show throughout this thesis, many fundamental questions arise when broaching this subject and the adoption of a uniform perspective will inevitably require uncomfortable value judgements. However, I also illustrate that the current lack of guidelines leads to an inconsistent treatment of future life effects which makes study-outcomes incomparable and undermines the role of economic evaluation to facilitate decision-making. A coherent view on the value of future lives in HTA is therefore required to enable transparent and consistent priority-setting.

Objective and outline of the thesis

The aim of this PhD thesis is to contribute to the development of a standardized framework by exploring the relevance of 'future life effects' for cost-effectiveness calculations, to study the different ways in which these effects can be taken into account and to build bridges between HTA and other disciplines that study this matter.

The first chapter explores the practical difficulties related to future life effects in HTA and studies how future lives are currently considered in cost-effectiveness evidence of programs that have substantial effects on unborn, future lives. In the first part, a systematic review is conducted of the cost-effectiveness evidence related to perinatal mental health programs. As these interventions have been shown to have long-term effects on the *Quality of Life (QoL)* of both the parents and their future children, this domain provides an example of healthcare that 'makes people healthy'. In the second part, cost-effectiveness studies of *Assisted Reproductive Technologies* (ARTs) are reviewed. ARTs give rise to the

conception of ‘new’ children that would otherwise not have existed and hence illustrate the programs that, in essence, ‘make healthy people’.

The second chapter discusses relevant findings from the population ethics literature. The field of population ethics assesses the ethical issues that arise when our present-day decisions affect future people. In order to address these issues, different theories have been proposed on how to evaluate alternative scenarios in which the number of future lives and their wellbeing varies. (For example, different views can be adopted to assess whether it would be better that a group A of 1 million people will exist in the future, or that a group B of 2 million people will exist in the future in which people experience 50% of the wellbeing experienced by people in group A.) This literature hence offers fertile grounds for all fields in which trade-offs have to be made between the wellbeing and existence of present and future generations.(27–32) The first part of chapter two discusses how these alternative theories can be applied in an HTA context, and whether they would lead to intuitive conclusions. Furthermore, the consideration of future offspring effects also raises the question how their future costs should be considered. This question is further elaborated in the second part of this chapter.

The third chapter of this thesis explores the public opinion about the value of future lives in a healthcare priority-setting context by means of a Dutch, Flemish population survey. It is investigated if people take into account future life effects when setting healthcare priorities and whether this depends on the type of intervention effect (affecting birth/conception rates or the QoL of future people) or the level of abstractness of the affected future life. Based on respondent’s individual choice patterns, the public support for the previously discussed ethical theories is studied.

Finally, the findings of the previous chapters become practically relevant in chapter four. This chapter presents a long-term cost-utility analysis of *Non-Invasive Prenatal Testing* (NIPT) in which neonatal *Quality-*

Adjusted Life Years (QALYs) are considered according to four plausible perspectives. The time horizon of a recent Belgian study is extended, by means of a Markov model that follows the screened mother and future child throughout the rest of their lives. It is illustrated that each view implies a different treatment of neonatal costs and QALYs, and consequentially gives rise to diverging cost-effectiveness outcomes.

This thesis consists of a compilation of research articles that have been, or will be published as peer-reviewed papers or book chapters. They can therefore all be read as stand-alone text and might occasionally overlap. A full list of publications related to this dissertation can be found at the end of the thesis.

References

1. Belga. New Kim uit Berlaar is ‘duurste duif ooit in Europa’ na veiling voor 1,6 miljoen euro. De Standaard [Internet]. 2020 Nov 15 [cited 2022 Mar 25]; Available from: https://www.standaard.be/cnt/dmf20201115_94001084
2. ‘Duiven met anticonceptie aanpakken werkt’ [Internet]. De Standaard. [cited 2022 Mar 25]. Available from: https://www.standaard.be/cnt/dmf20171129_03214491
3. Aantal duiven in jaar tijd met vijfde gedaald [Internet]. De Standaard Mobile. [cited 2022 Mar 25]. Available from: https://m.standaard.be/cnt/dmf20220313_98176818
4. Current health expenditure (% of GDP) - Belgium | Data [Internet]. [cited 2022 Mar 22]. Available from: https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?locations=BE
5. Lefèvre M. Evaluatie van de duurzaamheid van het Belgische gezondheidssysteem op basis van projecties- Aanvulling bij ‘de performantie van het belgische gezondheidssysteem-rapport 2019’. 2021 p. 37. (KCE). Report No.: KCE report 341 A.
6. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
7. Neumann P, Sanders G, Russell L, Siegel J, Ganiats T. Cost-effectiveness in health and medicine. 2nd ed. New York: Oxford University Press; 2017. 496 p.
8. Kearns B. NICE DSU Report. The relevance of future, unrelated health costs in economic evaluation in NICE appraisals. [Internet]. Decision Support Unit, ScHARR, University of Sheffield; 2020 p. 48. Available from: http://nicedsu.org.uk/wp-content/uploads/2020/11/Future_unrelated_costs_Final.pdf
9. de Vries LM, van Baal PHM, Brouwer WBF. Future Costs in Cost-Effectiveness Analyses: Past, Present, Future. Pharmacoeconomics. 2019;37(2):119–30.
10. Brouwer WBF. The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. PharmacoEconomics. 2019 Apr 1;37(4):451–6.
11. McCabe C. Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. PharmacoEconomics. 2019 Apr 1;37(4):457–60.

12. Braithwaite RS, Roberts MS. Are Discount Rates Too High? Population Health and Intergenerational Equity. Med Decis Making. 2021 Feb 1;41(2):245–9.
13. Scarborough H, Bennett J. Estimating intergenerational distribution preferences. Ecological Economics. 2008 Jul;66(4):575–83.
14. Blommaert A, Bilcke J, Vandendijck Y, Hanquet G, Hens N, Beutels P. Cost-effectiveness of seasonal influenza vaccination in pregnant women, health care workers and persons with underlying illnesses in Belgium. Vaccine. 2014 Oct;32(46):6075–83.
15. Lugnér AK, van der Maas N, van Boven M, Mooi FR, de Melker HE. Cost-effectiveness of targeted vaccination to protect new-borns against pertussis: Comparing neonatal, maternal, and cocooning vaccination strategies. Vaccine. 2013 Nov 4;31(46):5392–7.
16. Thompson KM, Kalkowska DA. An Updated Economic Analysis of the Global Polio Eradication Initiative. Risk Analysis. 2021;41(2):393–406.
17. WHO | Mother-to-child transmission of HIV [Internet]. WHO. [cited 2021 Jun 8]. Available from: http://www.who.int/hiv/topics/mtct/about/en/
18. van den Akker-van Marle ME, Dankert HM, Verkerk PH, Dankert-Roelse JE. Cost-effectiveness of 4 Neonatal Screening Strategies for Cystic Fibrosis. PEDIATRICS. 2006 Sep 1;118(3):896–905.
19. Fraser A, Lawlor DA. Long-Term Health Outcomes in Offspring Born to Women with Diabetes in Pregnancy. Curr Diab Rep. 2014 Mar 25;14(5):489.
20. Hoek J, Koster MPH, Schoenmakers S, Willemsen SP, Koning AHJ, Steegers EAP, et al. Does the father matter? The association between the periconceptional paternal folate status and embryonic growth. Fertility and Sterility. 2019 Feb 1;111(2):270–9.
21. Dimofski P, Meyre D, Dreumont N, Leininger-Muller B. Consequences of Paternal Nutrition on Offspring Health and Disease. Nutrients. 2021 Aug;13(8):2818.
22. Beal MA, Yauk CL, Marchetti F. From sperm to offspring: Assessing the heritable genetic consequences of paternal smoking and potential public health impacts. Mutation Research/Reviews in Mutation Research. 2017 Jul 1;773:26–50.
23. Lo Presti A, Ruvolo G, Gancitano RA, Cittadini E. Ovarian function following radiation and chemotherapy for cancer. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2004 Apr 5;113:S33–40.

24. Meirow D, Nugent D. The effects of radiotherapy and chemotherapy on female reproduction. Human Reproduction Update. 2001;Vol. 7 No.6:535–43.
25. Nezvalová-Henriksen K, Spigset O, Nordeng H. Effects of ibuprofen, diclofenac, naproxen, and piroxicam on the course of pregnancy and pregnancy outcome: a prospective cohort study. BJOG: An International Journal of Obstetrics & Gynaecology. 2013;120(8):948–59.
26. Weinstock M. The potential influence of maternal stress hormones on development and mental health of the offspring. Brain, Behavior, and Immunity. 2005 Jul 1;19(4):296–308.
27. Arrhenius G. Population Ethics. The Challenge of Future Generations. (Unpublished version May 2016).
28. Greaves H. Population Axiology. Philosophy Compass. 2017;12(11):e12442.
29. Broome. The Wellbeing of Future Generations. In: The Oxford Handbook of Wellbeing and Public Policy. Oxford University Press; 2016. p. 901–28.
30. Gosseries A. Theories of intergenerational justice: a synopsis. SAPIENS Surveys and Perspectives Integrating Environment and Society. 2008 Nov 26;(1.1).
31. Parfit D. Future Generations: Further Problems. Philosophy & Public Affairs. 1982;11(2):113–72.
32. Kamaga K. Intergenerational Social Welfare Evaluation with Variable Population Size. In: Kamaga K, editor. Social Welfare Evaluation and Intergenerational Equity [Internet]. Singapore: Springer; 2020 [cited 2020 Nov 19]. p. 87–109. (SpringerBriefs in Economics). Available from: https://doi.org/10.1007/978-981-15-4254-1_5

Chapter 1: The consideration of future lives in cost-effectiveness studies

Part 1: Cost-effectiveness of mental health interventions during and after pregnancy: a systematic review

Abstract

Background: Mental health problems during and after pregnancy such as depression, anxiety, post-traumatic stress disorder (PTSD) or addiction are common and can have lifelong implications for both parents and offspring. This review investigates the cost-effectiveness of interventions tackling these problems, assesses the methodological quality of included studies and indicates suggestions for further research.

Methods: Thirteen databases were searched for economic evaluations of interventions related to ante-, peri- and postnatal mental health conditions, published between 2000 and September 2021, in high-income countries.

Results: Thirty-nine studies met all inclusion criteria. Interventions considered were screening programs, pharmacological treatments and various forms of psychosocial and psychological support. Six studies reported that the intervention was cost-saving. Eighteen were cost-effective and seven likely to be cost-effective. Only six studies included health outcomes for the child; one study considered paternal health. The time horizon for which costs and consequences were considered was for

most evaluations limited to one year (n=18) or two years (n= 11) postpartum.

Conclusion: Given the importance of the subject, a relatively low number of studies has investigated the cost-effectiveness of interventions tackling mental health problems during and after pregnancy. The scant evidence available suggests good overall value-for-money. Likely, cost-effectiveness is underestimated as the costly long-term consequences on offspring are systematically excluded. No evidence was found for several frequently occurring conditions. Further research is required to obtain reliable, long-term effectiveness data and to address the methodological challenges related to measuring all relevant health outcomes for all parties affected.

This subchapter has been published as:

Verbeke, E., Bogaerts, A., Nuyts, T., Crombag, N., Luyten, J. (2022). Cost-effectiveness of mental health interventions during and after pregnancy: A systematic review. Birth. doi: 10.1111/birt.12623

1.1.1 Introduction

Globally, an estimated 10% of pregnant and 13% of postnatal women experience a mental health disorder.(1) In the UK, the National Institute for Clinical Excellence (NICE) guidelines consider the following problems most relevant for antenatal and postnatal mental health: depression, anxiety disorders, eating disorders, drug- and alcohol-use disorders, severe mental illness (such as psychosis, bipolar disorder, schizophrenia) and post-traumatic stress disorder (PTSD) following traumatic birth experiences.(2) Depression is recognized as the primary mental health condition with recent studies reporting prevalence rates of postpartum depression in mothers from 12% (3) up to 14% (4). Prevalence estimates with regard to paternal depression approximate 10%.(5–8) However, other mental health disorders during and after pregnancy, both in mothers and fathers, are also common but often overlooked. A prevalence rate of anxiety disorders in pregnant or post-partum women of 13% was reported in a large US population-based study.(9) A Canadian study estimated PTSD after childbirth to prevail in 4-17% of pregnant or postpartum women.(10) Pregnancy can furthermore be a catalyst for the start or remission of eating disorders.(11) Addiction is also a relevant problem, specifically when considering the high correlation of drug- and alcohol-use disorders with other mental health conditions.

Mental health problems during and after pregnancy should be of particular concern to health policy makers because of the significant long-term health consequences on parents and their offspring. Parental mental illness not only increases the risk of adverse obstetric outcomes, it also affects parent-infant attachment and the cognitive, emotional, social and behavioural development of the child, as well as its biological systems.(12–17) Therefore, ante-, peri- and postnatal mental health conditions are expected to generate significant long-term costs to be borne by health systems later on. In the UK, Bauer et al. estimated the

additional total lifetime costs of mother and child for perinatal depression to equal £75.728 per case, of which 69% relates to the child (fathers were not considered in this study).(18) As a consequence, effective prevention and treatment at an early stage is likely to be economically beneficial. Consensus on the general need and effectiveness of treatment is reflected in current clinical guidelines (in e.g. the UK, US and Australia).(2,19,20)

In order to increase access to treatment within the available healthcare budget, policy-makers do not only require proof of effectiveness but also evidence on the *cost-effectiveness* of interventions. This means assessing whether the costs of an intervention are worth the generated health benefit. For a better understanding, Table 1 provides an overview of key concepts and the different types of health economic analyses that are discussed throughout this paper. Previous reviews have summarized available evidence on the cost-effectiveness of preventing or treating maternal depression and anxiety during pregnancy.(21,22) However, there is no available overview of the state of knowledge of the broader scope of perinatal mental health conditions for both mother and father, which is the aim of this systematic review.

Table 1: Types of economic evaluation and key concepts in health economics (25)

Cost-effectiveness assessment		Assessing whether the costs of an intervention are worth the generated health benefit.
Incremental cost-effectiveness ratio	ICER	The incremental costs of an intervention divided by the incremental health gain generated. The ICER represents the additional cost for one extra unit of health outcome and is typically used to assess cost-effectiveness.

Cost-utility analysis	CUA	The costs of an intervention are compared to the generated health outcomes in terms of *QALYs*. (The ICER is expressed in terms of costs per QALY.)
Quality-adjusted life year	QALY	A generic measure of disease burden representing the time (in years) in a certain health state, adjusted for the quality of life (QoL) experienced in this health state. 1 QALY represents one life-year in perfect health.
Quality of life (weights)	QoL	A preference based weight of a certain health state defined by two reference points: zero (= a state perceived equal to death) and 1 (= a state perceived equal to perfect health).
Cost-effectiveness analysis	CEA	The costs of an intervention are compared to the generated health outcomes in terms of relevant *natural units* to express treatment success. (The ICER is expressed in terms of costs per e.g. case of postpartum depression avoided.)
Cost-benefit analysis	CBA	The costs of an intervention are compared to the generated health outcomes in terms of *monetary units*. (Cost-effectiveness is calculated by considering the net benefit or the benefit-cost ratio of an intervention.)
Dominant intervention		The intervention is less costly and more effective than the alternative to which it is compared in the evaluation. (The intervention *dominates* the alternative.)

We aim to provide an up-to-date synthesis of current knowledge about the cost-effectiveness of prevention or treatment of mental health

conditions in the ante-, peri- and postnatal period. This objective was translated into specific inclusion criteria using the PICO framework (Table 2). A secondary objective was to review the methodological quality of available studies, with particular attention to how studies have dealt with the complicated nature of interventions linked to pregnancy and as a consequence, the potentially lifelong effects on offspring.

Table 2: PICO characteristics of the systematic review

Patients	pregnant women and fathers up to 24 months after delivery
Intervention	screening, prevention or treatment of depression, anxiety disorders, eating disorders, drug- and alcohol-use disorders and severe mental illness (such as psychosis, bipolar disorder, schizophrenia) and PTSD from traumatic birth
Comparator	alternative interventions, usual care, no intervention, or placebo
Outcome	partial or full economic evaluation

1.1.2 Methods

1.1.2.1 Search strategy and selection criteria

A systematic literature search was conducted in February 2021 and updated in September 2021, as prescribed by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.(23) Electronic searches were performed on Pubmed Central, Embase, Web of Science, APA PsychArticles (via ProQuest), CINAHL (via Ebsco), Cochrane, NHS EED, INAHTA, DARE, CADTH, HAS, PBAC and CEA registry. Search terms included words related to (1) economic evaluation, health technology assessment, cost-effectiveness, cost-benefit or cost-utility analysis; (2) mental health conditions related to depression, anxiety, eating disorders, drug-and alcohol-use disorders, psychosis, bipolar disorder, schizophrenia and PTSD and (3) mothers and fathers before, during and after pregnancy (appendix 1). The search was not restricted by language.

The PICO framework was used as a guide to select and assess studies.(24) Included studies were all health-economic evaluations of preventive (including screening) or curative interventions; for mothers and/or fathers during and up to 2 years after pregnancy; for the previously mentioned mental health conditions. The search was restricted to high-income countries, to ensure the comparison of evidence in similar health care contexts. Treatment outcomes were compared with no intervention, usual care, alternative interventions or placebo. Only studies published from 2000 onwards were considered in order to exclude interventions that are less relevant for today's decision-makers.

Exclusion criteria were: studies describing only outcomes or only costs, effectiveness studies or studies of which only posters were available. Gestational obesity was not considered a purely mental health condition as such and was therefore excluded if no related mental health condition was mentioned in the study description.

1.1.2.2 Data extraction and quality assessment

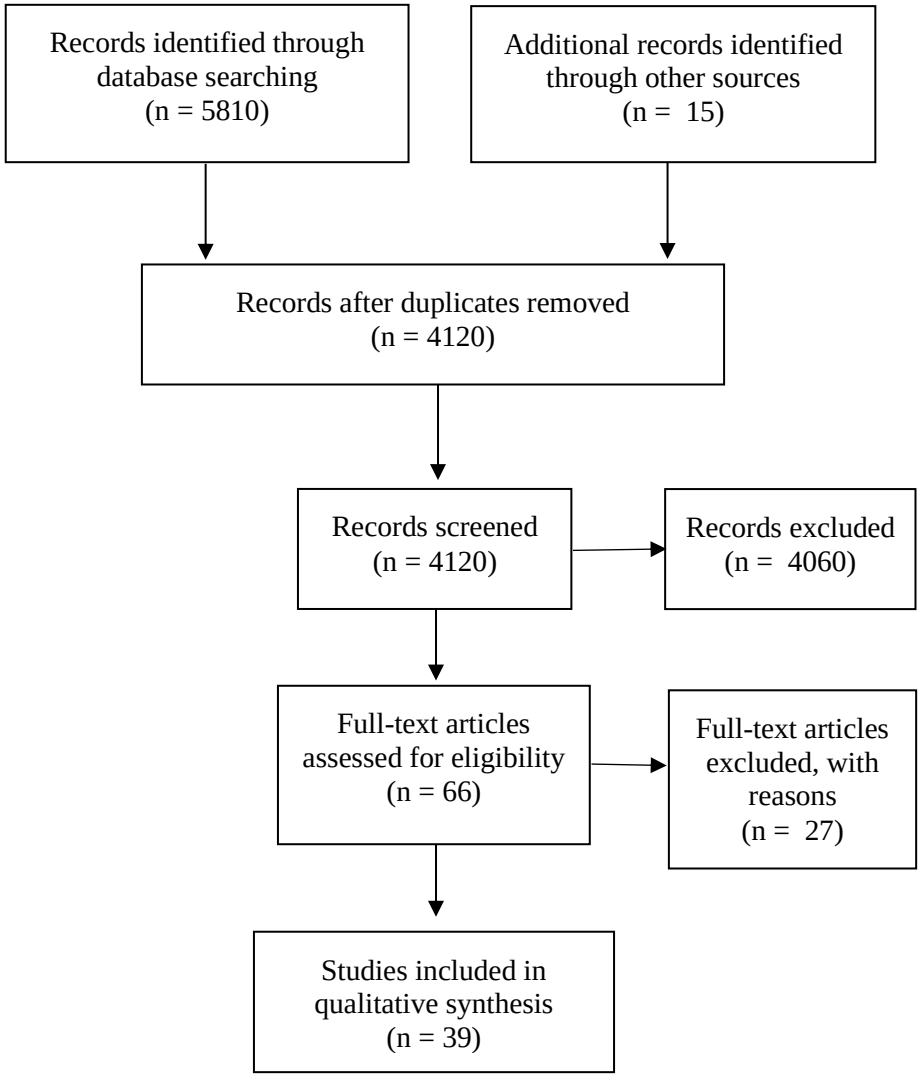
After abstract selection, 66 full articles were reviewed assessing eligibility for inclusion. A data collection form was constructed, summarizing the background and design of the studies, cost-effectiveness results, quality assessment and methodological limitations acknowledged by the study authors and the reviewers themselves. Studies excluded in the last review stage are listed in Appendix 2. Cost-effectiveness results were converted to 2019 euros (Appendix 3).

The methodological quality of the included studies was assessed by 2 authors (EV and JL), based on the 10-point checklist for assessing economic evaluations by Drummond et al..(25) Cases of disagreement on inclusion and quality assessment were resolved by discussion.

1.1.3 Results

1.1.3.1 Study characteristics

Figure 1: Study selection flow diagram



In total, 39 studies met the inclusion criteria. Table 3 (included at the end of this subchapter) describes the characteristics of each study. The evaluated interventions were related to: anxiety and depression (n=21 or

54%) (26–46), smoking cessation (n=12 or 31%) (47–58) and substance abuse (n=6 or 15%) (59–64). No results were found for eating disorders, drug-use disorders other than tobacco, severe mental illnesses (such as psychosis, bipolar disorder, schizophrenia) and PTSD from traumatic birth.

In terms of type of intervention, we identified four broad categories. *Psychosocial support* includes psycho-education, home visits, mentorship, financial incentives or supportive phone calls or text messages (n=14) (26,31,35,40,43,47,48,51–53,55–57,62). Psychosocial support is based on the social environment, while *psychological support* departs from psychological methods such as cognitive behavioural therapy and interpersonal therapy (n=7) (27,36,41,45,54,58,63).(65) Some programs combine both types of support (n=3) (39,49,59). Other categories are *pharmacological* interventions (treatment with medication) (n=3) (32,50,64), and *screening programs* (n=7) (28,30,34,38,45,46,60). Some interventions consisted of a mix of different categories (n=5). (29,33,37,42,61) The comparator in the studies was primarily usual care (n=27 or 69%), consisting of e.g. smoking cessation programs or perinatal care as prescribed by local guidelines.(26,27,29,31,33–43,45,47–52,55,59) The study population consisted of pregnant women (n=22 or 56%) (26,33,39,44–58,60–62,64), postpartum women (n=15 or 38%) (27,29–32,34–38,40–43,63) or both (n=1) (59), one study focused on postpartum fathers (28). No studies were found related to mental health in parents before or between pregnancies.

Studies were predominantly from the UK (n=16 or 41%) (29,34–39,41,45,47,48,50–52,56,63) and US (n=14 or 36%) (27,32,33,42,49,53–55,57–61,64).

Table 4 (included at the end of this subchapter) summarizes the design of the included studies. The type of decision-maker considered in an economic evaluation, determines the perspective adopted and hence, which costs and consequences should be included. For example, travel

costs might be relevant from a patient's perspective but not from the Ministry of Health's point of view.(25) The most commonly adopted perspective was that of a healthcare payer, considering only healthcare expenditure and excluding productivity losses or other economic costs (n=23) (26,27,30,32,33,35,37,40–42,44–49,52,53,55,57,59–61). Eight studies combined a healthcare payer perspective with a social services perspective (including e.g. home help costs), and one considered both the healthcare payer and broader patient costs (29). In one study these three perspectives were combined (43). A societal perspective, including productivity costs in addition to healthcare costs, was adopted in six studies (28,31,54,58,62,64). Effectiveness data included in the studies originated from Randomised Controlled Trials (RCTs) (26,27,31,33–35,39,45,48–50,52,53,55–57,63), clustered RCTs (36,43) or cohort studies (29,46,59). Seventeen studies were based on decision-analytic models populated with data from the scientific literature (28,30,32,37,38,40–42,44,47,51,54,58,60–62,64).

The type of economic analysis differed between studies. Fifteen evaluations considered health outcomes in terms of Quality-Adjusted Life Years (QALYs) in a cost-utility analysis (CUA) (26–28,30,32,34,38,40,41,44,45,51,55,58,61,64). QALYs are a generic measure of disease burden representing the time in a certain health state, adjusted for the quality of life (QoL) experienced in this health state.(25) Fifteen studies were identified as cost-effectiveness analyses (CEA) (29,31,33,35,36,39,46,47,49,50,53,54,56,57,62) as health outcomes were included in natural units. Seven studies reported incremental health gains both in QALYs (CUA) and natural units (CEA) (36,37,42,43,48,52,63). Lastly, two studies considered both costs and outcomes in monetary terms in a cost-benefit analysis (CBA) (59,60). In almost all studies (n=36) health outcomes for the mother were considered. Three studies did not include outcomes for the mother: Asper et al. considered screening for paternal depression(28), Pollack et al. only included the number of Sudden Infant Deaths (SIDS) due to gestational smoking (54) and Thanh et al. considered the number of Fetal Alcohol

Spectrum Disorder (FASD) cases in children (62). In total, 6 studies included health outcomes related to the offspring in terms of: numbers of FASD cases (60,62), number of SIDS averted (54) and adverse birth outcomes (51,58,64). Paternal health was once considered the main outcome in a study (28), and once included in the sensitivity analysis (32).

The time horizon considered was for most evaluations limited up to one year (n=18) or 2 years (+n=11) postpartum. Ten studies also considered a period longer than two years, of which five considered a patient's lifetime health. In case the time horizon exceeded one year, the studies reported a discount rate for costs, outcomes or both; only one study (60) did not report whether discounting was applied.

1.1.3.2 Critical appraisal

The quality of the included studies varied and recurring methodological challenges were identified. An overview of the quality assessment based on the 10-point checklist by Drummond et al. is included in Appendix 4. Two studies scored only three points (54,60), these studies did not provide sufficient information to assess whether reported results are conform guidelines. Nevertheless were these studies included because they were part of the scarce number of evaluations (n=3) that considered a lifetime time horizon for both women and offspring. All other studies scored at least five points or more, with a median score of 8.

The most common shortcoming of studies was that not all relevant costs and consequences were included. Besides the generally short time horizon considered, no study included health outcomes for mother, father and child altogether. Second, studies related to the same mental health condition expressed health effects in different outcome measures. Outcomes related to depression were for example considered in terms of the number of women no longer fulfilling diagnostic criteria (such as EPDS scale), number of depression free days, risk of depression

outcomes or SF-36 general health perception. This lack of uniformity between measures hampers cost-effectiveness comparisons, even for studies related to the same condition. Generally, QALYs are the preferred outcome-unit to improve comparability of results across disease categories. The QALY values in the included studies were however mainly based on secondary data, due to a lack of information on the QoL related to mental health during pregnancy for mothers, fathers and offspring. This introduces a significant risk of bias as specific aspects related to pregnancy or the mother-child-(father) relationship are not included in the evaluation. For example, four studies (30,37,38,42) included QALY values based on the QoL experienced during general depression while studying pregnancy related depression. Another challenge related to comparability of results is the difference between comparators. Even though the majority of studies (n=27) considered the same comparator: usual care, general practices still vary per country, potentially affecting the generalizability of results.(22) Last, while most (n=31) studies transparently discussed potential uncertainty of cost-effectiveness results, robust sensitivity analyses were lacking. Two studies did not report any sensitivity analysis (26,49), and six studies discussed uncertainty only to a limited extent (33,39,56,57,60,63). Because the lack of reliable (long-term) data (for all parties affected) was generally recognized, exploring the specific impact of uncertainty on outcomes is indispensable.

Discussing the limitations of the studies, authors reported recurring topics possibly leading to biased results. Most frequently mentioned were the consideration of a limited time horizon (29,31,34,39,42,51,52,61) and a limited perspective of the study which did not allow the inclusion of impacts on infants, family members or broader effects on society (29,30,33,34,38,39,44,60,62,63). Other risks of bias originate from self-reported results (31,33,45,48,54,59,63), low compliance or follow up rates (26,33,50,52,53,59) or the use of QALY values that are not specifically related to the assessed condition(30,32,37,38,40). Also non-randomized data (29,46,47,59),

missing data (47,59) and limited population size (29,45,55) were identified as factors possibly limiting reliability of clinical as well as health economic outcomes.

1.1.3.3 Synthesis cost-effectiveness results

Generally, cost-effectiveness is reported in terms of an *incremental cost-effectiveness ratio* (ICER), representing the incremental cost of the intervention for one extra unit of outcome (usually QALYs or specific effects such as depression-free-days achieved). A *cost-saving* intervention is less costly and more (or equally) effective compared to the alternative. In this case, the intervention *dominates* and should, at least from an economic perspective, be adopted.(25) Seven studies dominated (27,28,34,36,51,56,64), even though considerable statistical uncertainty was pointed out by two(64)(56). Alternatively, an intervention is cost-effective when the generated health gain is large enough to offset its additional costs, which was the case in sixteen studies (29,30,32,33,37,40,42,44,46,48,49,52,55,57,58,62). French et al. did not report an ICER but significant clinical improvements and a positive net benefit, indicating cost-effectiveness.(59) Cost-effective or cost-saving ICERs were found for psychological/social support (n=13), screening (n=4), pharmacological (n=2) or mixed (n=4) programs for depression (n=12), smoking cessation (n=8) and substance abuse (n=3), before (n=13) and after (n=10) birth. Seven studies reported potential cost-effectiveness, depending on the willingness to pay for the obtained health outcomes (31,39,47,53,54,60,63). Five studies emphasized uncertainty related to results and hence, more research is required (26,43,45,50,61). Only three studies reported that the intervention was not cost-effective, also indicating potential reasons for this conclusion. Morrell et al. did not find an additional health benefit generated by a community postnatal support program in addition to usual care. This could however depend on the measure of health, as 75% of the intervention group did indicate that the support was better than expected.(35) Paulden et al. reported a

lack of cost-effectiveness of screening for postnatal depression, this was mainly driven by the costs of managing false-positives, indicating the importance of accurate screening.(38) Last, Stevenson concluded that group CBT for postnatal depression does not appear to be cost-effective, but also here uncertainty and the need for further research was emphasized.(41) The cost-effectiveness results of the included studies are shown in Table 5, included at the end of this chapter.

1.1.4 Discussion

This systematic review demonstrated that the evidence related to the cost-effectiveness of mental health interventions during and up to 2 years after pregnancy is, given the importance of this subject, all-in-all limited. A total of 39 studies were identified, published between 2000 and September 2021, considering anxiety, depression, smoking or substance abuse. The majority of economic evaluations reported good value for money, even though results tend to be uncertain due to a lack of reliable data and difficulties of appropriately measuring all relevant health outcomes. These findings largely reflect challenges related to the underlying evidence base on effectiveness in the first place. Before the question of cost-effectiveness can be answered, more evidence is needed regarding the effectiveness of mental health programs in aspiring or young parents.

There were important gaps in the (cost-) effectiveness literature. First, while the evidence on perinatal depression is substantial, research on the broader range of mental health disorders is lacking even though the prevalence of these conditions during pregnancy is significant.(9,10,66–68) Second, no studies related to preconception mental health were found. While this can be a delicate topic because of e.g. stigmatising attitudes that can worsen mental disorders, preconception care offers an important window of opportunity to generate long-term health benefits.(68–71) Third, curative interventions that exclusively target

paternal mental health are scarce and as a consequence, so is rigorous research on this topic.(72–74) Fourth, the evidence on pharmaceutical interventions, including possible risks during pregnancy and lactation is poor.(75–77) Furthermore, no study considered the co-occurrence of mental health conditions in couples as well as co-morbidities and the implications on (cost-) effectiveness results. Incidence rates of paternal postpartum depression are however positively correlated with maternal depression, increasing from 1.2%-25.5% to 24%-50% when the mother experiences PND too.(7,8) Also, significant relations between schizophrenia and alcohol-use during pregnancy have been identified, but also smoking cessation, PTSD, anxiety, bipolar disorder and eating disorders have been associated with depression.(4,67,78–83) These conditions cannot be treated in isolation from each other and neither can cost-effectiveness be assessed. Finally, the importance of entangled risk factors such as the parent-infant relationship (e.g. hostile behaviour or domestic violence) is often overlooked. As these factors can be crucial for the behavioural outcomes of the child in the long-term, additional parenting support might be required to ensure (cost-)effective results.(84–87)

Regarding the studies that were available, it is noteworthy that none of them considered health outcomes for mother, child and father altogether, despite the potential long-term effects for each. Including *all* lifelong consequences is likely to increase cost-effectiveness significantly, but requires long-term follow-up data for all parties affected.(88) Furthermore, health outcomes need to be captured in uniform outcome measures that allow meaningful comparisons of study results. For example, considering the incremental change in number of depression-free days, does not allow inclusion of broader consequences such as future behavioural or cognitive problems of the child. In order to allow appropriate use of QALY values, further research is required to determine the QoL of mothers, children and fathers in the specific context of mental health during and after pregnancy.

Our recommendations for further research can be summarized as follows. First, access to reliable data on the costs and effectiveness of mental health interventions in the long term, for both mother, father and child need to be improved. Secondly, more research is required to identify appropriate measures of health that can capture all relevant health consequences of mental health interventions, for all parties affected. With regard to QALYs, QoL-values need to be determined for the specific context of mental health in the ante-, peri- and postnatal period. Finally, a holistic approach is required that allows the consideration of related disorders and risk factors to increase accuracy of results.

This review has several strengths and limitations. To our knowledge, this is the first review that provides a comprehensive overview of the different health domains relevant for budget allocation decisions related to perinatal mental health care. Studies were identified based on a systematic process and were assessed according to the quality criteria proposed by Drummond et al.. This assessment was used as a basis to identify fundamental methodological challenges of the included studies and to identify recommendations for future research.

However, despite our systematic search, some studies might not be included because they were not categorized under the searched mesh terms or did not include the specified search terms in the title or abstract. Second, while cost-effectiveness results were expressed in 2019 euros, cost data are health–system specific and study results should be compared and interpreted with caution.(89) Last, the time horizon of the literature search was limited to 2000-2021, entailing the risk that certain pioneering studies (at a methodological level) were omitted.

1.1.5 Conclusion

Mental health conditions are common during and after pregnancy, including long-term health consequences for mother, father and child. While guidelines generally recommend prevention and treatment of mental health conditions during this period, in many domains access to evidence-based care remains limited. This systematic review brings together cost-effectiveness evidence related to interventions targeting a broad range of mental health conditions during and after pregnancy. Overall, given the importance of this subject, there were relatively few studies available. The majority of interventions was found to be cost-effective in these studies. But these studies mostly illustrate a need for further research due to limited reliable long-term effectiveness data, or methodological challenges related to measuring all relevant health outcomes. Because of these challenges, it is likely that existing results systematically underestimate real world cost-effectiveness, as long-term costs of suboptimal child development in the first years of life will be substantial.

Systematic review tables

Table 3: Study characteristics

Ref	Author	Year	Title	Population	Coun-try	Intervention	Time	Condition	Type of intervention	Com-parator
ANXIETY & DEPRESSION										
26	Turkstra et al.	2016	An economic evaluation alongside a randomised controlled trial on psycho-education counselling intervention offered by midwives to address women's fear of childbirth in Australia	Women with high scores on childbirth fear n=184 (91 intervention), second trimester of pregnancy, English and 16 years or older	Aus	A midwife-led telephone psycho-education intervention for women fearful of birth (BELIEF)	antenatal	anxiety	psychosocial support	usual care
32	Eldar-Lissai et al.	2020	Cost-Effectiveness of Brexanolone Versus Selective Serotonin Reuptake Inhibitors for the Treatment	Mothers with moderate to severe PPD, on average 16 weeks postpartum and age 28 (similar to BRX clinical trial patients)	US	Brexanolone injection for postpartum depression	postnatal	depression	Pharmaco-logical	treatment with selective serotonin reuptake inhibitors (ssris) for PPD

			of Postpartum Depression in the United States							
27	Ammerman et al.	2017	Cost-effectiveness of In-Home Cognitive Behavioral Therapy for low-income depressed mothers participating in early childhood prevention programs	Low-income mothers enrolled in a home visiting program and diagnosed with Major depressive disorder (MDD) (n=93)	US	In-Home Cognitive Behavioral Therapy (IH-CBT)	postnatal	depression	psychological support	usual care
36	Morrell et al.	2009	Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PoNDER trial	Women registered with participating GP practices who became 36 weeks pregnant during the recruitment phase of the trial, had a live baby and were on a collaborating HV's caseload	UK	HV training in the assessment of postnatal women, combined with either cognitive behavioural approach (CBA) or person-centred approach (PCA) sessions for eligible women, plus the option	postnatal	depression	psychological support	usual care

				for 4 months postnatal. 103 clusters in 29 primary care trusts, n= 4084		of a selective serotonin reuptake inhibitor if indicated.				
41	Steven-son et al.	2010	The Cost-Effectiveness of Group Cognitive Behavioral Therapy Compared with Routine Primary Care for Women with Postnatal Depression in the UK	women with postnatal depression	UK	group cognitive behaviour therapy (1 session/week for 8 weeks, of 2-hours, in groups of 4-6)	postnatal	depression	psychological support	usual care
45	Trevilli on et al.	2020	An exploratory parallel-group randomised controlled trial of antenatal Guided Self-Help (plus usual care) versus usual care alone for pregnant women with	Pregnant women older than 16, meeting criteria for DSM-IV depression on the Structured Clinical Interview (n=53, 26 intervention)	UK	Guided self-help (GSH) modified for pregnancy plus usual care.	antenatal	depression	psychological support	usual care

			depression: DAWN trial							
33	Grote et al.	2017	Incremental Benefit-Cost of MOMCare: Collaborative Care for Perinatal Depression Among Economically Disadvantaged Women	Socioeconomically disadvantaged women with antenatal depression with and without comorbid PTSD (n=164) (women at 12-32 weeks gestation scoring 10 or higher on the PHQ-9 or with a diagnosis of probable dysthymia)	US	A multicomponent collaborative care intervention: including evidence-based depression treatment and active measurement of outcomes and follow-up according to stepped-care principles	antenatal	depression	psychological support & pharmacological	more intensive version of usual care (= mss-plus)
39	Petrou et al.	2006	Cost-effectiveness of a preventive counselling and support package for postnatal depression	Women at high risk of developing postnatal depression: predictive index score ≥ 24, 26-28 weeks of gestation (n= 151, 74 intervention)	UK	Counselling and specific support for mother-infant relationship: visit 35 and 37 weeks antenatally to establish supportive relationship, then visits on	antenatal	depression	psychological/ psychosocial support	usual care

						days 3,7 and 17 after delivery and then weekly up to 8 weeks				
29	Boath et al.	2003	When the cradle falls II: the cost-effectiveness of treating postnatal depression in a psychiatric day hospital compared with routine primary care	Women with postnatal depression (n=60, 30 intervention)	UK	A specialist psychiatric Parent and Baby Day Unit (PBDU): individual, high intensity, customized treatment, existing of: counselling, group therapy, creative therapy, hobbies and activities, stress management, assertiveness training, yoga and relaxation, a group for parents and older children and pharmacothera-py	postnatal	depression	psychological/ psychosocial support + pharmacologi cal	usual care

63	Barlow et al.	2019	A randomized controlled trial and economic evaluation of the Parents Under Pressure program for parents in substance abuse treatment	Parents receiving treatment for a drug or alcohol problem and are a primary carer of a child under the age of 2.5 years. (n= 100, 48 intervention)	UK	intensive one-to-one parenting program (Parents under Pressure, PuP) to reduce child abuse potential by enhancing parental emotional regulation	postnatal	substance abuse	psychological support	usual care
31	Dukhov-ny et al.	2013	Prospective Economic Evaluation of a Peer Support Intervention for Prevention of Postpartum Depression among High-Risk Women in Ontario, Canada	Postpartum woman from seven health regions across Ontario, Canada (n=610)	Can	A volunteer telephone-based peer support intervention for prevention of PPD	postnatal	depression	psychosocial support	usual care
35	Morrell et al.	2000	Costs and effectiveness of community postnatal support workers: randomised	Postnatal women aged 17 or over (n=493)	UK	Up to 10 home visits in the first postnatal month of up to 3h duration by a community	postnatal	depression	psychosocial support	usual care

			controlled trial			postnatal support worker				
40	Saing et al.	2018	Cost Effectiveness of a Community-Delivered Consultation to Improve Infant Sleep Problems and Maternal Well-Being	Mothers and infants (0-12 months)	Aus	A community delivered consultation aimed at improving infant sleep and maternal well-being	postnatal	depression	psychosocial support	usual care
28	Asper et al.	2018	Screening fathers for postpartum depression can be cost-effective: An example from Sweden	Postpartum fathers, initial study: a questionnaire was sent to 8,011 fathers of whom 3,656 (46%) responded.	Swe-den	Postpartum depression screening (EPDS screening) for fathers (1000 iterations)	postnatal	depression	screening	no screening
30	Camp-bell et al.	2008	Screening for postnatal depression within the Well Child Tamariki Ora Framework	New mothers in New Zealand who have given birth in any 12 month period (regardless of number or previous children) (n= 56635)	New Zealan d	Screening programme for post-natal depression	postnatal	depression	screening	no screening

46	Cham-bers et al.	2021	The clinical performance and cost-effectiveness of two psychosocial assessment models in maternity care: The Perinatal Integrated Psychosocial Assessment study	women attending their first antenatal visits (n= 3673 usual care, n= 3132 PIPA model)	AUS	Perinatal Integrated Psychosocial Assessment (PIPA)	perinatal	depression	screening	usual care
34	Hender son et al.	2018	Cost-effectiveness of PoNDER health visitor training for mothers at lower risk of depression: findings on prevention of postnatal depression from a cluster-randomised controlled trial	Mothers with lower-risk status at 6-weeks postnatal (n=1459)	UK	Health visitor training to assess postnatal depression	postnatal	depression	screening	usual care

38	Paul-den et al.	2009	Screening for postnatal depression in primary care: cost effectiveness analysis	A hypothetical population of women assessed for postnatal depression either via routine care only or supplemented by use of formal identification methods six weeks postnatally	UK	Alternative screening methods of postnatal depression in primary care (might detect women that are not detected by routine care but also incorrectly identify women who were not depressed	postnatal	depression	screening	usual care
44	Premji et al.	2021	Maximizing maternal health and value for money in postpartum depression screening: a cost-effectiveness analysis using the All Our Families cohort and administrative data in Alberta, Canada	Women during second trimester of pregnancy (n=2698, 87% screened)	Can	Postpartum depression screening	postnatal	depression	screening	no screening

42	Wilkin-son et al.	2017	Screening for and Treating Postpartum Depression and Psychosis: A Cost-Effectiveness Analysis	Hypothetical cohort of 1000 pregnant women experiencing one live birth over a 2-year time horizon	US	Physicians screening for and treating postpartum depression and psychosis in partnership with a psychiatrist	postnatal	depression	screening + psychological support	usual care
37	NCCMH	2014	Case identification and assessment, psychological and psychosocial interventions for the prevention or treatment of mental health problems. In Antenatal and postnatal mental health: the NICE guideline on clinical management and service guidance (update).	Women with subthreshold/mild to moderate depression in the postnatal period	UK	Different types of psychological and psychosocial interventions: facilitated self-help or listening visits	postnatal	depression	screening + psychological/ psychosocial support	usual care

43	Ride et al.	2016	Preventing postnatal maternal mental health problems using a psychoeducational intervention: the cost-effectiveness of What Were We Thinking	English-speaking first time mothers attending participating Maternal and Child Health Centres 6 months post partum(n= 362, 184 intervention)	Aus	What Were We Thinking, a psychoeducational intervention targeted at the partner relationship, management of infant behaviour and parental fatigue	postnatal	depression, anxiety, and adjustment disorders	psychosocial support	usual care
SMOKING CESSATION										
50	Essex et al.	2014	Cost-Effectiveness of Nicotine Patches for Smoking Cessation in Pregnancy: A Placebo Randomized Controlled Trial (SNAP)	Heavy smoking pregnant women(n=1050)	UK	Nicotine patches for smoking cessation	antenatal	smoking cessation	pharmaco-logical	usual care

54	Pollack et al.	2001	Sudden Infant Death Syndrome, Maternal Smoking During Pregnancy, and the Cost-Effectiveness of Smoking Cessation Intervention	Birth cohort 1995 (self-reported smoking status mothers)	US	Prototypical smoking cessation programs (field of psychological support)	antenatal	smoking cessation	psychological support	no intervention
49	Dornelas et al.	2006	Efficacy and cost-effectiveness of a clinic-based counseling intervention tested in an ethnically diverse sample of pregnant smokers	Low income, predominantly Hispanic, pregnant patients in an urban prenatal clinic, ≤ 30 weeks gestation, ≥ 18 years old (n=105, 53 intervention)	US	1.5h counselling plus telephone follow-up delivered by a masters prepared mental health counselor (bi-monthly during pregnancy and monthly after delivery)	antenatal	smoking cessation	psychological/ psychosocial support	usual care

58	Barchel ler et al.	2021	Behavioral Smoking Cessation Counselling During Pregnancy A Cost-Effectiveness Analysis	theoretical cohort of n=285,000 women	US	Behavioral smoking cessation counselling	perinatal	smoking cessation	psychological	Usual care
47	Bell et al.	2017	Evaluation of a complex healthcare intervention to increase smoking cessation in pregnant women: interrupted time series analysis with economic evaluation	n=10594 mothers smoking during pregnancy	UK	A package of measures implemented in trusts and smoking cessation services, comprising skills training for healthcare and smoking cessation staff; universal carbon monoxide monitoring with routine opt-out referral for smoking cessation support provision of carbon monoxide monitors and	antenatal	smoking cessation	psychosocial support	usual care

						supporting materials; and an explicit referral pathway and follow-up protocol.				
48	Boyd et al.	2016	Are financial incentives cost-effective to support smoking cessation during pregnancy?	Pregnant women (n=612), Markov model=1000 women with mean age= 28y	UK	Financial incentives for smoking cessation in pregnancy	antenatal	smoking cessation	psychosocial support	usual care
51	Jones et al.	2019	A dynamic, modifiable model for estimating cost-effectiveness of smoking cessation interventions in pregnancy: application to an RCT of self-help delivered by text message	Hypothetical cohort of 1000 singleton-pregnancy women who smoke	UK	12-week programme of tailored text messages	antenatal	smoking cessation	psychosocial support	usual care

57	Mundt et al.	2021	Cost-effectiveness of stop smoking incentives for Medicaid-enrolled pregnant women.	Medicaid-enrolled pregnant smoking women, mean gestation at enrolment = 15 weeks (n=1014 of which incentive group n=505)	US	Financial incentives in the form of gift cards	peri-, postnatal	smoking cessation	psychosocial support	lower incentive scheme
52	Naugh-ton et al.	2017	Large multi-centre pilot randomized controlled trial testing a low-cost, tailored, self-help smoking cessation text message intervention for pregnant smokers (MiQuit)	<25 weeks gestation, smoked at least 1 daily cigarette, able to receive and understand English SMS texts (n=407, n= 203 intervention)	UK	12-week programme of individually tailored, automated, interactive, self-help smoking cessation text messages	antenatal	smoking cessation	psychosocial support	usual care

53	Parker et al.	2006	Feasibility, cost, and cost-effectiveness of a telephone-based motivational intervention for underserved pregnant smokers	Women who have smoked at least one puff of a cigarette within the past 30 days, < 26 weeks pregnant, have access to a telephone and speak English or Spanish, n=1065 randomized between 3 experimental groups (n=358 intervention)	US	A proactively provided telephone-based motivational smoking cessation intervention	antenatal	smoking cessation	psychosocial support	group 1: self help quit kit, group 2: self help quit kit + monetary incentive lottery
55	Ruger et al.	2008	Cost-effectiveness of motivational interviewing for smoking cessation and relapse prevention among low-income pregnant women: A randomized controlled trial	Low-income pregnant women recruited from multiple obstetrical sites in the Boston metropolitan area (n=302) 2 groups: current smokers (smoking cessation: SC) and recent quitter within 3	US	Motivational interviewing (IM): 3 home visits (1h) client-centred technique exploring perceptions and concerns about smoking, clarifies conflicting motivations, focuses on the social context in which women	antenatal	smoking cessation	psychosocial support	usual care

				months (relapse prevention: RP)		live, and provides support and skills of training				
56	Ussher et al.	2015	The London Exercise And Pregnant smokers (LEAP) trial: a randomised controlled trial of physical activity for smoking cessation in pregnancy with an economic evaluation	Women aged 16–50 years, between 10 and 24 weeks' gestation, currently smoking at least one cigarette per day, were smoking at least five cigarettes per day before pregnancy, prepared to quit smoking 1 week after enrolment and they could confirm that they were able to walk continuously for at least 15 minutes. (n=785)	UK	Physical activity + behavioural support (moderate-intensity exercise was prescribed according to age and current activity levels)	antenatal	smoking cessation	psychosocial support	behavioural support only

SUBSTANCE ABUSE										
61	Premkumar et al.	2019	Methadone, Buprenorphine, or Detoxification for Management of Perinatal Opioid Use Disorder (Detoxification = medically supervised withdrawal over 5 days to 16 weeks with medications such as buprenorphine or clonidine)	Women with OUD after 16 weeks of pregnancy (100.000 simulations)	US	Methadone, buprenorphine, or detoxification treatment for the management of opioid use disorder (OUD) during pregnancy.	antenatal	substance abuse	Pharmaco-logical and psychological support	methadone, buprenorphine, or detoxification treatment
64	Robin et al.	2021	Cost-effectiveness of buprenorphine vs. methadone for pregnant people with opioid use disorder	Theoretical cohort of n=22,400 pregnant women	US	Buprenorphine	Perinatal	substance abuse (opioid use disorder)	Pharmaco-logical	Methad-one

63	Barlow et al.	2019	A randomized controlled trial and economic evaluation of the Parents Under Pressure program for parents in substance abuse treatment	Parents receiving treatment for a drug or alcohol problem (opioid replacement treatment, relapse prevention, counseling) and were a primary carer of a child under the age of 2.5 years. (n= 100, 48 intervention)	UK	Intensive one-to-one parenting program (Parents under Pressure, PuP) with the goal of reducing child abuse potential by enhancing parental emotional regulation	postnatal	substance abuse	psychological support	usual care
59	French et al.	2002	Benefit-cost analysis of addiction treatment in Arkansas: Specialty and standard residential programs for pregnant and parenting women	Pregnant and parenting substance abusers, (most women entered these programs as self or criminal justice referrals.) (n= 85, 44 intervention)	US	Speciality residential treatment: a comprehensive set of "wrap around" services on site (up to 12 months stay)	ante/postnatal	substance abuse	psychological/ psychosocial support	usual care
62	Thanh et al.	2014	An Economic Evaluation of the Parent-Child Assistance	Women abusing substances and are pregnant up to 6 months postpartum.	Can	Parent-Child Assistance Program: 3y home visitation/harm	antenatal	substance abuse	psychosocial support	no intervention

			Program for Preventing Fetal Alcohol Spectrum Disorder in Alberta, Canada	(n=366, of which 161 alcohol abuse)		reduction intervention to prevent alcohol exposed births (thereby births with fetal alcohol spectrum disorder)				
60	Gifford et al.	2010	Assessment of Benefits of a Universal Screen for Maternal Alcohol Use during Pregnancy	Pregnant women	US	Universal meconium screening for maternal drinking during pregnancy (combined with 3 possible interventions)	antenatal	substance abuse	screening	no interventi on

Table 4: Design of included studies

Ref	Author	Type of evaluation	Data-source	Persp-ective	Health outcomes included	Health outcomes included for: mother	child	father	Costs included	Time horizon	Curr-ency	Disc rate
ANXIETY & DEPRESSION												
26	Turkstra et al.	CUA	RCT	healthcare	health-related quality of life baseline and 6 weeks post-partum (EQ-5D-3L)	v			self-reported visits to GP, midwives, obstetricians, nurse, home visits, ultrasound scans, hospital emergency department visits, hospital admissions (pre- and post-birth), special care nursery and mode of birth.	baseline until 6 weeks post-partum	AUSD 2013	
32	Eldar-Lissai et al.	CUA	model	healthcare (third-party payer perspective)	QALYs	v	v (0-4y hospitalization but no utility loss, 5-12y behavioral issues, 13-18y	v (in sensitivity analysis: risk of major depressive disorder)	treatment costs + direct medical costs (from literature)	11y base case (model up to 18y)	$ 2018	3%

							depression)					
27	Ammerman et al.	CUA	RCT + model	healthcare	QALYs	v			treatment costs + medical care	3y	$ 2013	3%
36	Morrell et al.	CEA	cluster RCT/ model	healthcare + social services	proportion of at-risk women with a 6-month Edinburgh Postnatal Depression Scale (EPDS) score ≥ 12, QALY (SF-6D at 6 weeks, 6, 12 and 18 months)	v			health visitor training costs, primary analysis: costs of mother at 6 months, further analysis: mother and baby costs at 12 months	6 months	£ 2003-2004	
41	Stevenson et al.	CUA	model	healthcare	QALYs (mapped EPDS scores)	v			costs of treatment (health worker*time)	12 months	£ 2007-2008	
45	Trevillion et al.	CUA	RCT	healthcare	QALYs based on SF-6D at baseline, 14-week post randomisation and 3 months post delivery	v			treatment costs, health and social care costs	3 months post delivery	£ 2015-2016	

33	Grote et al.	CEA	RCT	healthcare	depression-free days (DFDs)	v			treatment costs + mental health services costs directly related to depression treatment	18 months	$ 2013	no
39	Petrou et al.	CEA	RCT	healthcare + social services	duration of postnatal depression, SCID-II assessment 8 weeks, 18 weeks, 12 months and 18 months postpartum	v			intervention costs, all health and social care services (via interviewing women and diaries care providers - coupled to prices in literature) also for child	18 months (after delivery)	£ 2000	cost s: 6%, effe cts 1,5 %
29	Boath et al.	CEA	prospect ive cohort study	healthcare + broader patient costs	women no longer fulfilling Research Diagnostic Criteria for major or minor depressive disorder after 6 months	v			medication, cost of transport, child care, opportunity costs of women: loss of employment, house work, leisure time	6 months	£ 1992-1993	6%
63	Barlow et al.	CEA/ CUA	RCT	healthcare + social services	child abuse potential, parental emotional regulation, QALYs	v		(v)	program costs, hospital, community health and social services , legal services and	12 months	£ 2016	

					(baseline, 6 and 12 months) (if both parents had alcohol or drug problem primary caregiver was assessed)				costs borne by parents			
31	Dukhovny et al.	CEA	RCT + model	societal	Edinburgh Postnatal Depression Scale (EPDS) score >12 at 12weeks postpartum	v			direct medical and program costs to the healthcare system, costs absorbed by all of the stakeholders, including child care and household help, missed work, and the opportunity cost of volunteer time	first 12 weeks postpartum	CAD 2011	
35	Morell et al.	CEA	RCT	healthcare	short form-36 (SF-36) general health perception domain measured at 6 weeks (costs and outcomes compared at 6 weeks and 6 months after delivery)	v			Costs of visits, GP contacts and prescriptions, hospital contacts => medical care	6 months	£ 1996	

40	Saing et al.	CUA	model	healthcare	EPDS scores mapped to published utility scores -> QALYs, interruption free-nights	v			consultation, training, residential stay and health care resource costs	16 months	AUSD 2014-2015	5% for QALYs
28	Asper et al.	CUA	model	societal	QALYs (health utilities related to depression for fathers)			v	screening costs, direct and indirect costs related to depression (anti-depressants, productivity losses	lifetime	€ 2016	3%
30	Campbell et al.	CUA	model	healthcare	(1) the number of mothers with resolved PND (who are not depressed at endpoint), (2) the number of PND cases detected, and (3) maternal Quality Adjusted Life Years (QALY). (screening 6 weeks and 4 months postpartum)	v			annual cost of implementing a routine screening programme for PND: social support, psychological therapy, combination of antidepressants and psychological therapy	12 months	New Zealand dollars 2006-2007	

46	Chambers et al.	CEA	Prospective cohort study	healthcare	rate of true positives and false positives (after additional screening by midwife)	v			time of screening midwives, staff and clinicians	12 months	AUSD 2017	
34	Henderson et al.	CUA	RCT + model	healthcare + social services	QALYs + risk-of-depression outcomes at follow-up as a secondary outcome	v			HV training, ongoing clinical supervision; HV contacts; infant immunisations; GP contacts; prescriptions for all conditions; social worker contacts; admissions to Mother and Baby psychiatric units and other mental health contacts	6 months	£2003 -2004	
38	Paulden et al.	CUA	model based on systematic review	healthcare + social services	QALYs	v			cost of screening, subsequent treatment and incorrect diagnosis (national reference costs	1 year	£2006 -2007	

44	Premji et al.	CUA	model	Healthcare	QALYs sf-6d	v			screening, treatment and healthcare costs	2y	CAD 2019	
42	Wilkinson et al.	CEA/ CUA	model	healthcare	number of remissions and QALYs	v			treatment costs and healthcare costs related to depression	2y	$ 2014	3%
37	NCCMH	CEA/CUA	model (based on guideline meta analyses)	healthcare	1) Number of women who improved and did not relapse at the end of 1-year follow-up 2) Number of quality adjusted life years (QALYs) gained at the end of 1-year follow-up.	v			treatment costs, health and social care costs for mother–infant dyad. (deterministic costing)	1y after initiation treatment	£ 2013-2014	
43	Ride et al.	CEA/CUA	clustered RCT + model	healthcare + social services + patient costs	the 30-day prevalence of depression, anxiety and adjustment disorders, and quality-adjusted life years (QALYs)	v			healthcare, early childhood and social service costs + participant's out-of-pocket costs	6 months	AUSD 2013-2014	

SMOKING CESSATION												
50	Essex et al.	CEA	RCT	healthcare + social services	biochemically validated smoking cessation	v			treatment costs+ medical care	7 months	£ 2009-2010	
54	Pollack et al.	CEA	model	societal	SIDS deaths averted, life-years saved		v			1 year (number of deaths+ extra-polation lifetime for life-years)	$ 1998	5%
49	Dornelas et al.	CEA	RCT	healthcare	self-reported smoking abstinence for the previous 7 days (at end of pregnancy and 6 months post-partum) - confirmed by a carbon monoxide reading	v			costs of training mental health counsellors and healthcare providers, counselling time, telephone time, clerical staff time	6 months post-partum	$ 2002	

58	Bachelle r et al.	CUA	model	societal	maternal and neonatal outcomes + QALYs	v	v		counselling costs + long-term outcome costs (stillbirth, smoking, preterm delivery)	Lifetime	$ 2020	3%
47	Bell et al.	CEA	model	healthcare	probability of quitting smoking during pregnancy	v			training of staff, investment in equipment, consumables and changes in workload	5y (for costs, data only 4 months after inter-vention)	£ 2013	1.5 %
48	Boyd et al.	CEA/ CUA	RCT/ model	healthcare	number of quitters/ QALYs	v			direct costs to NHS/ model also includes LT costs: post-birth hospitalisation costs + LT cost of treating smoking related diseases	34-38 weeks pregnanc y / model= lifetime, consideri ng relapse up to 8y post quit	£ 2013	3,50 %

51	Jones et al.	CUA	Economics of Smoking in Pregnancy (ESIP) model, estimates the life-time cost-effectiveness of smoking cessation interventions in pregnancy applied to an RCT	healthcare + social services	costs of treating disease burdens, adverse birth outcomes, life-years and QALYs	v	v		antenatal care, perinatal care, delivery, neonatal care, treatment costs of life-time morbidities (NHS reference costs + literature)	women's and offspring life-time (up to 100y)	£ 2014-2015	3.5 %
57	Mundt et al.	CEA	RCT/model	healthcare	number of quitters	v			incentives, services, staff and medication costs	6 months postpartum	$ 2020	
52	Naughton et al.	CEA	RCT	healthcare	quit rate (+7 measures of smoking cessation)	v			intervention costs	>25 weeks gestation up to 36 week gestation	£ 2014-2015	

53	Parker et al.	CEA	RCT (smokers not randomized between call groups)	healthcare (agency provider)	quit rate (7days abstinence) assessment at 32 weeks and 6 weeks and 6 months postpartum (self-report controlled by urine sample)	v			intervention costs	6 months	$ 2006 (assumed)	
55	Ruger et al.	CUA	RCT	healthcare	1) smoking cessation & relapse prevention, biochemically verified.(responses at baseline, 1 month after intervention and 6 months postpartum) 2) infant: birth weight and post-delivery status 3) QALYs, life-years	v			patient time, net resource costs: 1)intervention costs; 2) cost savings for neonatal intensive care, chronic medical conditions, and acute conditions during the first year of life; and 3) cost savings for maternal health care (cardiovascular and lung diseases).	6 months postpartum	$ 1997	3% for QALYs

56	Ussher et al.	CEA	RCT (Leap trial)	healthcare + social services	biochemically validated abstinence from smoking between a quit date and the end of pregnancy	v			Intervention cost + costs of caring for each woman and her infant during the period between randomisation and the immediate postnatal period (in terms of expected annual cost)	up to 9 months (10-24 weeks gestation up to 10 weeks post-partum	£ 2012-2013	
SUBSTANCE ABUSE												
61	Premkumar et al.	CUA	model	healthcare	QALYs (maternal perspective: maternal health + maternal disutilities related to neonatal health state)	v			treatment costs, medical care costs including neonatal care (related to health)	26 weeks	$ 2017	
64	Robin et al.	CUA	model	societal	maternal and neonatal outcomes in terms of QALYs	v	v		drug and healthcare costs	lifetime	$ 2020	3% for QALYs
63	Barlow et al.	CEA/ CUA	RCT	healthcare + social services	child abuse potential, parental emotional regulation, QALYs	v		(v)	program costs, hospital, community health and social services, legal services and	12 months	£ 2016	

					(baseline, 6 and 12 months) (if both parents had alcohol or drug problem primary caregiver was assessed)				costs borne by parents			
59	French et al.	CBA	Non-randomized effectiveness study + model	healthcare	$ equivalent of quality adjusted life day related to substance abuse, psychiatric status. Selected variables from the Addiction Severity Index (ASI) converted into monetary equivalents, self-reported at treatment entry and 6 month post discharge	v			reimbursed costs of treatment	6 months	$1998	

62	Thanh et al.	CEA	Decision analytic modelling	societal	number of FASD cases avoided		v		cost of intervention compared to lifetime cost of case FASD	3 years	CAD 2013	5%
60	Gifford et al.	CBA	model	healthcare	FASD cases prevented (in terms of costs)	v	v		medical, education, social services and out-of-pocket costs	lifetime per child and per woman	$2006	? not ment-ionned

Table 5: Cost-effectiveness results

Ref	Author	Type of economic evaluation	Incremental health gain	Incremental costs	ICER	Key conclusion	Sensitivity analysis/ key determinants outcomes	Quality/ bias considerations
ANXIETY & DEPRESSION								
26	Turkstra et al.	CEA	(-) 0,024 (EQ-5D-3L level)	-€ 66	€2.758 (cost per 0.1 point improvement on EQ-5D-3L)	The intervention did not increase costs; however, it might be cost-effective for those women with very high childbirth fear.	The probability that the intervention was more effective was 12%, while the probability that the intervention was less costly was 58%. (no sensitivity analysis)	follow-up retention rate of 54%, only public hospitals, self-reported patient data

32	Eldar-Lissai et al.	CUA	0.286 QALYs (0.25 mother, 0.036 child)	€ 21.776	€76.074 (and €48.940 for women with severe PPD)	BRX is a cost-effective therapy compared with SSRIs for treating women with PPD in the US.	probabilistic sensitivity analysis	ADHD used as a proxy for behavioural difficulties of child due to lack of appropriate utility values. Key factor is duration of treatment: 4 week model yields ICER of $5million, 18 year: $60,000.
27	Ammerman et al.	CUA	0,07 QALYs	-€ 1,50	The intervention dominates.	IH-CBT is a more cost-effective treatment for low-income, depressed mothers than current standards of practice. (driven by reduction in expected depression days).	Results were most sensitive to: transition from remission to MDD, transition from MDD to remission, cost of MDD medications.	Relatively small number of mothers, located in the same region. Constant transition probabilities from remission to MDD were assumed, limited follow-up window.
36	Morrell et al.	CUA	at-risk women: 0.003 QALY, all women: 0.002 QALY	at-risk women 6 months: €-48.505 (non-sign.), all women: €-27.865	(not reported)	HV intervention was highly likely to be cost-effective compared with the control. There was no difference in outcomes between the	Calculation of CEACs, modelling of missing data: 30% missing.	Three issues were noted: the impact of missing data, the applied clustering and costing method and data were not normally distributed, hence parametric tests were possibly biased.

						CBA and the PCA groups.		
41	Stevenson et al.	CUA	0.032 QALY	€1908 per women	€59.099/ QALY gained (base case) €45.870/ QALY (PSA incorporating stochastic values)	Group CBT is unlikely to be cost-effective based on used assumptions.	considerable uncertainty in the model parameters (probabilistic sensitivity analysis)	No data available to compare group-CBT with CBT, unknown role of concurrent medication, only one RCT was used to populate efficacy data.
45	Trevillion et al.	2020	CUA	no significant difference in QALYs	€8850 (of which €448 intervention)	€7723 per ALY	GSH was cheaper but less effective on average than usual care alone, the probability of being cost-effective compared with usual care is around 50% at a threshold of £20 000–£30 000 per QALY. Results remain uncertain.	Results based on the secondary analysis using EQ-5D-5L-based QALYs, and results of the sensitivity analyses did not alter the significance of outcomes.

33	Grote et al.	CEA	with comorbid PTSD: 68 more DFDs major depression alone: 13 more DFDs	with comorbid PTSD: €1.008 major depression alone: €897	with comorbid PTSD: €15 per DFD major depression alone: €69 per DFD	Women with comorbid PTSD: MOMCare intervention was more effective than MSS-Plus; major depression alone: similar improvement in both treatment conditions	Limited sensitivity analysis: correction for missing data and skewed distribution of costs.	No inclusion of observational measures of the mother-child relationship or a standardized assessment of child development. Mental health service use was self-reported and there was possible bias due to missing follow up data.
39	Petrou et al.	CEA	-0,49 months less depressed on average (nonsign)	€293 (nonsign)	€63 per month of postnatal depression avoided	At a willingness-to-pay (WTP) threshold of £1000 per month of postnatal depression avoided, the probability of cost-effectiveness is 0,71 (likely to be cost-effective even at low WTP)	Univariate analysis, non-parametric bootstrapping for CE acceptability curves and alternative WTP thresholds, broad confidence intervals (CI).	A broader, societal perspective would allow the consideration of direct nonmedical costs (e.g., travel and childcare costs), indirect costs (e.g., lost productivity), and intangible costs (e.g., costs of fear, pain, and suffering). Limited time horizon, no preference based outcome measure. Characteristics of declined population were not reported.

29	Boath et al.	CEA	14 less women depressed	€ 46.830	€3345 per successfully treated woman	The treatment should be recommended to health care decision-makers.	Results were sensitive to inclusion of primary care contacts and costs of medication.	Initial study dates from 1992-1993, low number of participants, uncertainty because of a lack of spontaneous recovery rates estimates. Infant health is not considered by source study, limited time horizon: utilisation of services extended beyond this period. Non-randomized data.
60	Barlow et al.	CEA/CUA	0,07 QALYs or 2,376 improvement on the Risk Abuse Scale from the Brief Child Abuse Potential Inventory (BCAP)	€ 2.647	€36.391 per QALY or €1072 per unit of improvement in BCAP	The probability that the program is cost-effective was ~51.8% if decision-makers are willing to pay £1000 for a unit improvement in BCAP. Significant improvements in emotional regulation, and measures of mood and borderline psychopathology.	Scenario analysis (limited discussion), further research is needed.	ICER in terms of QALYs does not capture effects on child. The primary outcome measure was the parent report of child abuse potential. There is likely to be variability in the quality and nature of the community-based addiction services supporting parents.

31	Dukhovny et al.	CEA	11% absolute risk reduction in PPD in the peer support group	€ 739	€6622 per case of PPD averted (50% probability)	The costs are within the range for other accepted interventions for this population.	Results were sensitive to the health region costs to implement the program and opportunity costs of family/friend, partner time off work.	Effectiveness measure (EPDS) score is a screening tool and not diagnostic. Limited time horizon, possible recall bias questionnaire
35	Morell et al.	CEA	no difference in SF-36 scores	no difference except for costs support worker service	/	No health benefit of additional home visits compared with traditional community midwifery visiting as measured by the SF36, 'no savings' to the NHS: mean difference in total costs was €281.	limited sensitivity analysis/ discussion of uncertainty	RCT from 1996-1997, 79% response rate, SF-36 likely too insensitive to detect changes (more than 75% found the support better than expected)

40	Saing et al.	CUA	0,017 QALYs	€ 37	€2171/QALY	Infant sleep consultations are cost-effective and led to improvements in quality of life through a reduction in postnatal depression.	Univariate and probabilistic sensitivity analysis, model was most sensitive to probability of overnight residential stays and baseline EPDS mean score.	Utility mapping was based on other study: different postnatal population. Costs are self-reported. Costs of treating postnatal depression were not included.
28	Asper et al.	CUA	0.03 QALYs	-€ 1.118	Base case analysis resulted in a negative ICER (€–37.266)	Program dominates the no screening program.	70% probability of being cost-effective, results were sensitive to variables of QALYs for the depressed fathers, probabilities of remission in treatment and no treatment groups, start age and productivity losses.	Study is based on only secondary data-further research required. Screening costs was assumed to be equal to cost of nurses' time (excluding administration and training costs). Possible overestimation productivity loss based on human capital approach.
30	Campbell et al.	CUA	616 QALYs	€ 1.248.822	€2027 per QALY gained	Introducing formalised screening for PND appears to represent good value for money.	Uni- and multivariate sensitivity analysis, results depend on treatment uptake and subsidy level GP.	Out-of-pocket costs and broader social impact were not included. Base case assumes 100% treatment uptake (unrealistic). Utility values related to a

								general population with depression using antidepressants.
	Campbell et al.	CEA	true positive rate: +0.035, false positive rate: -0.128	true positives: €-0.13, false positives: €-0.37	true positives: €-2, false positives: €2,9	PIPA model was cost-saving and more effective at eliminating false positives and identifying 'at-risk' women.	PSA conducted, great degree of uncertainty in outcomes (large CI's)	not randomized cohorts in different years, only intermediate outcomes (and costs) included
34	Henderson et al.	CUA	0.002 QALY	€ -114	€-4884/QALY	PoNDER HV training was highly cost-effective in preventing symptoms of PND in a population of lower-risk women.	Multivariate sensitivity analysis, impact of the intervention appears to have been relatively uniform over the whole of the lower-risk sample.	Included costs were limited to health and social care services, not included: longer term adverse effects on child development and costs, employment related productivity losses.

38	Paulden et al.	CUA	EPDS cut point 16: 0.0006 QALY	€ 32	EDPS at a cut point of 16: €53.806/QALY compared with routine care only. ICER all other strategies: €65.358 to €356.667/QALY	Probability that no formal identification strategy was cost-effective was 88% (59%) at a cost effectiveness threshold of £20 000 (£30 000) per QALY.	Cost of managing incorrectly identified depression (false positive result) was an important driver of the model.	Probability that depression is detected and utility weights were based on values for 'moderate depression' in general. Whooley questions not considered in base case because of lack of data. Family members were not considered, insufficient data for subgroup analysis.
44	Premji et al.	2021	CUA	0,0021 QALYs	€29	€13.666 per QALY	Screening is a favourable strategy, resulting in 11% more cases being diagnosed annually relative to not screening.	With 100% attending referral, the ICER fell to €8.113 per QALY. Probabilistic analysis, model most sensitive to % of women receiving pharmaceutical/mixed treatment.
42	Wilkinson et al.	CEA/CUA	21.43 QALYs or 29 more healthy women	€713 per woman	€7.698 per remission, €10.477 per QALY	Screening for and treating postpartum depression is a cost-effective intervention	Results were robust in both the deterministic and probabilistic sensitivity analyses of input parameters.	limited time horizon to capture all relevant outcomes, variety of data sources (some before the year 2000)

37	NCCMH	CEA/CUA	Whooley questions followed by EPDS: 0,113 QALYs	€ 5.628	ICER of Whooley questions followed by EPDS versus Whooley questions followed by PHQ-9: €49.696/QALY (which is above threshold)	The use of formal identification comprises a cost-effective strategy when compared with standard care case identification.	Threshold sensitivity analyses showed that the results were sensitive to the diagnostic characteristics of formal case identification tools and consultation time required to administer case identification tool.	Due to lack of available evidence, a number of the estimates used in the economic model were based on single studies and where necessary supplemented by the GDG expert opinion. Utility values for general depression were used.
43	Ride et al.	CEA/CUA	complete case: 1.77 percentage point lower prevalence of depression, anxiety and adjustment disorders and +0.007 QALYs, multiple imputation of missing data: 0.33 pp and +0.006 QALYs	€148 per participant; with imputation of missing data: €137	€20/ QALY	Neither costs nor outcomes were statistically significantly different.	Probabilistic sensitivity and scenario analysis, there was considerable uncertainty surrounding the effectiveness of the intervention (55% prob. of being CE at a threshold of $A 55.000/QALY)	The cluster-randomised nature of the trial and small but non-negligible intra-cluster correlation coefficient for QALYs may also have reduced the ability to detect an effect of the intervention on QALYs in this trial.

SMOKING CESSATION								
50	Essex et al.	CEA	1.8% quit rate (non-sign)	€110	€5967 per additional quitter	Due to high levels of statistical uncertainty, it was hard to determine the cost-effectiveness of NRT in this population.	Sensitivity analysis including only singleton births yielded an ICER of €5033 per quitter.	low compliance: only 7.2% of women in the NRT group and 2.8% in the placebo arm used trial patches for longer than 1 month.
54	Pollack et al.	CEA	108 SIDS deaths averted	€49 per participant	€231.091 per SIDS death averted	Prenatal smoking cessation programs are estimated to cost less than $11.000 per life-year. Typical prenatal smoking cessation programs are highly cost-effective but have limited impact on the population incidence of SIDS.	No sensitivity analysis, only CI were reported.	Self-reported smoking data, postnatal maternal smoking or household members not considered. Impact of race/ethnicity not considered. Only SIDS was considered, no other benefits of smoking cessation such as reduced incidence of low birthweight, maternal complications in pregnancy, childhood asthma, lung cancer, and cardiovascular disease.

49	Bacheller et al.	CUA	1050 QALYs	€74 million	€70,800 per QALY	The intervention is cost-effective and leads to fewer adverse neonatal outcomes.	intervention was CE at probabilities of smoking cessation > 11.6% or cost of the intervention < $475.21	simplified model assumptions, long-term down stream effects of smoking not included
58	Dornelas et al.	CEA	incremental quit rate at end of pregnancy= 18.7	€56 per patient	€298 per quitter at end of pregnancy	Intervention is cost-effective and is most effective early in pregnancy and for women under age 25.	(missing)	Non-generalizable population (specific segment: overall smoking rate was very high), 68% attendance rate counselling (these people were still contacted and still had similar quit rates).
47	Bell et al.	CEA	increased quitting rate: 1.81	€34 per additional delivery	€1045 per additional quit	The intervention was associated with a significant increase in rates of quitting by delivery.	(limited)	Routinely collected data from different sources, some had high levels of missing data and were non-randomized or observational.
48	Boyd et al.	CEA/CUA	trial: 0.14 quit rate lifetime: 0,036 QALY	trial: €172 lifetime: €19	€1231 per quitter Lifetime model: €519/QALY	Financial incentives (shopping vouchers) for smoking cessation in pregnancy are	Probabilistic sensitivity analysis indicates uncertainty of results related to relapse after birth.	risk of untruthful reporting

						highly cost-effective.		
51	Jones et al.	CUA	0.04 QALYs	€-41	Dominant: €-1036 per QALY	Intervention was very likely to be cost-effective in the longer term and to generate health-care savings. Greatest benefit comes from long-term perspective which was not included in original analysis.	probabilistic sensitivity analysis	The initial evaluation of MiQuit found non-significant effectiveness results. RCT with only short-term results. Trial did not collect pregnancy outcomes and ended at 36 weeks of gestation. Smoking of household members was not included. Potential underestimation of long-term abstinence because success of quitting increases per attempt (but was kept constant).
57	Mundt et al.	CEA	5.5% increase in 6 months post-birth biochemically-confirmed tobacco abstinence	€182 per participant	€3952 per additional woman tobacco abstinent at 6m post-birth	financial incentives are cost-effective for socio-economically disadvantaged	(limited)	pregnancy and infant health outcomes not included

						pregnant women who smoke		
52	Naughton et al.	CEA/CUA	incremental quit rate= 3.46%	€ 5	€144 per quitter based on lifetime utility gain values: €75 per QALY	Probability of MiQuit being cost-effective was 96.5% (at a threshold of £10.000).	As only 34% of those setting a quit date achieve longer-term abstinence, the cost-per-quitter, inflated accordingly, is probably closer to €650.	High dropout rate (adjusted conservatively by assuming that all dropouts were still smoking), short time horizon.
53	Parker et al.	CEA/CBA	3 calls: 38 quitters, no call: 5 quitters	3 calls: €2.829 no call: €123	reported effectiveness to cost ratio: 1: €74 for 3 calls	Telephone counselling is a potentially cost-effective approach to help a low-income, underserved population of pregnant women quit smoking.	Results were tested for: exclusion of the no-call group, changes in price and covariates.	Low number of urine samples collected: 114 first visit, 113 third trimester and 23 after 6 months postpartum. 46% of participants received all 3 calls.

55	Ruger et al.	CUA	smoking cessation: -0.04 QALYs, relapse prevention: 0.49 QALYs	€ 340	Smoking cessation: MI is more costly but less effective relapse prevention: €949/LY saved and €701/QALY saved	Among low-income pregnant women, MI helps prevent relapse at relatively low cost, and may be cost-saving when net medical cost savings are considered.	Including savings in maternal lifetime medical costs in sensitivity analyses resulted in cost savings for MI for relapse prevention compared to UC.	Specific patient sub-group (low income), long-term morbidity and mortality data for children were not included, sample size was limited.
56	Ussher et al.	CEA	1,3% incremental quit rate, 0.76 expected annual quitters	€-2253 (expected annual cost) or €-39 per participant	not calculated, intervention dominates alternatives	Physical activity seems cost-effective but results are quite uncertain.	considerable statistical uncertainty, limited sensitivity analysis (no adjustment for quit rates)	Low attendance may have affected efficacy, intervention contamination might have occurred in control group (insufficient difference in PA between two groups). Participants seem to overestimate self-reported PA levels.

SUBSTANCE ABUSE								
61	Premkumar et al.	CUA	Buprenorphine: - Methadone: -0.23 Detoxification: 0.13	Buprenorphine: - Methadone: €6.448 Detoxification: €17.237	methadone & detoxification are dominated	Initiation of buprenorphine was the dominant strategy.	Nonetheless, Buprenorphine was not found cost-effective in almost one out of three of simulations suggesting limited robustness of the model - further research is needed.	Long-term developmental outcomes of the offspring associated with each of the strategies remains unknown, and there are limited data focusing on adherence in the postpartum period for women using methadone or buprenorphine, and no data for detoxification.
64	Robin et al.	CUA	558 QALYs	€121.5 million	methadone is dominated	Buprenorphine is cost saving and reduces neonatal morbidity and mortality.	61% probability of being CE, impactful parameters: probability of neonatal opioid withdrawal syndrome, discontinuing buprenorphine, preterm birth and stillbirth	only self-reported data during pregnancy, high dropout rates, comparator is not no-intervention

63	Barlow et al.	CEA/CUA	0,07 QALYs or 2,376 improvement on the Risk Abuse Scale from the Brief Child Abuse Potential Inventory (BCAP)	€ 2.647	€36.391 per QALY or €1072 per unit of improvement in BCAP	The probability that the program is cost-effective was ~51.8% if decision-makers are willing to pay £1000 for a unit improvement in BCAP. Significant improvements in emotional regulation, and measures of mood and borderline psychopathology.	Scenario analysis (limited discussion), further research is needed.	ICER in terms of QALYs does not capture effects on child. The primary outcome measure was the parent report of child abuse potential. There is likely to be variability in the quality and nature of the community-based addiction services supporting parents.
59	French et al.	CBA	€ 17.150,14 (Specialty relative to standard)	specialty: €8.821 per client standard: €1.611 per client	per client net benefit €18.820 for specialty, €8.881 standard. Benefit-cost ratio: 3,1 specialty and 6,5 standard	Both Specialty and Standard clients showed significant clinical improvements and positive net benefits resulting from treatment.	Economic benefits were not distributed evenly over different outcome categories. No statistical significance of total benefit or net benefit.	Nonrandomized field study, patients receiving specialized treatment had more severe issues: samples differed at baseline. Unknown reliability and validity of interview instrument. Missing data: drop rate #cases= 32%, low follow up rate (56%), self-reported data, selection bias (substance abusers

								that voluntarily seek treatment).
62	Thanh et al.	CEA	31 prevented FASD cases	€ 62.650	NMB= €24,4 million	The program is cost–effective and the net monetary benefit is significant.	Increasing the use of contraceptives had a significant impact on the outcomes.	Benefits from reduction in unemployment, welfare income dependence and potential drug abuse were not included (results likely underestimated).
60	Gifford et al.	CBA	Benefits from preventing FASD births range between €30,7 - €37,3 billion	Total costs of screening and treatment for all positive tests range between €0,31 - 4,26 billion.	Potential financial savings range from €6 to €97 for every €1 spent on screening and treatment.	Universal meconium analysis of newborns and subsequent intervention could be cost-effective to reduce the incidence of FAS and FASD.	Sensitivity analysis was conducted to test result for social drinkers instead of binge drinkers (sensitivity of 80%).	Costs of intervention after screening not included, economic values most likely underestimate full impact of FASD, psychological burden - quality of life not included, costs based on literature past 20 years.

References

1. World Health Organization. WHO | Maternal mental health [Internet]. WHO. 2020 [cited 2020 Sep 3]. Available from: https://www.who.int/mental_health/maternal-child/maternal_mental_health/en/
2. NICE. Antenatal and postnatal mental health: clinical management and service guidance [Internet]. Antenatal and postnatal mental health: clinical management and service guidance. 2020 [cited 2020 Sep 3]. Available from: https://www.nice.org.uk/guidance/cg192
3. Shorey S, Chee CYI, Ng ED, Chan YH, Tam WWS, Chong YS. Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis. Journal of Psychiatric Research. 2018 Sep 1;104:235–48.
4. Wisner KL, Sit DKY, McShea MC, Rizzo DM, Zoretich RA, Hughes CL, et al. Onset Timing, Thoughts of Self-harm, and Diagnoses in Postpartum Women With Screen-Positive Depression Findings. JAMA Psychiatry. 2013 May 1;70(5):490–8.
5. Rao W-W, Zhu X-M, Zong Q-Q, Zhang Q, Hall BJ, Ungvari GS, et al. Prevalence of prenatal and postpartum depression in fathers: A comprehensive meta-analysis of observational surveys. Journal of Affective Disorders. 2020 Feb 15;263:491–9.
6. Gressier F, Tabat-Bouher M, Cazas O, Hardy P. Dépression paternelle du post-partum : revue de la littérature. La Presse Médicale. 2015 Apr 1;44(4, Part 1):418–24.
7. Paulson JF, Bazemore SD. Prenatal and Postpartum Depression in Fathers and Its Association With Maternal Depression: A Meta-analysis. JAMA. 2010 May 19;303(19):1961–9.
8. Goodman JH. Paternal postpartum depression, its relationship to maternal postpartum depression, and implications for family health. J Adv Nurs. 2004 Jan;45(1):26–35.
9. Howard LM, Molyneaux E, Dennis C-L, Rochat T, Stein A, Milgrom J. Non-psychotic mental disorders in the perinatal period. The Lancet. 2014 Nov 15;384(9956):1775–88.
10. Verreault N, Da Costa D, Marchand A, Ireland K, Banack H, Dritsa M, et al. PTSD following childbirth: A prospective study of incidence and

risk factors in Canadian women. Journal of Psychosomatic Research. 2012 Oct 1;73(4):257–63.
11. Bulik CM, Von Holle A, Hamer R, Knoph Berg C, Torgersen L, Magnus P, et al. Patterns of remission, continuation, and incidence of broadly defined eating disorders during early pregnancy in the Norwegian Mother and Child Cohort Study (MoBa). Psychol Med. 2007 Aug;37(8):1109–18.
12. Manning C, Gregoire A. Effects of parental mental illness on children. Psychiatry. 2009 Jan 1;8(1):7–9.
13. Van den Bergh BRH, van den Heuvel MI, Lahti M, Braeken M, de Rooij SR, Entringer S, et al. Prenatal developmental origins of behavior and mental health: The influence of maternal stress in pregnancy. Neurosci Biobehav Rev. 2017 Jul 28;
14. Nilsson E, Lichtenstein P, Cnattingius S, Murray RM, Hultman CM. Women with schizophrenia: pregnancy outcome and infant death among their offspring. Schizophr Res. 2002 Dec 1;58(2–3):221–9.
15. Andres RL, Day MC. Perinatal complications associated with maternal tobacco use. Semin Neonatol. 2000 Aug;5(3):231–41.
16. Jaakkola JJK, Gissler M. Maternal Smoking in Pregnancy, Fetal Development, and Childhood Asthma. Am J Public Health. 2004 Jan 1;94(1):136–40.
17. Durmus B, Kruithof CJ, Gillman MH, Willemsen SP, Hofman A, Raat H, et al. Parental smoking during pregnancy, early growth, and risk of obesity in preschool children: the Generation R Study. Am J Clin Nutr. 2011 Jul;94(1):164–71.
18. Bauer A, Knapp M, Parsonage M. Lifetime costs of perinatal anxiety and depression. Journal of Affective Disorders. 2016 Mar 1;192:83–90.
19. Robakis T, Jernick E, Williams K. Recent advances in understanding maternal perinatal mood disorders. F1000Res. 2017 Jun 15;6:916.
20. Austin M-P, Highet N, Expert Working Group (2017). Mental Health Care in the Perinatal Period. Australian Clinical Practice Guideline [Internet]. Centre of Perinatal Excellence; 2017. Available from: https://www.cope.org.au/wp-content/uploads/2018/05/COPE-Perinatal-MH-Guideline_Final-2018.pdf
21. Gurung B., Jackson L.J., Monahan M., Butterworth R., Roberts T.E. Identifying and assessing the benefits of interventions for postnatal

depression: A systematic review of economic evaluations. BMC Pregnancy Childbirth [Internet]. 2018;18(1). Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L622211140
22. Camacho EM, Shields GE. Cost-effectiveness of interventions for perinatal anxiety and/or depression: a systematic review. BMJ Open. 2018 Aug;8(8):e022022.
23. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. PLoS Med. 2009 Jul 21;6(7):e1000100.
24. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons; 2019. 726 p.
25. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
26. Turkstra E, Mihala G, Scuffham PA, Creedy DK, Gamble J, Toohill J, et al. An economic evaluation alongside a randomised controlled trial on psycho-education counselling intervention offered by midwives to address women's fear of childbirth in Australia. Sexual & Reproductive Healthcare. 2017 Mar 1;11:1–6.
27. Ammerman RT, Mallow PJ, Rizzo JA, Putnam FW, Van Ginkel JB. Cost-effectiveness of In-Home Cognitive Behavioral Therapy for low-income depressed mothers participating in early childhood prevention programs. J Affect Disord. 2017 Jan 15;208:475–82.
28. Asper MM, Hallén N, Lindberg L, Månsdotter A, Carlberg M, Wells MB. Screening fathers for postpartum depression can be cost-effective: An example from Sweden. J Affect Disord. 2018 Dec 1;241:154–63.
29. Boath E, Major K, Cox J. When the cradle falls II: the cost-effectiveness of treating postnatal depression in a psychiatric day hospital compared with routine primary care. Journal of Affective Disorders. 2003 Apr 1;74(2):159–66.
30. Campbell S, Norris S, Standfield L, Suebwongpat A. Screening for postnatal depression within the Well Child Tamariki Ora framework: an economic analysis of implementation of a screening programme [Internet]. Christchurch, N.Z.: Health Services Assessment

Collaboration (HSAC), University of Canterbury; 2008 [cited 2020 Aug 20]. (HSAC Report). Available from: http://www.healthsac.net/downloads/publications/HSAC01%20PND%20170608%20Final.pdf

31. Dukhovny D, Dennis C-L, Hodnett E, Weston J, Stewart DE, Mao W, et al. Prospective economic evaluation of a peer support intervention for prevention of postpartum depression among high-risk women in Ontario, Canada. Am J Perinatol. 2013 Sep;30(8):631–42.
32. Eldar-Lissai A, Cohen JT, Meltzer-Brody S, Gerbasi ME, Chertavian E, Hodgkins P, et al. Cost-Effectiveness of Brexanolone Versus Selective Serotonin Reuptake Inhibitors for the Treatment of Postpartum Depression in the United States. Journal of Managed Care & Specialty Pharmacy. 2020 May;26(5):627-+.
33. Grote NK, Simon GE, Russo J, Lohr MJ, Carson K, Katon W. Incremental Benefit-Cost of MOMCare: Collaborative Care for Perinatal Depression Among Economically Disadvantaged Women. Psychiatr Serv. 2017 Nov 1;68(11):1164–71.
34. Henderson C, Dixon S, Bauer A, Knapp M, Morrell CJ, Slade P, et al. Cost-effectiveness of PoNDER health visitor training for mothers at lower risk of depression: findings on prevention of postnatal depression from a cluster-randomised controlled trial. Psychol Med. 2019 Jun;49(8):1324–34.
35. Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A. Costs and effectiveness of community postnatal support workers: randomised controlled trial. BMJ. 2000 Sep 9;321(7261):593–8.
36. Morrell C.J., Warner R., Slade P., Dixon S., Walters S., Paley G., et al. Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PoNDER trial. Health Technol Assess. 2009;13(30):iii–iv, xi–xiii, 1–153.
37. National Collaborating Centre for Mental Health. Case identification and assessment, psychological and psychosocial interventions for the prevention of treatment of mental health problems. In Antenatal and postnatal mental health: the NICE guideline on clinical management and service guidance (update). [Internet]. The British Psychological Society and The Royal College of Psychiatrists; 2018. Available from: https://www.nice.org.uk/guidance/cg192/evidence/full-guideline-pdf-4840896925

38. Paulden M, Palmer S, Hewitt C, Gilbody S. Screening for postnatal depression in primary care: cost effectiveness analysis. BMJ [Internet]. 2009 Dec 22 [cited 2020 Aug 14];339. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2797050/
39. Petrou S, Cooper P, Murray L, Davidson LL. Cost-effectiveness of a preventive counseling and support package for postnatal depression. International Journal of Technology Assessment in Health Care. 2006 Oct;22(4):443–53.
40. Saing S, Parkinson B, Church J, Goodall S. Cost Effectiveness of a Community-Delivered Consultation to Improve Infant Sleep Problems and Maternal Well-Being. Value Health Reg Issues. 2018 May;15:91–8.
41. Stevenson MD, Scope A, Sutcliffe PA. The cost-effectiveness of group cognitive behavioral therapy compared with routine primary care for women with postnatal depression in the UK. Value Health. 2010 Aug;13(5):580–4.
42. Wilkinson A, Anderson S, Wheeler SB. Screening for and Treating Postpartum Depression and Psychosis: A Cost-Effectiveness Analysis. Matern Child Health J. 2017 Apr;21(4):903–14.
43. Ride J, Lorgelly P, Tran T, Wynter K, Rowe H, Fisher J. Preventing postnatal maternal mental health problems using a psychoeducational intervention: the cost-effectiveness of What Were We Thinking. Bmj Open. 2016;6(11):e012086.
44. Premji S, McDonald SW, Metcalfe A, Faris P, Quan H, Tough S, et al. Examining postpartum depression screening effectiveness in well child clinics in Alberta, Canada: A study using the All Our Families cohort and administrative data. Prev Med Rep [Internet]. 2019 May 3 [cited 2020 Aug 11];14. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6517566/
45. Trevillion K, Ryan EG, Pickles A, Heslin M, Byford S, Nath S, et al. An exploratory parallel-group randomised controlled trial of antenatal Guided Self-Help (plus usual care) versus usual care alone for pregnant women with depression: DAWN trial. J Affective Disord. 2020;261((Trevillion K., kylee.trevillion@kcl.ac.uk; Ryan E.G.; Pickles A.; Heslin M.; Byford S.; Nath S.; Mycroft R.; Domoney J.; Pariante C.; Hunter M.S.; Howard L.M.) Institute of Psychiatry,

Psychology & Neuroscience at King's College London, United Kingdom):187–97.

46. Chambers GM, Botha W, Reilly N, Black E, Kingston D, Austin M-P. The clinical performance and cost-effectiveness of two psychosocial assessment models in maternity care: The Perinatal Integrated Psychosocial Assessment study. Women and Birth [Internet]. 2021 Jun 5 [cited 2021 Oct 28]; Available from: https://www.sciencedirect.com/science/article/pii/S187151922100086X
47. Bell R, Glinianaia SV, Waal Z van der, Close A, Moloney E, Jones S, et al. Evaluation of a complex healthcare intervention to increase smoking cessation in pregnant women: interrupted time series analysis with economic evaluation. Tobacco Control. 2018 Jan 1;27(1):90–8.
48. Boyd KA, Briggs AH, Bauld L, Sinclair L, Tappin D. Are financial incentives cost-effective to support smoking cessation during pregnancy? Addiction. 2016 Feb;111(2):360–70.
49. Dornelas EA, Magnavita J, Beazoglou T, Fischer EH, Oncken C, Lando H, et al. Efficacy and cost-effectiveness of a clinic-based counseling intervention tested in an ethnically diverse sample of pregnant smokers. Patient Education and Counseling. 2006 Dec 1;64(1):342–9.
50. Essex HN, Parrott S, Wu Q, Li J, Cooper S, Coleman T. Cost-Effectiveness of Nicotine Patches for Smoking Cessation in Pregnancy: A Placebo Randomized Controlled Trial (SNAP). Nicotine Tob Res. 2015 Jun 1;17(6):636–42.
51. Jones M, Smith M, Lewis S, Parrott S, Coleman T. A dynamic, modifiable model for estimating cost-effectiveness of smoking cessation interventions in pregnancy: application to an RCT of self-help delivered by text message. Addiction. 2019 Feb;114(2):353–65.
52. Naughton F, Cooper S, Foster K, Emery J, Leonardi-Bee J, Sutton S, et al. Large multi-centre pilot randomized controlled trial testing a low-cost, tailored, self-help smoking cessation text message intervention for pregnant smokers (MiQuit). Addiction. 2017 Jul;112(7):1238–49.
53. Parker D, Windsor R, Roberts M, Hecht J, Hardy N, Strolla L, et al. Feasibility, cost, and cost-effectiveness of a telephone-based motivational intervention for underserved pregnant smokers. Nicotine & Tobacco Res. 2007 Oct;9(10):1043–51.

54. Pollack HA. Sudden infant death syndrome, maternal smoking during pregnancy, and the cost-effectiveness of smoking cessation intervention. Am J Public Health. 2001 Mar;91(3):432–6.
55. Ruger JP, Weinstein MC, Hammond SK, Kearney MH, Emmons KM. Cost-Effectiveness of Motivational Interviewing for Smoking Cessation and Relapse Prevention among Low-Income Pregnant Women: A Randomized Controlled Trial. Value Health. 2008 Apr;11(2):191–8.
56. Ussher M, Lewis S, Aveyard P, Manyonda I, West R, Lewis B, et al. The London Exercise And Pregnant smokers (LEAP) trial: a randomised controlled trial of physical activity for smoking cessation in pregnancy with an economic evaluation. Health Technol Assess. 2015 Oct;19(84):vii–xxiv, 1–135.
57. Mundt MP, Fiore MC, Piper ME, Adsit RT, Kobinsky KH, Alaniz KM, et al. Cost-effectiveness of stop smoking incentives for medicaid-enrolled pregnant women. Preventive Medicine. 2021 Dec;153:106777.
58. Bacheller HL, Hersh AR, Caughey AB. Behavioral Smoking Cessation Counseling During Pregnancy: A Cost-Effectiveness Analysis. Obstetrics & Gynecology. 2021 Apr;137(4):703–712.
59. French M.T., McCollister K.E., Cacciola J., Durell J., Stephens R.L. Benefit-cost analysis of addiction treatment in Arkansas: Specialty and standard residential programs for pregnant and parenting women. Subst Abuse. 2002;23(1):31–51.
60. Gifford A.E., Farkas K.J., Jackson L.W., Molteno C.D., Jacobson J.L., Jacobson S.W., et al. Assessment of benefits of a universal screen for maternal alcohol use during pregnancy. Birth Defects Res Part A Clin Mol Teratol. 2010;88(10):838–46.
61. Premkumar A, Grobman WA, Terplan M, Miller ES. Methadone, Buprenorphine, or Detoxification for Management of Perinatal Opioid Use Disorder: A Cost-Effectiveness Analysis. Obstet Gynecol. 2019 Nov;134(5):921–31.
62. Thanh NX, Jonsson E, Moffatt J, Dennett L, Chuck AW, Birchard S. An Economic Evaluation of the Parent–Child Assistance Program for Preventing Fetal Alcohol Spectrum Disorder in Alberta, Canada. Adm Policy Ment Health. 2015 Jan 1;42(1):10–8.

63. Barlow J, Sembi S, Parsons H, Kim S, Petrou S, Harnett P, et al. A randomized controlled trial and economic evaluation of the Parents Under Pressure program for parents in substance abuse treatment. Drug Alcohol Depend. 2019;194((Barlow J.) Department of Social Policy and Intervention, University of Oxford, Oxford, United Kingdom):184–94.
64. Robin AM, Hersh AR, John C, Caughey AB. Cost effectiveness of buprenorphine vs. methadone for pregnant people with opioid use disorder. The Journal of Maternal-Fetal & Neonatal Medicine. 2021 Jan 17;1–9.
65. Dennis C-L, Dowswell T. Psychosocial and psychological interventions for preventing postpartum depression. Cochrane Database of Systematic Reviews [Internet]. 2013 [cited 2020 Oct 5];(2). Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD001134.pub3/full
66. Lange S, Probst C, Rehm J, Popova S. National, regional, and global prevalence of smoking during pregnancy in the general population: a systematic review and meta-analysis. The Lancet Global Health. 2018 Jul 1;6(7):e769–76.
67. Merrill L, Mittal L, Nicoloro J, Caiozzo C, Maciejewski PK, Miller LJ. Screening for bipolar disorder during pregnancy. Arch Womens Ment Health. 2015 Aug 1;18(4):579–83.
68. Howard LM, Khalifeh H. Perinatal mental health: a review of progress and challenges. World Psychiatry. 2020;19(3):313–27.
69. Patton GC, Romaniuk H, Spry E, Coffey C, Olsson C, Doyle LW, et al. Prediction of perinatal depression from adolescence and before conception (VIHCS): 20-year prospective cohort study. Lancet. 2015 Aug 29;386(9996):875–83.
70. Dolman C, Jones IR, Howard LM. Women with bipolar disorder and pregnancy: factors influencing their decision-making. BJPsych Open. 2016 Sep;2(5):294–300.
71. Holmes S. Responses to warnings about the impact of eating disorders on fertility: a qualitative study. Sociology of Health & Illness. 2018;40(4):670–86.

72. Rominov H, Pilkington PD, Giallo R, Whelan TA. A systematic review of interventions targeting paternal mental health in the perinatal period. Infant Ment Health J. 2016 May;37(3):289–301.
73. Shorey S, Chan V. Paternal mental health during the perinatal period: A qualitative systematic review. Journal of Advanced Nursing. 2020;76(6):1307–19.
74. Goldstein Z, Rosen B, Howlett A, Anderson M, Herman D. Interventions for paternal perinatal depression: A systematic review. Journal of Affective Disorders. 2020 Mar 15;265:505–10.
75. Khan SJ, Fersh ME, Ernst C, Klipstein K, Albertini ES, Lusskin SI. Bipolar Disorder in Pregnancy and Postpartum: Principles of Management. Curr Psychiatry Rep. 2016 Jan 19;18(2):13.
76. Doucet S, Jones I, Letourneau N, Dennis C-L, Blackmore ER. Interventions for the prevention and treatment of postpartum psychosis: a systematic review. Arch Womens Ment Health. 2011 Apr 1;14(2):89–98.
77. Jones I, Chandra PS, Dazzan P, Howard LM. Bipolar disorder, affective psychosis, and schizophrenia in pregnancy and the post-partum period. The Lancet. 2014 Nov 15;384(9956):1789–99.
78. Park ER, Chang Y, Quinn V, Regan S, Cohen L, Viguera A, et al. The association of depressive, anxiety, and stress symptoms and postpartum relapse to smoking: A longitudinal study. Nicotine Tob Res. 2009 Jun 1;11(6):707–14.
79. Seng JS, Rauch SAM, Resnick H, Reed CD, King A, Low LK, et al. Exploring posttraumatic stress disorder symptom profile among pregnant women. Journal of Psychosomatic Obstetrics & Gynecology. 2010 Sep 1;31(3):176–87.
80. Simoila L, Isometsä E, Gissler M, Suvisaari J, Halmesmäki E, Lindberg N. Schizophrenia and pregnancy: a national register-based follow-up study among Finnish women born between 1965 and 1980. Arch Womens Ment Health. 2020 Feb 1;23(1):91–100.
81. Micali N, Simonoff E, Treasure J. Pregnancy and post-partum depression and anxiety in a longitudinal general population cohort: The effect of eating disorders and past depression. Journal of Affective Disorders. 2011 Jun 1;131(1):150–7.

82. Whitaker RC, Orzol SM, Kahn RS. The co-occurrence of smoking and a major depressive episode among mothers 15 months after delivery. Preventive Medicine. 2007 Dec 1;45(6):476–80.
83. Grisbrook M-A, Letourneau N. Improving maternal postpartum mental health screening guidelines requires assessment of post-traumatic stress disorder. Can J Public Health [Internet]. 2020 Aug 11 [cited 2020 Sep 9]; Available from: https://doi.org/10.17269/s41997-020-00373-8
84. Velders FP, Dieleman G, Henrichs J, Jaddoe VWV, Hofman A, Verhulst FC, et al. Prenatal and postnatal psychological symptoms of parents and family functioning: the impact on child emotional and behavioural problems. Eur Child Adolesc Psychiatry. 2011 Jul 1;20(7):341–50.
85. Barker ED, Jaffee SR, Uher R, Maughan B. The contribution of prenatal and postnatal maternal anxiety and depression to child maladjustment. Depression and Anxiety. 2011;28(8):696–702.
86. Plant DT, Barker ED, Waters CS, Pawlby S, Pariante CM. Intergenerational transmission of maltreatment and psychopathology: the role of antenatal depression. Psychological Medicine. 2013 Mar;43(3):519–28.
87. Flach C, Leese M, Heron J, Evans J, Feder G, Sharp D, et al. Antenatal domestic violence, maternal mental health and subsequent child behaviour: a cohort study. BJOG: An International Journal of Obstetrics & Gynaecology. 2011;118(11):1383–91.
88. Milgrom J, Holt C, Holt CJ, Ross J, Ericksen J, Gemmill AW. Feasibility study and pilot randomised trial of an antenatal depression treatment with infant follow-up. Arch Womens Ment Health. 2015 Oct;18(5):717–30.
89. Turner HC, Lauer JA, Tran BX, Teerawattananon Y, Jit M. Adjusting for Inflation and Currency Changes Within Health Economic Studies. Value in Health. 2019 Sep 1;22(9):1026–32.
90. OECD.stat. Consumer price indices (CPIs) - Complete database [Internet]. [cited 2020 Sep 1]. Available from: https://stats.oecd.org/Index.aspx?DataSetCode=PRICES_CPI#
91. OECD Data. Conversion rates - Purchasing power parities (PPP) [Internet]. Purchasing power parities (PPP) - OECD Data. [cited 2021 Feb 26]. Available from: http://data.oecd.org/conversion/purchasing-power-parities-ppp.htm

Appendices

Appendix 1

(Overview of 3 main search strategies that were adapted for every specific database.)

1. PUBMED

Pregnancy
Postpartum[tiab] OR post-partum[tiab] OR postnatal[tiab] OR post-natal[tiab] OR perinatal[tiab] OR peri-natal[tiab] OR antepartum[tiab] OR ante-partum[tiab] OR antenatal[tiab] OR ante-natal[tiab] OR pregnan*[tiab] OR gestation [tiab] OR father*[tiab] OR paternal[tiab] OR matern*[tiab] OR mother*[tiab] OR parent* [tiab] OR “Birth Interval”[tiab] OR “Birth Spacing” [tiab] OR “Birth Spacings”[tiab] OR “Interpregnancy”[tiab] OR
"Pregnancy"[Mesh] OR "Maternal Health Services"[Mesh] OR "Pregnant Women"[Mesh] OR "Parents"[Mesh] OR “Postpartum Period”[Mesh] OR “Birth Intervals”[Mesh]

Cost-effectiveness
Economic-evaluation[tiab] OR economic-analys*[tiab] OR cost-effective*[tiab] OR costeffective*[tiab] OR cost-benefit*[tiab] OR cost-utilit*[tiab] OR cost-and-benefit*[tiab] OR costs-and-benefit*[tiab] OR benefit-and-cost*[tiab] OR benefits-and-cost*[tiab] OR "cost efficiency"[tiab] OR “Health Technology Assessment”[tiab] OR
"Cost-Benefit Analysis"[Mesh]

Mental health
depress*[tiab] OR anxi*[tiab] OR "mental health"[tiab] OR "mental disorder"[tiab] OR eating-disorder*[tiab] OR anorexia[tiab] OR bulimia[tiab] OR smoking[tiab] OR (substance[tiab] AND (abuse*[tiab] OR addiction[tiab] OR dependence[tiab])) OR ((drug*[tiab] AND (abuse*[tiab] OR addiction[tiab] OR dependence[tiab]) OR psycho*[tiab] bipolar[tiab] OR schizophren*[tiab] OR PTSD[tiab] OR Post-Traumatic[tiab] OR

posttraumatic[tiab] OR stress[tiab] OR "Mental Disorders"[Mesh] OR "Smoking Cessation"[Mesh]

2. EMBASE

Pregnancy:
postpartum:ti,ab,kw OR ‘post partum’:ti,ab,kw OR postnatal:ti,ab,kw OR ‘post natal’:ti,ab,kw OR perinatal:ti,ab,kw OR ‘peri natal’:ti,ab,kw OR antepartum:ti,ab,kw OR ‘ante partum’:ti,ab,kw OR antenatal:ti,ab,kw OR ‘ante natal’:ti,ab,kw OR pregnan*:ti,ab,kw OR gestation:ti,ab,kw OR father*:ti,ab,kw OR paternal:ti,ab,kw OR matern*:ti,ab,kw OR mother*:ti,ab,kw OR parent*:ti,ab,kw OR ‘birth Interval’:ti,ab,kw OR ‘birth Spacing’:ti,ab,kw OR ‘birth Spacings’:ti,ab,kw OR ‘interpregnancy’:ti,ab,kw OR 'pregnancy'/exp OR ‘perinatal period’/exp OR ‘parent’/exp OR 'pregnant woman'/exp OR ‘perinatal care’/exp

Cost-benefit
‘Economic evaluation’:ti,ab,kw OR ‘economic analys*’:ti,ab,kw OR ‘cost effective*’:ti,ab,kw OR costeffective*:ti,ab,kw OR ‘cost benefit*’:ti,ab,kw OR ‘cost utilit*’:ti,ab,kw OR ‘cost and benefit*’:ti,ab,kw OR ‘costs and benefit*’:ti,ab,kw OR ‘benefit and cost*’:ti,ab,kw OR ‘benefits and cost*’:ti,ab,kw OR ‘cost efficiency’:ti,ab,kw OR ‘Health Technology Assessment’:ti,ab,kw OR 'economic evaluation'/exp

Mental health
‘depression’:ti,ab,kw OR ‘anxi*’:ti,ab,kw OR ‘mental health’:ti,ab,kw OR ‘mental disorder’:ti,ab,kw OR ‘eating disorder’:ti,ab,kw OR anorexia:ti,ab,kw OR *bulimia:ti,ab,kw OR smoking:ti,ab,kw OR ((substance:ti,ab,kw OR alcohol:ti,ab,kw) AND (abuse*:ti,ab,kw OR addiction:ti,ab,kw OR dependence:ti,ab,kw)) OR ((drug*:ti,ab,kw OR tobacco:ti,ab,kw OR nicotine:ti,ab,kw OR amphetamine:ti,ab,kw OR cocaine:ti,ab,kw OR marijuana:ti,ab,kw OR narcotic*:ti,ab,kw) AND (abuse*:ti,ab,kw OR addiction:ti,ab,kw OR dependence:ti,ab,kw)) OR psycho*:ti,ab,kw bipolar:ti,ab,kw OR schizophren*:ti,ab,kw OR PTSD:ti,ab,kw OR ‘Post Traumatic’:ti,ab,kw OR posttraumatic:ti,ab,kw OR stress:ti,ab,kw OR

'mental disease'/de OR ‘perinatal depression’/exp OR ‘anxiety disorder’/exp OR ‘drug abuse’/exp OR ‘substance abuse’/exp OR ‘substance use’/exp OR ‘psychosis’/exp OR ‘bipolar disorder’/exp OR ‘schizophrenia’/exp OR ‘smoking cessation’/exp OR 'eating disorder'/exp OR 'psychological well-being'/exp OR 'mood disorder'/exp

3. WOS CORE COLLECTION

Pregnancy:
TS=(postpartum OR “post partum” OR postnatal OR “post natal” OR perinatal OR “peri natal” OR antepartum OR “ante partum” OR antenatal OR “ante natal” OR pregnan* OR gestation OR father* OR paternal OR matern* OR mother* OR parent* OR “birth Interval” OR “birth Spacing” OR “birth Spacings” OR “interpregnancy”)
Cost-benefit
“Economic evaluation” OR “economic analys*” OR “cost effective*” OR costeffective* OR “cost benefit*” OR “cost utilit*” OR “cost and benefit*” OR “costs and benefit*” OR “benefit and cost*” OR “benefits and cost*” OR “cost efficiency” OR “Health Technology Assessment”

Mental health
“depression “ OR “anxi* “ OR “mental health “ OR “mental disorder “ OR “eating disorder “ OR anorexia OR *bulimia OR smoking OR ((substance OR alcohol) AND (abuse* OR addiction OR dependence)) OR ((drug* OR tobacco OR nicotine OR amphetamine OR cocaine OR marijuana OR narcotic*) AND (abuse* OR addiction OR dependence)) OR psycho* OR bipolar OR schizophren* OR PTSD OR “Post Traumatic “ OR posttraumatic OR stress

Appendix 2: Full text articles excluded

Title	Year	Author	Reason
Cost-benefit analysis of varenicline vs. Existing smoking cessation strategies in pregnant women	2010	Barnard et al.	poster
Lifetime costs of perinatal anxiety and depression	2016	Bauer et al.	does not consider interventions (only economic impact of perinatal anxiety & depression)
Perinatal depression and child development: exploring the economic consequences from a South London cohort	2015	Bauer et al.	no economic evaluation
The clinical effectiveness, cost-effectiveness and acceptability of community-based interventions aimed at improving or maintaining quality of life in children of parents with serious mental illness: a systematic review	2014	Bee et al.	no economic evaluation
The cost-effectiveness of screening tools used in the diagnosis of foetal alcohol spectrum disorder: a modelled analysis	2019	Berrigan et al.	only newborn screening is evaluated
An experimental evaluation of the benefits and costs of providing fertility information to adolescents and emerging adults	2018	Boivin et al.	not directly related to mental health condition. no economic evaluation.
Mums 4 Mums: structured telephone peer support for women experiencing postnatal depression. Pilot and exploratory RCT of its clinical and cost effectiveness	2011	Caramlau et al.	study protocol

Relapse prevention in UK Stop Smoking Services: current practice, systematic reviews of effectiveness and cost-effectiveness analysis	2010	Colemman et al.	no specific evaluation for pregnant smokers
The SNAP trial: a randomised placebo-controlled trial of nicotine replacement therapy in pregnancy--clinical effectiveness and safety until 2 years after delivery, with economic evaluation.	2014	Cooper et al.	same evaluation as included study of Essex et al.
243: The cost-effectiveness of counselling interventions for young women at-risk of perinatal depression	2020	Franta et al.	poster
Early Start A Cost-Beneficial Perinatal Substance Abuse Program	2012	Goler et al.	no economic evaluation (cost savings calculated)
A culturally tailored intervention to reduce risk of alcohol-exposed pregnancies in American Indian communities: Rationale, design, and methods	2021	Hanson et al.	study protocol
The Cost-Effectiveness of Parent-Child Interaction Therapy: Exmining Standard, Intensive, and Group Adaptations	2021	Hare et al.	young children are older than our target group
Improving infant sleep and maternal mental health: a cluster randomised trial	2007	Hiscock et al.	no economic evaluation
A model for cost-effectiveness analyses of smoking cessation interventions applied to a Quit-and-Win contest for mothers of small children.	2005	Johansson et al.	pre-school children (not within 2 years after birth)
Comparing the Acceptability, Clinical-, and Cost-effectiveness of Mental Health E-screening to Paper-based Screening in	2015	Kingston et al.	poster

Pregnant Women: a Randomized, Parallel-group, Superiority Trial			
Training health visitors in cognitive behavioural and person-centred approaches for depression in postnatal women as part of a cluster randomised trial and economic evaluation in primary care: the PoNDER trial	2011	Morell et al.	no economic evaluation
Effectiveness and cost-effectiveness of an electronic mindfulness-based intervention (eMBI) on maternal mental health during pregnancy: the mindmom study protocol for a randomized controlled clinical trial	2020	Müller et al.	study protocol
Postpartum depression screening in Alberta, Canada: a cost effectiveness analysis using administrative data	2017	Premji et al.	poster
Setting the Boundaries for Economic Evaluation: Investigating Time Horizon and Family Effects in the Case of Postnatal Depression	2018	Ride et al.	no specific intervention evaluated
Lithium use during pregnancy for bipolar disorder: a cost-effectiveness analysis	2020	Saito	poster
Group cognitive behavioural therapy for postnatal depression: a systematic review of clinical effectiveness, cost-effectiveness and value of information analyses	2010	Stevenson et al.	same study as other paper of Stevenson et al. 2010
The Benefits of Family Action: An Economic Assessment of the potential Benefits from Family action Interventions for Women at Risk of Perinatal Depression	2014	Taylor et al.	no economic evaluation
A cost effectiveness analysis of midwife psycho-education for fearful pregnant women	2017	Toohill et al.	same data as the study of Turkstra et al. (Turkstra et al.

- a health system perspective for the antenatal period			was preferred because of the use of outcomes in terms of health related quality of life)
Cost-utility analysis of a one-time supervisor telephone contact at 6-weeks post-partum to prevent extended sick leave following maternity leave in The Netherlands: results of an economic evaluation alongside a randomized controlled trial	2011	Uegaki et al.	not directly related to one of the specific defined mental health conditions
Costs of a Motivational Enhancement Therapy Coupled with Cognitive Behavioral Therapy versus Brief Advice for Pregnant Substance Users	2014	Xu	no economic evaluation (only cost analysis)

Appendix 3: inflation and currency conversion

Reported costs were first adjusted to a target price year using the OECD Consumer Price Indices (90), in order to take into account general inflation in the country of the study. If two base-years were reported, the average of both annual rates was calculated. If no price year was reported, the year the study was reported to be 'received' was considered. Second, these price-year adjusted costs were converted to euros based on Purchasing Power Parities (91).

Author	**Country**	**Currency - price year**	**CPI in price year (2015=100) [1]**	**CPI 2019**	**PPP**
Ammerman et al.	US	$ 2013	98,3	107,9	1
Asper et al.	Sweden	€ 2013	100,2	105	0,7
Bacheller et al.	US	$ 2020	109,2	107,9	1

Barlow et al.	UK	£ 2016	101	107,8	0,7
Bell et al.	UK	£ 2013	98,2	107,8	0,7
Boath et al.	UK	£ 1992-1993	62,7	107,8	0,7
Boyd et al.	UK	£ 2013	98,2	107,8	0,7
Campbell et al.	New Zealand	NZD 2006-2007	84,3	105,8	1,5
Chambers et al.	AUS	AUSD 2020	107,8	106,9	1,4
Dornelas et al.	US	$ 2002	75,9	107,9	1
Dukhovny et al.	Can	CAD 2011	94,7	107,4	1,2
Eldar-Lissai et al.	US	$ 2018	105,9	107,9	1
Essex et al.	UK	£2009-2010	89	107,8	0,7
French et al.	US	$1998	68,8	107,9	1
Gifford et al.	US	$2006	85,1	107,9	1
Grote et al.	US	$ 2013	98,3	107,9	1
Henderson et al.	UK	£2003-2004	77,25	107,8	0,7
Jones et al.	UK	£ 2014-2015	99,8	107,8	0,7
Morell et al.	UK	£ 1996	68,5	107,8	0,7
Morrell et al.	UK	£ 2003-2004	77,25	107,8	0,7
Mundt et al.	US	$ 2020	109,2	107,9	1
Naughton et al.	UK	£ 2014-2015	99,8	107,8	0,7
NCCMH	UK	£ 2013-2014	98,9	107,8	0,7
Parker et al.	US	$ 2006 (assumed)	85,1	107,9	1
Paulden et al.	UK	£2006-2007	82,35	107,8	0,7
Petrou et al.	UK	£ 2000	73,4	107,8	0,7
Pollack et al.	US	$ 1998	68,8	107,9	1
Premji et al.	Can	CAD 2019	107,4	107,4	1,2
Premkumar et al.		$ 2017	103,4	107,9	1
Ride et al.	AUS	AUSD 2013-2014	97,3	106,9	1,4
Robin et al.	US	$ 2020	109,2	107,9	1

Ruger et al.	US	$ 1997	67,7	107,9	1
Saing et al.	Aus	AUSD 2014-2015	99,25	106,9	1,4
Stevenson et al.	UK	£ 2007- 2008	84,75	107,8	0,7
Thanh et al.	Can	CAD 2013	97	107,4	1,2
Trevillion et al.	UK	£ 2015-2016	100,5	107,8	0,7
Turkstra et al.	Aus	AUSD 2013	96,1	106,9	1,4
Ussher et al.	UK	£ 2012-2013	97,1	107,8	0,7
Wilkinson et al.	US	$ 2014	99,9	107,9	1

Appendix 4: Quality assessment

Study	Was a well-defined question posed in answerable form?	Was a comprehend-sive description of the competing alternatives given?	Was the effectiveness of the programmes or services established?	Were all the important and relevant costs and consequences for each alternative identified?	Were costs and consequences measured accurately in appropriate physical units prior to valuation?	Were costs and conseque n-ces valued credibly?	Were costs and conseque n-ces adjusted for differenti al timing?	Was an increment al analysis of costs and conseque n-ces performed?	Was uncertainty in the estimated of costs and consequences adequately characterized?	Did the presentation and discussion of study results include all issues of concern to users?	Total score (on a scale from 0 to 10)
Barlow et al.	1	0	1	0	1	0	1	1	0	0	5
Ammerman et al.	1	1	1	0	1	1	1	1	1	1	9
Asper et al.	1	1	0	0	1	1	1	1	1	1	8
Bacheller et al.	1	1	1	0	1	1	1	1	1	1	9
Bell et al.	0	0	1	1	1	1	1	1	1	0	7
Boath et al.	1	1	0	0	1	1	1	1	1	1	8
Boyd et al.	1	1	1	0	1	1	1	1	1	1	9
Campbell et al.	1	1	1	0	1	0	1	1	1	0	7
Chambers et al.	1	1	1	0	1	1	1	1	1	1	9
Dornelas et al.	1	1	1	0	1	0	1	1	0	0	6
Dukhovny et al.	1	1	1	0	1	1	1	1	1	1	9
Eldar-Lissai et al.	1	1	1	0	1	0	1	1	1	1	8

Essex et al.	1	1	0	0	1	1	1	1	1	0	7
French et al.	1	1	1	0	1	0	1	1	1	1	8
Gifford et al.	1	1	0	0	1	0	0	0	0	0	3
Grote et al.	1	1	1	0	1	1	1	1	1	1	9
Henderson et al.	1	0	1	0	1	1	1	1	1	1	8
Jones et al.	1	1	1	0	0	1	1	1	1	1	8
Morell et al.	1	1	1	0	0	0	1	1	0	0	5
Morrell et al. (2009)	1	0	1	0	1	1	1	1	1	1	8
Mundt et al.	1	1	0	0	1	0	1	1	0	1	6
Naughton et al.	1	1	1	0	1	1	1	1	1	1	9
NCCMH	1	1	1	0	0	0	1	1	1	1	7
Parker et al.	1	1	1	0	0	1	1	0	1	1	7
Paulden et al.	1	1	0	0	0	0	1	1	1	1	6
Petrou et al.	1	1	1	0	1	1	1	1	1	1	9
Pollack et al.	0	0	0	0	1	0	1	1	0	0	3
Premji et al.	1	1	1	0	1	1	1	1	1	1	9
Premkumar et al.	1	1	1	0	1	1	1	1	1	1	9
Ride et al.	1	0	1	0	1	1	1	1	1	1	8
Robin et al.	1	1	0	0	1	1	1	1	1	1	8
Ruger et al.	1	1	0	0	1	1	1	1	1	1	8
Saing et al.	1	1	0	0	0	0	0	1	1	1	5

Stevenson et al.	1	1	1	0	0	1	1	1	1	1	8
Thanh et al.	1	1	1	0	1	1	1	1	1	1	9
Trevellion et al.	1	1	1	0	1	1	1	1	1	0	8
Turkstra et al.	0	0	1	0	1	1	1	1	1	0	6
Ussher et al.	1	1	1	0	1	1	1	1	0	0	7
Wilkinson et al.	1	1	0	0	1	0	1	1	1	1	7

Part 2: Public funding of Assisted Reproductive Technologies

Abstract

Subfertile couples have a wide range of assisted reproductive technologies (ART) at their disposal. However, there are wide international differences in how these programs are funded within public health insurance and for many people coverage largely determines access to these interventions. In this chapter we discuss the question which ART programs should be state-funded and why. We evaluate the role of health-economic evaluation in resource allocation for ART and summarize recent cost-effectiveness research on three common ART-techniques: IVF, ICSI and cryopreservation of embryos, estimating the incremental gain for the costs incurred by these programs. Whereas these studies offer important information to decision-makers that have to allocate limited budgets, the methods are also ill-equipped to estimate the full value-for-money of ART. We explain how broader costs and benefits need to be considered, that behavioral changes of patients and providers need to be taken into account and that important equity dimensions of ART funding are neglected in cost-effectiveness analysis. An important challenge for the future is to develop appropriate methods that can help policy makers in determining the scope of resources that should be allocated to which ART programs.

This subchapter has been published as:

Verbeke, E., Luyten, J., D'Hooghe, T. (2020). The economics of IVF: evaluating the necessity and value of public funding. In: *Patient-Centered Assisted Reproduction: How to Integrate Exceptional Care with Cutting-Edge Technology.*, (106-135). Cambridge: Cambridge University Press. ISBN: 9781108796774.

1.2.1 Overview of public funding of ARTs (in selected countries)

Infertility is recognized by the WHO as a condition leading to disability, and it is widely acknowledged that patients have a right to treatment.(1) One of the treatment options is Assisted Reproductive Technology (ART). The International Glossary on Infertility and Fertility care defines procedures of ART as "all interventions that include the in vitro handling of both human oocytes and sperm or of embryos for the purpose of reproduction. This includes, but is not limited to, IVF and embryo transfer (ET), intracytoplasmic sperm injection (ICSI), embryo biopsy, preimplantation genetic testing (PGT), assisted hatching, gamete intrafallopian transfer (GIFT), zygote intrafallopian transfer, gamete and embryo cryopreservation, semen, oocyte and embryo donation, and gestational carrier cycles. Thus, ART does not, and ART-only registries do not, include assisted insemination using sperm from either a woman's partner or sperm donor."(1) This right to treatment and the wide scope of ART that is available to patients is however confronted with limited and often already pressurized healthcare budgets. The unavoidable question is therefore which ART programs should be funded by public health insurance and which ones can be referred to private resources or private health insurance? This chapter discusses the public funding of ARTs and hence the access to them by large segments of the population.

Even though most common ARTs are technically available worldwide, substantial differences in access exist between and within countries (Figure 1). The most generous funding policies exist in Europe and the resulting highest utilization rates are found in Denmark, the Czech Republic and Belgium. In Denmark 15,449 ART cycles per million females of reproductive age were performed, compared to 3844 cycles in Portugal. With a need for fertility treatments that is likely to be comparable over these countries, these large differences illustrate the impact of regulation and public funding on access to ARTs and hence the importance of finding a fair and sustainable funding scheme.

Figure 1: Data for 2014, for 12 European countries of which complete ART data are available. ART cycles include: IVF, ICSI, FER, egg donation (ED), IVM, frozen oocyte replacement (FOR) and PGD combined. For women 15-45 years old.(2)

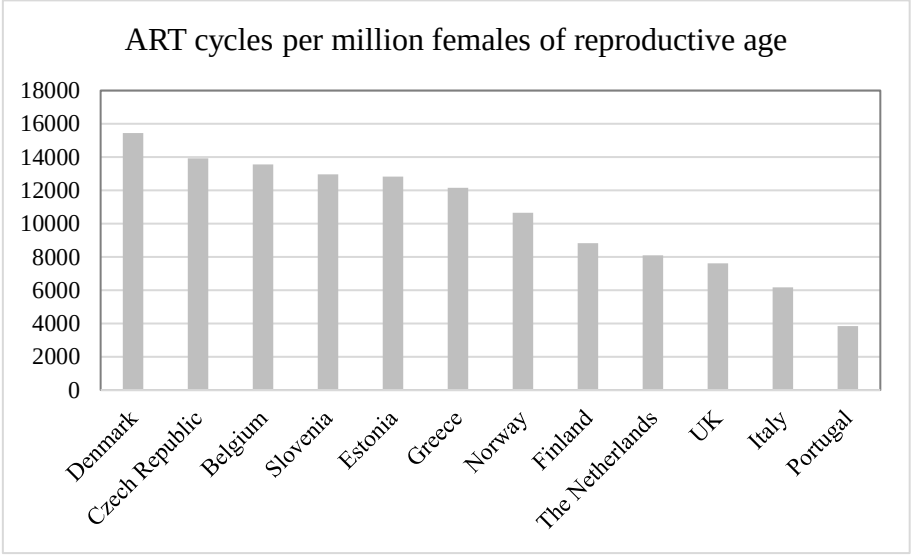


Table 1 (included at the end of this subchapter) provides an overview of funding policies for the three most frequently used ARTs in Europe (in terms of treatment cycles in 2014): ICSI (46,6%), followed by frozen-thawed embryo transfer (FET) (24,7%) and IVF (18,8%).(2) Reported in this table are the European Countries included in the IFFS and ESHRE reports and, in order to provide a broader comparison, six non-European countries (China, India, Israel, Japan, Russia and the United States).

Ireland, Poland (since 2016) and Switzerland are the only European countries not providing any public coverage for ART programs. Of the countries that do offer reimbursement, all of them cover IVF. ICSI is not covered in the Czech Republic, Romania and Slovakia. However, 'complete coverage' does not imply that costs are fully reimbursed for *all* patients. Additional restrictions often apply: restrictions can be related to the patient's marital status, a maximum number of reimbursed cycles can be stated ranging from one cycle (Romania) to six (Belgium, Croatia and Japan) and most countries impose an age limit for women, most commonly between 40 and 45 years. This age limit for

reimbursement can differ from the legal age limit to undergo fertility treatment. In Belgium for example, the legal age limit for egg retrieval is 45 and 48 for embryo transfers whereas these treatments are only reimbursed when the patient is younger than 43 on the day of the insemination of the eggs.(3)

These substantial discrepancies in public insurance coverage and usage of ART are the result of differences in priority setting and resource allocation amongst countries. How can policymakers decide which ART programs to fund, to which extent (fully or partially?), on which scale (for whom?) and how much money should in total be spend on ARTs? In this chapter, we explain some of the answers to these questions offered by cost-effectiveness analyses and their limitations in prescribing a fair and efficient public funding scheme.

1.2.2 Health-economic evaluation of ARTs

1.2.2.1 Available methods

Important first answers to the question which ART to fund are offered by economic evaluation studies. These analyses systematically compare two or more interventions on their costs and expected outcomes and express the value of the more promising option in terms of an *Incremental Cost-effectiveness Ratio (ICER)*: the additional costs required by a program per effect gained. Different types of economic evaluation exist, with the unit in which the health effects are expressed determining the type of evaluation. The broadest form of evaluation is Cost-Benefit Analysis (CBA), expressing outcomes in monetary terms (e.g. the money equivalent of a clinical pregnancy achieved) that can be compared to the costs incurred. The advantage of this approach is that everything is expressed in money terms and this allows incorporating *all* possible sources of value of the outcome (not just clinical aspects of an ART but also its impact on e.g. wellbeing to parents or economic benefits of having a child). The estimated return-on-investment can then be

compared to the money returns of any other way of spending available resources. The disadvantage is that such a broad monetization of outcomes (especially of those that are typically kept out of the economic sphere) is considered ethically-sensitive, highly challenging and available estimates often lack reliability and validity.

Alternative evaluation techniques are cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). In CEA, the outcomes considered are restricted to particular natural units of effects gained by the program. The ICER will hence be reported in terms of e.g. a cost per live birth achieved, without specifying how valuable a live birth is. The advantage is its straightforwardness and, often, such estimates are meaningful to clinicians or decision-makers who have to allocate a ring-fenced budget (e.g. a fixed budget only to be used for ARTs). The disadvantage is that cost-comparisons are only possible within the sphere of the particular effect that was considered (e.g. per live birth/ per clinical pregnancy or per cycle) and a consensus on the most meaningful denominator for ARTs is currently lacking.(4) Furthermore, cost-effectiveness comparisons of interventions that operate in different healthcare domains, and where different health outcomes are 'produced', becomes impossible. This means that from cost-effectiveness estimates we can judge which intervention is the most 'technically efficient' one (producing desired outcomes at minimal cost), but we cannot judge whether the overall ART budget size is appropriate, and whether from a more global perspective on healthcare funding, resources are used there where they create the most value (i.e. 'allocative efficiency').

A third evaluation technique is somewhere in between CBA and CEA in terms of scope: cost-utility analysis (CUA). Here, health effects are translated into 'utilities' such as a *Quality-Adjusted Life Years (QALY's)*. A QALY is a generic 'unit' of health, representing one life year in perfect health. Health effects of an intervention will thus consist of a certain number of QALYs gained, determined by the expected increase in life years, adjusted for the expected quality-of-life during those years. Quality of life-weights are attributed to each year ranging from zero for

a life that has the same value as death and one for a life in perfect health. CUA is the commonly-used method in the economic evaluation of healthcare programs, as it allows cost-effectiveness-comparisons of interventions in different health domains in terms of €/QALY, whilst avoiding the problems of monetizing health in CBA.(5)

1.2.2.2 Cost-effectiveness of IVF, ICSI and cryopreservation of embryo's

We summarized 18 recent cost-effectiveness (CE) studies in Table 2 (included at the end of this subchapter). This table is based on a targeted (rather than systematic) review, selecting recent economic evaluations related to the assessment of cost-effectiveness of IVF, ICSI and cryopreservation of embryo's (as these subjects were also included in section 1). This search was conducted for articles published between 2009 and 2019, on Pubmed and in high-impact journals focusing on fertility, human reproduction or health economics. We discuss the reported CE-results for the most common research questions and explore existing discrepancies in methodology.

Elective single embryo transfer (eSET) versus double embryo transfer (DET)

The most common focus was the CE of elective single embryo transfer (eSET), compared to double embryo transfer (DET). Do the higher success rates of DET justify the higher risk of multiple gestation and related higher healthcare costs? In general, success rates of ART decline when women get older, leading to a lower risk of multiple gestation and hence higher CE of DET. CE will consequently be dependent of the patient's age: according to Hernandez Torres et al. eSET is not significantly more cost-effective for women younger than 38 years.(6) However, Papaleo et al. (10) found eSET to be only cost-effective for women under 32 years, Roque et al. (11) reported 36 years as the age up to when eSET should be preferred. Fiddelers et al. (7) found eSET more cost effective only for a willingness-to-pay (WTP) per program below € 7350. Veleva et al. reported eSET to be more cost-effective in general.(8)

As will be discussed later in this chapter, the public funding scheme will determine the patient's and clinics' behavior in choosing the most cost-effective option from their perspective.

Freeze-only versus fresh embryo transfers

Freeze-all strategies in which all embryos are cryopreserved and transferred with delay, is currently preferred for patients with high risk of ovarian hyperstimulation syndrome (OHSS), the need for pre-implantation genetic diagnosis (PGD), pre-implantation genetic screening (PGS) and impairment in endometrial receptivity owing to serum progesterone elevation during ovarian stimulation.(9) Improvements in cryopreservation techniques have led freeze-all strategies to be an interesting alternative, as it has been proposed that the embryo transfer can be performed in a more favorable intrauterine environment, possibly improving ART outcomes and hence increasing popularity of eSET compared to DET.(10) A freeze-all strategy might however also imply additional costs related to embryo cryopreservation, endometrial priming, extra medication use and ultrasound scanning.(11) CEA provides the possibility of weighing these potentially higher success rates against the potential increase in costs.

Success rates of the freeze-all group were indeed found to be higher in a number of studies (9–11). Papaleo et al. also reported a lower average cost for the freeze-all group and hence found it to be more cost-effective.(9) Roque et al. reported a higher average cost per patient for the freeze-all strategy, but still identified freeze-all as the most cost-effective option because of the higher success rates.(10) Apart from other study-specific factors, an important difference between studies is the observed population: Roque et al. only considered patients with a progesterone level on trigger day > 1.5 ng/mL, Papaleo et al. also included patients for OHHS risk, detection of sacto-and hydrosalpinx and suspected endometrial pathology. Because freeze-all strategies are not yet used for routine IVF treatments in all patient populations,

broader evidence of ‘random patients’ is still limited. Le et al. evaluated cost-effectiveness based on one of the already available randomized controlled trials (RCT’s) for women without polycystic ovary syndrome (PCOS).(11) Success rates and costs for both strategies were similar for both randomized groups. Hence, the probability of freeze-all to be more cost-effective was found to be very low. In summary, a freeze-all strategy was found to be cost-effective for particular patient groups, but evidence of success rates and hence cost-effectiveness for more general groups is still very limited. Furthermore, a potential decrease in indirect costs of freeze-all because of lower maternal and perinatal morbidity was not considered in these studies, which might benefit the freeze-all case when considering cost-effectiveness in the long run.(9)

Immediate versus delayed IVF

Temporarily postponing IVF might be cost-effective as natural conception might occur during this delay, naturally or by the use of cheaper, less invasive techniques such as an improved diet or increased physical activity. The cost-effectiveness of delaying IVF depends on the type of infertility and the age of female patients (as success rates of young women decrease less rapidly). Pham et al. reported a substantial decrease in out-of-pocket and Medicare costs by delaying IVF with six months for couples with unexplained infertility and a good prognosis for natural conception, while birth rates were minimally impacted.(12) Eijkemans et al. considered a delay of 12 months and concluded that couples with unexplained fertility and female partner younger than 31 should wait 3 years before considering IVF.(13) For patients with endometriosis however, the cost-effectiveness of delaying IVF was found to be low and only slightly dependent on the duration of fertility and female age.

1.2.2.3 Methodology used in reviewed CEA's

The types of costs that were considered, the time horizon in which costs and effects are measured and the effectiveness measure used can all have an important impact on the study results. The available literature did not always account for the uncertainty regarding these issues. *Costs* were typically reported in terms of average cost per cycle or per patient (for the total number of cycles). The types of costs included are however very divergent and often not clearly specified. Costs considered in the selected articles included: direct medical costs associated with ART treatment, pregnancy, childbirth or neonatal care, travel expenses and income lost due to visits. General guidelines, determining which of these and other costs to include, would facilitate direct comparisons of costs between different studies. The longest *time horizon* considered was 6 weeks after delivery (e.g.(10), (12) and (19)), with the exception of (16) considering a time horizon of 20-years. This in general very limited time horizon might lead to suboptimal decision-making, as important long term costs (borne by patients or the public health care system) are neglected. Also, all selected articles reported *effectiveness* in natural units and are hence to be labelled as CEA's. (17) reported the number of ongoing pregnancies, and (10,18) the number of clinical pregnancies, all other studies reported outcomes in terms of live birth per cycle or cumulated live birth rate. Only Scotland et al. considered health effects for the mother and expressed outcomes in terms of QALYs.(16) A QALY loss was for example included when mothers had a disabled child or were infertile with the desire of having a child. The authors did not specify whether QALYs related to newborns were also taken into account.(16) Not including these QALYs however implies that a loss of quality-of-life of the newborn (related to e.g. prematurity or multiple gestation) would not be considered. This absence of information illustrates the vagueness of QALY-usage in the context or ART, which we will discuss more extensively in section 3.1. A uniform effectiveness measure (the unit in which the ICER is reported) would facilitate comparisons between studies but this requires consensus on the most meaningful

denominator. In general, a cost per success (whether this is a live birth, clinical pregnancy or ongoing pregnancy) is preferred, enabling the assessment of a completed program in terms of the desired outcome. Alternatively, evidence regarding the cost-effectiveness of one additional cycle could also be valuable. Taking into account decreasing marginal success rates (the success rate of every additional cycle will go down), might offer valuable information for policymakers when determining the maximum number of reimbursed cycles. However, evaluating a cost per cycle is ill-suited for the assessment of cryocycles (frozen embryo transfers) as in this case only cumulative success rates provide meaningful information. A possible solution may be to calculate a cost per oocyte retrieval.(4) These last two alternatives might however both give rise to behavioral changes when competition amongst providers exists, which we will discuss in section 2.3.2.

1.2.3 Limitations of cost-effectiveness studies

1.2.3.1 Broader Costs and benefits of ARTs

CEA was the dominant technique found to be used in evaluating ARTs. In general, cost-effectiveness studies only consider a limited variety of costs and effects and this for a limited period of time. In order to be able to assess the full value of ART, i.e. the value of an additional child, policy-makers need to be aware of the much broader cost- and benefit-dimensions.

On the cost side, ART treatment leads to various *indirect costs* such as costs for psychological assistance or productivity loss during medical visits. These costs occur during, but also after treatment and are often not included as analyses are generally performed from a more narrow, payer perspective. Estimating productivity loss is difficult and can be country- and patient-dependent (according to the social security system, occupation of the parents). Moreover, should these costs be corrected for the fact that a newborn will itself become a productive factor? Including these broader costs can hence lead to complex and incomparable

evaluations. But, more fundamentally, ARTs can also give rise to *long-term costs* as significant medical, social or educational costs might manifest related to e.g. maternal complications, prematurity, twin births or disability or other health problems of the newborn.(19) These costs occur throughout the life of mother and child, posing a significant burden on patients or society. For instance, there is some emerging evidence that ARTs in obese women leads to substantial and long-during complications in the newborn.(20)

On the benefit side, commonly reported outcomes are 'live birth' or 'clinical pregnancy achieved'. CEA-evidence hence allows the ranking of alternative treatments, but does not provide an answer to how valuable a birth or pregnancy is and, hence, how much of the available budget should be spend on ART's. Some have attempted to expand CEA to a CBA perspective by calculating a monetary value to society of an additional child based on proxies such as e.g. expected tax revenues. Connolly et al. calculated a lifetime discounted value of net taxes from an IVF-conceived child of £109,939 compared to £122,127 for a naturally conceived child.(21) A CBA perspective might appear specifically relevant for societies where population growth is below replacement level. However, taking into account the 'expected societal revenues' of the newborn, holds an important risk of discrimination. Calculating expected tax revenues would imply estimating the future educational or income-level of the newborns, leading to a valuation of the potential newborn that will most likely depend on the socio-economic background of their parents. Alternatively, a parental WTP perspective can be applied: how much money are people willing to pay for a newborn child? Settumba et al. estimated the mean WTP of the general Australian population for one cycle of IVF to be AUS$ 6135-$13,561 (depending on the success rate of IVF).(22) An Israeli study, reported a WTP per cycle of $5482 for patients undergoing IVF and $4398 for the general public.(23) These values do however not reflect an objective valuation of a live birth, as results will be influenced by various

factors such as the perspective applied, the features of the survey and the current national health insurance policy. Using a single monetary value for every live birth is hence a very abstract approach where significant amounts of information will be lost. (22,23) On the other hand, a live birth can also be translated into QALYs, but a consensus on how this can be done is lacking. Some authors argue that the use of QALYs is not appropriate, as the conception of a new life cannot be considered an improvement of health.(24,25) If QALYs would be used to account for a healthy birth, ARTs would become extremely cost-effective compared to treatments in other health domains as it would produce a 'life-time' amount of QALYs for the newborn. The creation of QALYs per se would hence be considered a reason to bring someone into existence, leading to large investments in ARTs while neglecting the health of already existing people.

Lastly, although often forgotten in CEA's, undergoing ARTs influences the *wellbeing of the parents* substantially. Failure of treatments might have severe psychological consequences. Achieving a live birth can increase happiness but having children might however again decrease life satisfaction in the long term.(26) In theory, these consequences can be included in CEA by using QALYs. Scotland et al. illustrated this approach by including maternal QALYs depending on the outcome of the ART treatment. (The wellbeing of the partner was by contrast not considered.) They however also illustrated the significant impact of a change in QALY-weights on reported outcomes. This demonstrates the importance of a transparent and systematic use of QALYs, currently hampered by the lack of quality-of-life information related to infertility problems and ART.(27)(16) Furthermore, the question also rises whether psycho-social consequences of fulfilling or not fulfilling the desire to have children can be captured sufficiently by the QALY-approach.(24)

Considering these much broader effects of ARTs raises the currently-debated question whether ARTs fulfill a purely medical need (curing

infertility) or a mere desire to have children and hence should be covered by public health insurance or not.(26,28) Whereas this issue is fundamental to the question how to fund ART, it is in any case hard to argue that only purely clinical effects should be considered in funding decisions as health interventions always serve multiple purposes beyond generating health itself and, moreover, non-clinical problems (e.g. an unfulfilled child wish) can lead to medical problems and associated costs more downstream (e.g. depression). Cost-effectiveness evidence can unfortunately not provide an answer to these more fundamental questions, even when a method to systematically include broader costs and benefits is found.

A broader, related point that should be made in the context of evaluating ARTs and their broader costs and benefits is that what counts as a cost or benefit also depends upon the underlying philosophical view on appropriate population size. Are more people better? We should be aware that cost-effectiveness studies of health interventions (especially CEA) are in general based on the ethical perspective of total utilitarianism, in which every live birth is perceived as a gain to society. In reality, this might not be the case, especially in countries with overpopulation. The field of population ethics provides interesting alternative perspectives based on what counts as an optimal population size. A discussion of these would however go beyond the scope of this chapter.

1.2.3.2 Behavioral responses to ART funding schemes

A second aspect that is not considered in CEA's, is the impact of behavioral responses as a result of changes in public funding. These unanticipated behavioral factors may, when not included, alter the assumptions underlying CEA and affect the resulting cost-effectiveness estimates.

Effects of price-elasticity. Changes in public funding might give rise to changes in out-of-pocket costs and will, hence, influence demand. When the patient's demand is *inelastic*, a percentage change in price leads to a smaller-than-one-percent change in the number of treatments demanded. When demand is *elastic*, a percentage change in price gives rise to a larger change in demand. Chambers et al. estimated the price elasticity of ART-demand in developed countries and reported a relatively elastic demand when prices were in the mid-range and a less elastic demand in the upper and lower price ranges. When out-of-pocket costs were considerably high (e.g. in the United States), a small decrease in price did not have a significant impact on demand.(19) Decision-makers want to be aware of these possible changes in demand and their related economic consequences. For illustration, Connolly et al. estimated the change in IVF/ICSI-demand after the change of German policy in 2004, introducing a 50% co-payment for patients. The short-term (12 months) price elasticity of demand after introducing patient payments of €1500-€2000 was estimated to be -0.36. (A 10% price increase results in a reduction of demand with 3.6%.)(29)

Moral hazard and delayed IVF. Secondly, generous reimbursement of ARTs might encourage patients to undergo more ART cycles (even when expected success rates are limited) or to choose for ART without first considering other less invasive methods. [15] and [16] illustrated the possible cost reductions by delaying IVF for young women with unexplained fertility without compromising expected live birth rates. On the other hand, delaying ART for women with advanced age is not recommended, as success rates more rapidly decrease. Higher age limits for women hence hold the risk that older women will also postpone treatment, leading to lower effectiveness of the reimbursed cycles.(12,13), (30)

Cross-border reproductive care. Existing national differences in regulation and public funding of ART gave rise to the increased use of cross-border reproductive care (CBRC), in which patients seek ART

abroad. Slovakia is now the first country even initiating public insurance coverage of CBRC.(31) The main reasons patients travel abroad for ART are lower costs, higher quality of treatments or limited access to services (because of e.g. income, sexual orientation or advanced age). Popular countries for CBRC are the USA, Israel, Spain, Czech Republic, Denmark and Belgium. Even though CBRC provides alternative opportunities to fulfill a child-wish, it also holds some important risks. Firstly, CBRC is often facilitated by 'brokers', connecting foreign patients online with local ART centers. Brokers might have other financial objectives and language barriers hamper communication with local providers. The resulting asymmetrical information can lead to supplier-induced demand, encouraging unaware patients to undergo unnecessary or suboptimal treatments. DETs can for example be recommended to increase probability of a live birth without patients being aware of the risks and long-term consequences, which will be borne by the home country's health care system. Furthermore, increased demand of foreign patients (with a higher WTP) might decrease local access for patients, as they can be priced out of the market. Consequences of CBRC might hence be substantial for patients and governments. Current CBRC-regulation is however very limited and detailed information about the magnitude and scope of CBRC is lacking. Different international organizations such as the ESHRE and IFFS started data collection and are setting out general guidelines.(32,33)

Patients maximizing success rates. In the end, a patient's objective will be to maximize the probability of a live birth while minimizing out-of-pocket costs. When ART is not fully funded by public health insurance, patients are likely to choose the program with the highest probability of a live birth, while limiting the number of cycles (and their related costs). In the United States for example, ART treatment costs are typically borne by patients but pregnancy and infant-associated medical costs by health insurance. As a consequence, patients might be discouraged to choose eSET compared to DET because they only consider the lower

direct treatment cost of DET without the related medical 'twin'-costs borne by insurance.(34) Secondly, ART might be fully funded but only for a limited number of cycles. This again encourages patients to increase their success rates per cycle without considering long-term consequences. Lastly, we also need to consider the fact that some patients might just prefer giving birth to twins. A Danish study (n=588, 2004) reported 58.7% of patients preferring twins.(35) A smaller but more recent American study (n=106, 2017) found that 61% of patients that never gave live birth preferred a singleton and 39% preferred twins.(36)

Competition among clinics. Reported success rates are not only important for patients but also for providers when competition amongst ART clinics exists (e.g. in the context of CBRC). Reporting live birth rates per cycle encourages clinics to improve their reputation by increasing the number of DETs, disregarding the conclusions of cost-effectiveness studies we previously discussed. Attention for this problem has increased and eSET is now encouraged by raising awareness of patients and by adaptations of ART funding policies.(37) A higher uptake of eSET was for example found to be closely related to more generous ART funding schemes.(38) Countries can also specify the number of embryos allowed during embryo transfer in order to obtain reimbursement. Belgium was the first country to link reimbursement of laboratory costs to transfer of a restricted number of embryos (dependent on female age and rank of cycle), in order to reduce the risk of multiple gestation. Furthermore, a maximum number of six fresh cycles was determined, not including cryocyles and hence encouraging additional frozen embryo transfers. (3) Peeraer et al. demonstrated a 50% reduction in multiple pregnancy rates as a result of this policy implementation, without a negative impact on the cumulative delivery rate of patients.(39) As a consequence, in this context, calculating the cost per oocyte retrieval seems the most meaningful cost-effectiveness measure. But this again holds a risk as it might encourage more aggressive ovarian

stimulation when the number of oocyte retrievals is limited.(4). This risk needs to be managed taking into account the risk of OHSS by balancing the intensity of ovarian stimulation (gonadotrophin dosing), by the choice of ovulation triggering (GnRH agonist when OHSS risk is high), and the option of freezing all retrieved oocytes/embryos.

1.2.3.3 Equity

Lastly, CEA's do not consider distributional effects of public funding schemes. General global access to ART is increasing, but barriers to utilization still exist. We summarize three grounds on which ART funding can lead to the exclusion of specific patient groups and consequently to unequal, and potentially inequitable access.

Socio-economic status. Firstly, access to ART will depend on the socio-economic status of a patient. Access to financial means will determine whether the patient can pay out-of-pocket costs of publicly funded programs or supplementary treatments in private clinics. Additionally, access is determined by indirect factors such as access to information, flexibility at work and support from families and friends. Furthermore, the risk of infertility and hence the demand for ART is also related to socio-economic status, as it is determined by the work and living environment of patients and behavioral factors such as smoking, drug use or obesity.

Cultural differences. Secondly, cultural differences exist with regards to the perceived optimal family size and hence the importance of fertility. In developing countries for example, demand for ART is most prevalent, as infertility or childlessness causes a severe burden on individuals. Paradoxically, access in these countries is very limited as priority is given to other health domains.(40) The question rises whether parents should be given equal opportunity to fulfill their child wish, or to achieve their desired family size. In the first case, we can argue that achieving

one healthy birth is sufficient to fulfill this wish and public funding can be limited to cases of primary infertility. Limiting the number of children a couple can have when seeking ART treatment seems reasonable, considering the limited health care budget. On the other hand, it will be difficult to justify a universal limit that will allow some cultural groups to achieve their desired family size and others not.

Sexual orientation. Lastly, barriers to utilization of ART based on the sexual orientation of patients are not uncommon (Table 3).(31) When patients do not have access to treatment, they will have to turn to (often more expensive) private providers or cross-border reproductive care. In 13 of the 32 selected countries, a stable heterosexual relationship is required. Discrimination based on sexual orientation might result from the idea that patients should only be provided the opportunity of achieving 'normal functioning'. Even when same sex couples are fertile, achieving a live birth is just not part of their normal functioning. On the other hand, one might consider same sex couples and single parents more eligible for reimbursement because they suffer the injustice of not being able to achieve the same family goals heterosexual couples might have.(26)

Table 3: Legal access to IVF or ART services, based on sexual orientation. (IFFS,2016)
(n.a.= no answer was stated.)

Country	Stable heterosexual relationship required	Access for single women	Access for single men	Access for female same sex couples	Access for male same sex couples
Austria	Yes	N.a.	N.a.	N.a.	N.a.
Belarus	No	Yes	No	N.a.	N.a.
Belgium	No	Yes	Yes	Yes	Yes
Bulgaria	No	Yes	No	Yes	No
China	Yes	N.a.	N.a.	N.a.	N.a.
Czech republic	Yes	N.a.	N.a.	N.a.	N.a.
Croatia	N.a.	N.a.	N.a.	N.a.	N.a.
Denmark	No	Yes	No	Yes	No
Estonia	No	Yes	No	Yes	No
Finland	No	Yes	Yes	Yes	Yes
France	Yes	N.a.	N.a.	N.a.	N.a.
Germany	No	Yes	No	Yes	No
Greece	No	Yes	No	No	No
Hungary	Yes	N.a.	N.a.	N.a.	N.a.
India	N.a.	N.a.	N.a.	N.a.	N.a.
Ireland	No	Yes	No	Yes	Yes
Israel	No	Yes	N.a.	No	No
Italy	Yes	N.a.	N.a.	N.a.	N.a.
Japan	Yes	N.a.	N.a.	N.a.	N.a.
Netherlands	No	Yes	N.a.	Yes	N.a.
Norway	Yes	N.a.	N.a.	N.a.	N.a.
Poland	N.a.	N.a.	N.a.	N.a.	N.a.
Portugal	Yes	N.a.	N.a.	N.a.	N.a.
Romania	No	Yes	No	No	No
Russia	No	Yes	No	N.a.	No
Slovak republic	Yes	N.a.	N.a.	N.a.	N.a.
Spain	No	Yes	N.a.	Yes	No
Sweden	Yes	N.a.	N.a.	N.a.	N.a.
Switzerland	Yes	N.a.	N.a.	N.a.	N.a.
Turkey	Yes	N.a.	N.a.	N.a.	N.a.
United Kingdom	No	Yes	Yes	Yes	Yes
United States	No	Yes	Yes	Yes	Yes

1.2.4 Conclusion

Differences in national health insurance coverage of ARTs give rise to significant variations in utilization rates between countries, illustrating the importance of a well-conceived funding scheme. This chapter discussed the challenges of setting out an optimal and justifiable public funding scheme for ARTs.

Economic evaluation evidence allows the ranking of alternative treatments based on their value-for-money, with value defined in either more narrow or broader terms. We summarized evidence of recent cost-effectiveness studies on ARTs. When interpreting outcomes however, we need to be aware of the differences between studies in terms of costs and time horizon considered and effectiveness measure used. CEA was the dominant technique found to be used in evaluating ARTs. It is currently unclear how CUA's should be conducted in this context, as it remains an open question if, and how, neonatal QALYs should be included in the analysis. Reaching consensus on a uniform denominator for cost-effectiveness studies of ARTs is not straightforward as broader consequences and the local policy framework need to be taken into account.

The narrow perspective of CEA has nevertheless some important limitations in prescribing funding. Firstly, policy makers need to consider the broader social value of ARTs and the broader range of relevant costs that can emerge. However, considering broader costs and benefits raises the question where to draw the line and how to determine to which extent ARTs fulfill a purely medical need or whether their goals are broader. This is important to determine as the answer here will affect which type of (broad or narrow) evaluation framework is most appropriate to prescribe public funding. Secondly, CEA's neglect behavioral reactions and the consequences of patients and providers trying to optimize their own outcomes and not choosing the most cost-

effective alternative from a societal perspective. Lastly, how can equal access be provided without discriminating groups based on socio-economic status, cultural differences or sexual orientation?

Including these broader dimensions in the evaluation of ARTs and in subsequent priority-setting choices poses a major challenge for policy-makers. A systematic and transparent method is needed to enable more consistent and uniform decision-making regarding resource allocation for ARTs, on a micro level to determine which particular ART programs should be funded and to which extent but also on a macro level to determine the scope of public resources that should go to ARTs as a whole.

Review tables

Table 1: overview of public health insurance coverage in selected countries (n.a.: not available)

Country	Public coverage regulated by federal/ national law (31)	Type of coverage provided (41)	Public insurance typically covers:(31)				Age limit for coverage(31) (age women)
			IVF (max. number of cycles)	ICSI	Cryopreservation of embryos for fertility preservation for medical indications	Cryopreservation embryos from IVF cycle	
Austria	Yes	Partial (~70 %) (42)	Yes	Yes	Yes	Yes	Yes (40, men 50)
Belarus	No	None	No	No	No	No	Yes
Belgium (3)	Yes	Complete	Yes (6)	Yes	No	Yes	Yes (43)
Bulgaria	Yes	Partial	Yes	Yes	No	Yes	Yes
China	No	None	No	No	No	No	N.a.
Czech republic (43)	Yes	Complete	Yes (4)	No	No	No	Yes (40)
Croatia (42)	N.a.	Partial	Yes (6)	N.a.	N.a.	N.a.	Yes (42)

Denmark	Yes	Partial	Yes	Yes	Yes	Yes	Yes
Estonia (42)	Yes	n.a.	Yes	Yes	No	Yes	Yes (40)
Finland	Yes	Partial	Yes	Yes	Yes	Yes	Yes
France (43)	Yes	Complete	Yes	Yes	Yes	Yes	Yes (43)
Germany (43)	Yes	Partial (50% + 25% federal states)	Yes	Yes	No	No	Yes (40, men 50)
Greece	N.a.	Partial	No	No	No	No	Yes (50 unrelated to reimbursement)
Hungary (42)	Yes	Partial (75%)	Yes (5)	Yes	No	Yes	Yes (45-49)
India	N.a.	None	No	No	No	No	N.a.
Ireland	No	None	No	No	No	No	No
Israel	Yes	Complete	Yes	Yes	Yes	Yes	Yes
Italy (43)	Yes	Partial (~65%)	Yes	Yes	No	Yes	Yes (50)
Japan	Yes	Partial	Yes (6)	Yes	No	Yes	Yes
Netherlands	Yes	Complete	Yes (3)	Yes	Yes	Yes	Yes (43)
Norway	Yes	Partial	Yes	Yes	Yes	Yes	No
Poland (43)	Yes	0	No	No	No	No	/
Portugal (42)	Yes	Partial	Yes (3)	Yes	No	Yes	Yes (40)

Romania (43)	Yes	Partial (up to 1375€)	Yes (1)	No	Yes	Yes	Yes (40)
Russia	Yes	Complete	Yes	Yes	No	No	No
Slovak republic (42)	Yes	N.a.	Yes	No	No	No	Yes (40)
Spain (43)	No	Complete	Yes	Yes	Yes	Yes	Yes
Sweden (43)	Yes	Complete	Yes (3)	Yes	Yes	Yes	Yes (40, men 46)
Switzerland	N.a.	No	No	No	No	No	/
Turkey	Yes	Partial	Yes	Yes	Yes	Yes	Yes
United Kingdom (43)	No	Partial	Yes (Scotland 3(44), Wales 2, Northern Ireland one 2-part cycle)	Yes	Yes	Yes	Yes (40)[1]
United States	No	No	No (6 – state specific)	N.a.	N.a.	No	N.a.

[1] Applicable in Scotland, Wales and Nothern Ireland; in England funding is region-dependent.

Table 2: Included cost-effectiveness studies of ARTs. (CE= cost-effectiveness, eSET= elective single embryo transfer, DET= double embryo transfer)

First author (year) country of publication	Focus of study	No. of observations	Costs	Effectiveness	Reported cost-effectiveness	Comments
Elective single embryo transfer (eSET) vs double embryo transfer (DET)						
Crawford (2016) USA (34)	Assessment of costs and outcomes of DET and projection of the difference in costs and outcomes had the double ET cycles been performed as sequential SETs.	10.001 cycles DET, 4129 cycles eSET, patients <35 years with no prior ART use consisting of cryopreservation	*Average total cost per cycle, including infant associated medical costs during first year of life:* $38,600 SET $58,100 DET [2006 USD]	*Cumulative birth rate* 68% SET 57.7% DET Multiple birth rates: 1.2% SET 24.7% DET	If 10,001 DETs would have been performed as sequential SETs, this would have saved $195 million and increased live birth rates with 10.3%. *(Own calculations:)* $56,765 per live birth SET $100,693 per live birth DET	Assumes no infants stillborn in a multiple birth, assumes same success rates for DET as the projected sequential SET group. Ignores higher risk of drop out in the case of SET.

Fiddelers (2009) The Netherlands (7)	The CE of 7 different strategies is calculated. Three options are combined: I) eSET in all patients II) eSET in good prognosis patients and DET in the remainder of patients (=STP) III) DET in all patients The 7 strategies consist of: 1) 3 eSET 2) eSET + 2 STP 3) eSET + STP + DET 4) eSET + 2 DET 5) 3 STP 6) STP + 2 DET 7) 3 DET	308 couples that started their first IVF cycle	*Expected cost per couple per strategy*: 1) € 16,381 2) € 16,655 3) € 17,092 4) € 17,440 5) € 16,999 6) € 17,444 7) € 18,046 [2003 €]	Mean live birth rate per strategy: 1) 39.6% 2) 43.3% 3) 43.3% 4) 45.7% 5) 47.5% 6) 49.9% 7) 53.4%	Combining several transfer policies was found not to be cost-effective. The choice between eSET, STP and DET will be determined by the willingsness to pay: WTP< €7350: 3x eSET is to be preferred €7350 < WTP < €15,250: 3x STP WTP > € 15,250: 3x DET	Costs from societal perspective including health care costs and costs outside health care sector. From 2 weeks before randomization up to 6 weeks after birth/ 2 weeks after last ovum pick-up.

Hernandez Torres (2015) Spain (6)	CE of IVF-ICSI with eSET followed by the transfer of cryopreserved embryo's (eSFET), versus DET.	121 women < 38y undergoing their first or second IVF cycle. 57 eSET +eSFET and 64 DET	*Average cost per patient:* €5641 eSET+ eSFET €5562 DET [2012 €]	*Cumulative live birth :* 38.6% eSET+ eSFET 42.19% DET group (Differences were not statistically significant)	The chance of eSET+eSFET being cost-effective is <50% for all threshold values. *(ICER is not reported because no alternative was more effective for a higher cost.)*	Costs include direct medical costs associated with AR treatment, pregnancy, childbirth and neonatal care.
Scotland (2011) UK (16)	Assessment of cumulative costs and consequences of eSET and DET in women commencing IVF with treatment ages 32,36 and 39 years.	6153 women, 10.511 fresh cycles, 3106 frozen cycles. Up to three full cycles per patient	Not stated	Not stated	*Compared to eSET, DET leads to additional costs per live birth:* £27,356 age 32 £15,539 age 39 *Additional cost per QALY of women for DET:* £28,300 age 32 £20,300 for age 39.=> eSET more CE for women ≤ 36 years old, older women: case by case decision	Time horizon of 20 years. In addition use of QALY's to assess the cost-utility of the alternatives. (Discountrate costs and effects: 3.5%) [2007 £]

Van Loendersloot (2017) The Netherlands (14)	CE of SET followed by an additional frozen-thawed (eSFET), compared to DET, in relation to female age.	3390 patients undergoing IVF/ICSI	*Cost per strategy:* 1) SET+ eSFET: €4,491 for age 30 - €3,623 for age 43 2) DET: €5,291 for age 30 - €3,627 for age 43 [2010 €]	*Cumulative live birth rate:* 1) SET + eSFET: 36.8% at age 30 -> 49.7% at age 40 2) DET: 34.1% at age 30 -> 11.7% at age 40	*ICER (DET compared to SET eSFET):* Age 30: €-29,032 Age 31: €-114,576 starting from age 32 DET becomes more cost-effective: Age 32: €50,348 -> Age 43: €53 SET followed by eSFET is dominant over DET for women under 32 years. After the age of 32 DET was more effective but also more costly.	Including direct medical costs, including cost of singleton/twin birth until 6 weeks after birth.
Veleva (2009) Finland (8)	CE of eSET compared to DET.	DET period ('95-'99) : 1359 fresh embryo cycles+589 frozen cycles eSET period ('00-'04): 107 fresh and 683 frozen cycles	*Total discounted costs per woman:* €4942 DET period €4584 eSET period [2008 €], discountrate 3%	*Cumulative live birth rate:* eSET period: 41.7% DET period: 36.6%	€19,889 euros saved per term live birth in the eSET period: eSET with cryopreservation is more effective and less expensive than DET.	Total treatment charges and medication costs for fresh and FET cycles until pregnancy test. Costs from complications ovarian stimulation were not included.

Freeze-only vs fresh embryo transfers						
Le (2018) Vietnam (11)	CE of freeze-only strategy compared to fresh embryo transfer after one complete IVF/ICSI cycle in women without polycystic ovary syndrome (PCOS)	782 infertile couples	Estimated total cost per couple: €3906 freeze-only €3512 fresh [2016 €]	Live birth rate after one completed cycle: Freeze only: 48,6% Fresh ET: 47,3%	Estimated cost per live birth: €8037/live birth for freeze only strategy €7425/live birth for fresh ET strategy	Data frome one single private IVF center in Vietnam. Costs include direct medical costs from randomization to delivery and treatment complications. Travel expenses and accommodation costs included. *Should be noted that these costs are significantly lower compared to other studies.*
Papaleo (2017) Italy (9)	CE of IVF freeze-all policy (entire cohort of embryos is cryopreserved) compared with fresh embro transfer (only supernumerary embryos are cryopreserved).	Patients: 67 freeze all policy, 189 fresh transfer	*Mean cost per patient:* Freeze-all: €6863 Fresh: €6952	*Cumulative live birth rate:* Freeze-all: 52.4% Fresh: 45,5%	*Mean cost per live birth:* Freeze all: €13,101 Fresh= €15,279 => Freeze-all more cost-effective although difference not significant	Only direct healthcare costs considered. Freeze-all patients: patients for OHSS risk, high progresterone levels on trigger day > 1.5 ng/mL, detection of sacto and hydrosalpinx and suspected

						endometrial pathology.
Roque (2015) Brazil (10)	CE of freeze-all cycles compared to fresh embryo transfers.	530 ICSI cycles: 351 fresh embryo cycles and 179 freeze-all cycles	*1) from perspective of private center policy and charges* Fresh: $7424 Freeze-all: $7598 *2) including cryopreservation and thawing costs from all patients in freeze-all group* Fresh: $7417 Freeze-all: $9020 [2015 USD]	*Pregnancy rate (clinical pregnancy after confirmation of fetal heartbeat after 7-8 weeks of gestation) per cycle:* 31.1% fresh group 39.7% freeze-all group	*Total cost per ongoing pregnancy:* *Cost scenario 1* Fresh group: $23,060 Freeze-all group: $19,157 Cost scenario 2 Fresh group: $23,040 Freeze-all group: $22,742 Freeze all policy is CE compared to fresh embryo transfer	Freeze all patients: progesterone level on trigger day > 1.5 ng/mL. Exclusion of patients with OHHS risk. Exclusion of indirect costs.

Delayed vs immediate IVF						
Eijkemans (2017) The Netherlands (13)	CE of immediate versus delayed IVF in relation to prognostic characteristics of the couple Scenario 1: wait 1 year and then undergo IVF for 1 year Scenario2: immediate IVF during 1 year + 1 year trying to conceive naturally	5962 couples with primary infertility	*Cost per scenario (2years)* *Unexplained infertility* Scenario 1: €6792 (age 30) - €6695 (age 38) Scenario 2: €7379 - €7331 *Endometriosis* Scenario 1: €6787 - €6759 Scenario 2: €6985-€7047 [2014 €]	*Live birth rate after scenario (=2 years)* Unexplained infertility Scenario 1: 55% (age 30) - 33,8% (age 38) Scenario 2: 55,1%-37,6% Endometriosis Scenario 1: 43,5%-24,6% Scenario 2: 44,2%-28,5%	*Discounted cost efficiency ratio (cost per live birth gained by immediate IVF)* *Unexplained infertility:* €56.500 (age 30) - €16.600 (age 38) *Endometriosis* €19.400 (age 30) - €8700 (age 38) Recommendation for couples with unexplained infertility and female partner <31y to wait 3 years before IVF can be considered.	Excluding frozen embryo's. Discount rate of 3,5% for costs and effects. Costs from societal perspective: direct and indirect medical and non-medical costs (treatment, delivery and neonatal period)
Pham (2018) Australia (12)	CE of delaying IVF for 6 months in couples with unexplained infertility compared to immediate IVF treatment.	8781 couples aged <40 years, 17,418 fresh and 10,230 frozen ETcycles	*Estimated out of pocket costs for 3 cycles IVF*: AUS $7125-$19,500 (Assuming results are reported in AUS dollars, not stated)	*Live birth rate:* women <30 years: 64% women age 35-39: 48,7%	*Potential out-of-pocket savings if 90% of couples delayed IVF:* AUS $4.7-12.2 million for 27,648 cycles (with total Medicare cost savings up to $15.1 million) Delaying IVF for 6 months was found to substantially decrease	

					costs without compromising live births.	

Other studies						
Groen (2013) The Netherlands (45)	CE of modified natural cycle IVF or ICSI (MNC) as an alternative for controlled ovarian hyperstimulation IVF or ICSI (COH). 3 scenario's: 1) 3 cycles MNC vs 1 cycle COH with SET or (DET) + subsequent transfer of cryopreserved embryos 2) 6 cycles MNC vs 1 cycle COH with strictly SET + subsequent transfer of cryopreserved embryos 3) 6 cycles MNC with minimized medication (hCG ovulation trigger only) vs 1 cycle COH with SET or DET + subsequent transfer of cryopreserved embryos	1994 MNC cycles and 392 fresh COH cycles with subsequent transfer of cryopreserved embryos.	*Total costs per patient:* MNC-IVF: €5415 MNC-ICSI: €5096 COH-IVF: €3221 COH-ICSI: €4167 [2009 €]	*Cumulative live birth rate* MNC-IVF: 21,9% MNC-ICSI: 26,2% COH-IVF: 21,9% COH-ICSI: 32,1%	*Costs per live birth:* MNC-IVF: €24,728 MNC-ICSI: €19,452 COH-IVF: €14,710 COH-ICSI: €12,980 -> COH dominates	Cost of singleton, twins included. Costs of medication, costs of treatment procedures, cost of ongoing pregnanties for singletons and twins including costs of preganncy, delivery and costs up to 6 weeks after delivery.

Messinger (2015) USA (17)	CE of tubal anastomosis (TA) to IVF in women who desire fertility after tubal ligation	2.256 tubal anastomosis procedures, number of IVF observations not specified .	Average cost not stated	*Ongoing pregnancy after TA*Age < 35y: 63%Age 35-40: 44%Age > 40: 5%*Ongoing pregnancy after IVF*Age < 35y: 40%Age 35-40: 28%Age > 40: 10%*Ongoing pregnancy after frozen IVF cycle*Age < 35y: 39%Age 35-40: 35%Age > 40: 21%	*Costs per ongoing pregnancy:*$16,466 - $223,482 after tubal anastomosis$32,902 - $111,679 for IVF=> For women < 40y, tubal anastomosis was found to be more CE, for women> 41y IVF was the most cost effective approach.	The charges for IVF included physician visits, ultrasound and laboratory evaluation, oocyte retrieval, intracytoplasmic sperm injection (ICSI), embryo transfer (ET), embryology fees, and all medications. [2014 USD]

Moolenaar (2011) The Netherlands (46)	CE of ovarian reserve testing in IVF. Scenario's: 1) no treatment 2) up to three cycles of IVF limited to women <41 years + no ovarian reserve testing 3) up to three cycles of IVF with dose individualization of gonadotropins according to ovarian reserve 4) up to three cycles of IVF with ovarian reserve testing and exclusion of expected poor responders after the first cycle	Computer-simulated cohort of subfertile women 20-45 years old, based on prospective cohort study following 4928 couples during 12 months.	*Absolute costs per woman per scenario:* 1) €0 2) €6917 3) €6678 4) €5892 [2008 €]	*Cumulative live birth:* 1) 9% 2) 54.8% 3) 70.6% 4) 51.9%	*ICER per scenario [€/additional live birth]:* 1) base case 2) €15.166 3) €10.837 4) €13.743 If society's WTP > €10.900, dose individualization according to ovarian reserve has the highest probability of being CE.	Costs of IVF cycle, ovarian reserve testing and Gonadotropin dose increase (limited cost inclusion)

Tjon-Kon-Fat (2015) The Netherlands (47)	CE of 1) IVF with ovarian stimulation, SET and subsequent cryocylces (3 cycles) 2) IVF in a modified natural cycle (MNC) (6 cycles) 3) IUI-COH (controlled ovarian hyperstimulation) (6 cycles) as a first-line treatment for patients with unexplainded subfertility.	602 couples with unexplained infertility, age 18-38.	*Mean costs per couple:* €7187 for IVF-SET €8206 for IVF-MNC €5070 for IUI-COH [2013 €]	*Live birth rate:* IVF-SET: 52% IVF-MNC: 43% IUI-COH: 47%	*ICER of IVF-SET compared with IUI-COH*= €43,375 (IVF-MNC is more expensive and less effective -> dominated) Both IVF strategies significantly more expensive than IUI-COH without being significantly more effective. CE of IVF-SET depends on society's WTP.	ART treatment costs, medication & pregnancy leading to delivery. Costs of pregnancy and delivery included.
Vitek (2013) USA (48)	CE of split IVF-ICSI for the treatment of couples with unexplained fertility	154 couples	*One cycle:*All IVF: $13,842Split IVF-ICSI: $15,605All ICSI: $15,605*2 cycles:*All IVF: $21,197Split IVF-ICSI: $22,176All ICSI: $22,552[2012 USD]	*Cumulative birth ratesOne cycle:*All IVF: 38.8%Split IVF-ICSI: 41.8%All ICSI: 41.8%*Two cycles:*All IVF: 61.6%Split IVF-ICSI: 64.9%All ICSI: 64.8%	*ICEROne cycle:*- All IVF= base case (=preferred approach)- Split IVF-ICSI: $58,766- All ICSI: $58,766*Two cycles:*- All IVF= base case- Split IVF-ICSI: $29,666 (= preferred approach)- All ICSI: $42,343	Direct costs of IVF & ICSI (procedure, medication and cryopreservation, ET, donor insemination)

Yilmaz (2017) Turkey (18)	Perinatal outcomes and CE of patients with advanced age. (Assessment of the current age cut-off in Turkey at 40 years old.)	456 patients: 158 older than 39, 298 < 39 years old	Mean cost per cycle for hormonal stimulation: ≥ 39 $1058 < 39 $732	*Clinical pregnancy rate:* ≥ 39= 11,3% <39= 38,6%	*Mean expense per pregnancy:* ≥ 39= $9294 < 39= $1874	No cost-details provided

References

1. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Hum Reprod. 2017 01;32(9):1786–801.
2. De Geyter C, Calhaz-Jorge C, Kupka MS, Wyns C, Mocanu E, Motrenko T, et al. ART in Europe, 2014: results generated from European registries by ESHREThe European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE). Hum Reprod. 2018 Sep 1;33(9):1586–601.
3. Centrum voor Reproductieve Geneeskunde. Information about the cost of an assisted reproductive treatment (ART) for belgian patients. Universitair Ziekenhuis Brussel; 2016.
4. Martins WP, Niederberger C, Nastri CO, Racowsky C. Making evidence-based decisions in reproductive medicine. Fertil Steril. 2018;110(7):1227–30.
5. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
6. Hernandez Torres E, Navarro-Espigares JL, Clavero A, López-Regalado Ml, Camacho-Ballesta JA, Onieva-García Má, et al. Economic evaluation of elective single-embryo transfer with subsequent single frozen embryo transfer in an in vitro fertilization/intracytoplasmic sperm injection program. Fertility and Sterility. 2015 Mar 1;103(3):699–706.
7. Fiddelers AAA, Dirksen CD, Dumoulin JCM, van Montfoort APA, Land JA, Janssen JM, et al. Cost-effectiveness of seven IVF strategies: results of a Markov decision-analytic model. Hum Reprod. 2009 Jul;24(7):1648–55.
8. Veleva Z, Karinen P, Tomas C, Tapanainen JS, Martikainen H. Elective single embryo transfer with cryopreservation improves the outcome and diminishes the costs of IVF/ICSI. Human Reproduction. 2009 Jul 1;24(7):1632–9.
9. Papaleo E, Pagliardini L, Vanni VS, Delprato D, Rubino P, Candiani M, et al. A direct healthcare cost analysis of the cryopreserved versus

fresh transfer policy at the blastocyst stage. Reproductive BioMedicine Online. 2017 Jan 1;34(1):19–26.
10. Roque M, Valle M, Guimarães F, Sampaio M, Geber S. Cost-Effectiveness of the Freeze-All Policy. JBRA Assist Reprod. 2015 Aug 1;19(3):125–30.
11. Le KD, Vuong LN, Ho TM, Dang VQ, Pham TD, Pham CT, et al. A cost-effectiveness analysis of freeze-only or fresh embryo transfer in IVF of non-PCOS women. Hum Reprod. 2018 Oct 1;33(10):1907–14.
12. Pham CT, Karnon JD, Norman RJ, Mol BW. Cost-effectiveness modelling of IVF in couples with unexplained infertility. Reproductive BioMedicine Online. 2018 Nov 1;37(5):555–63.
13. Eijkemans MJC, Kersten F a. M, Lintsen AME, Hunault CC, Bouwmans C a. M, Roijen LH, et al. Cost-effectiveness of 'immediate IVF' versus 'delayed IVF': a prospective study. Hum Reprod. 2017 May 1;32(5):999–1008.
14. van Loendersloot LL, Moolenaar LM, van Wely M, Repping S, Bossuyt PM, Hompes PGA, et al. Cost-effectiveness of single versus double embryo transfer in IVF in relation to female age. Eur J Obstet Gynecol Reprod Biol. 2017 Jul;214:25–30.
15. Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PMM, Koks C, Oosterhuis GJE, Hoek A, et al. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod. 2015 Oct;30(10):2331–9.
16. Scotland GS, McLernon D, Kurinczuk JJ, McNamee P, Harrild K, Lyall H, et al. Minimising twins in in vitro fertilisation: a modelling study assessing the costs, consequences and cost-utility of elective single versus double embryo transfer over a 20-year time horizon. BJOG. 2011 Aug;118(9):1073–83.
17. Messinger LB, Alford CE, Csokmay JM, Henne MB, Mumford SL, Segars JH, et al. Cost and efficacy comparison of in vitro fertilization and tubal anastomosis for women after tubal ligation. Fertility and Sterility. 2015 Jul;104(1):32-38.e4.
18. Yilmaz N, Kara M, Coskun B, Kaba M, Erkilinc S, Yenicesu O, et al. Perinatal outcomes and cost-effectivity of the assisted reproduction pregnancies with advanced age: A retrospective analysis. J Obstet Gynaecol. 2017 May;37(4):450–3.

19. Chambers GM, Sullivan EA, Ishihara O, Chapman MG, Adamson GD. The economic impact of assisted reproductive technology: a review of selected developed countries. Fertility and Sterility. 2009 Jun;91(6):2281–94.
20. Williams CB, Mackenzie KC, Gahagan S. The Effect of Maternal Obesity on the Offspring. Clin Obstet Gynecol. 2014 Sep;57(3):508–15.
21. Connolly M, Gallo F, Hoorens S, Ledger W. Assessing long-run economic benefits attributed to an IVF-conceived singleton based on projected lifetime net tax contributions in the UK. Human Reproduction. 2008 Dec 4;24(3):626–32.
22. Settumba SN, Shanahan M, Botha W, Ramli MZ, Chambers GM. Reliability and Validity of t he Contingent Valuation Method for Estimating Willingness to Pay: A Case of In Vitro Fertilisation. Appl Health Econ Health Policy. 2018 Oct 13;
23. Spiegel U, Gonen LD, Templeman J. Economic implications of in vitro fertilization using willingness to pay. J Public Health. 2013 Dec 1;21(6):535–57.
24. Devlin N, Parkin D. Funding fertility: issues in the allocation and distribution of resources to assisted reproduction technologies. Hum Fertil (Camb). 2003 May;6 Suppl 1:S2-6.
25. Collins JA. An international survey of the health economics of IVF and ICSI. Hum Reprod Update. 2002 May 1;8(3):265–77.
26. McTernan E. Should Fertility Treatment be State Funded? Journal of Applied Philosophy. 2015 Aug;32(3):227–40.
27. Hubens K, Arons AMM, Krol M. Measurement and evaluation of quality of life and well-being in individuals having or having had fertility problems: a systematic review. The European Journal of Contraception & Reproductive Health Care. 2018 Nov 30;0(0):1–10.
28. McMillan J. Allocating fertility services by medical need. Human Fertility. 2001 Jan;4(1):11–3.
29. Connolly MP, Griesinger G, Ledger W, Postma MJ. The impact of introducing patient co-payments in Germany on the use of IVF and ICSI: a price-elasticity of demand assessment. Hum Reprod. 2009 Nov 1;24(11):2796–800.

30. Dolan P, Rudisill C. Babies in waiting: Why increasing the IVF age cut-off might lead to fewer wanted pregnancies in the presence of procrastination. Health Policy. 2015 Feb 1;119(2):174–9.
31. IFFS Surveillance 2016 : Global Reproductive Health [Internet]. LWW. [cited 2019 Feb 20]. Available from: https://journals.lww.com/grh/Fulltext/2016/09000/IFFS_Surveillance_2016.1.aspx
32. Salama M, Isachenko V, Isachenko E, Rahimi G, Mallmann P, Westphal LM, et al. Cross border reproductive care (CBRC): a growing global phenomenon with multidimensional implications (a systematic and critical review). Journal of Assisted Reproduction and Genetics. 2018 Jul;35(7):1277–88.
33. Präg P, Mills MC. Assisted Reproductive Technology in Europe: Usage and Regulation in the Context of Cross-Border Reproductive Care. In: Kreyenfeld M, Konietzka D, editors. Childlessness in Europe: Contexts, Causes, and Consequences [Internet]. Cham: Springer International Publishing; 2017 [cited 2019 Feb 20]. p. 289–309. (Demographic Research Monographs). Available from: https://doi.org/10.1007/978-3-319-44667-7_14
34. Crawford S, Boulet SL, Mneimneh AS, Perkins KM, Jamieson DJ, Zhang Y, et al. Costs of achieving live birth from assisted reproductive technology: a comparison of sequential single and double embryo transfer approaches. Fertility and Sterility. 2016 Feb 1;105(2):444–50.
35. Højgaard A, Ottosen LDM, Kesmodel U, Ingerslev HJ. Patient attitudes towards twin pregnancies and single embryo transfer—a questionnaire study. Hum Reprod. 2007 Oct 1;22(10):2673–8.
36. Libby V, Thakore S, Weinerman R, Goldfarb J. Change in patient attitude towards multiple gestation pregnancies and selective fetal reduction over 20 years. Fertility and Sterility. 2018 Mar 1;109(3):e36.
37. Multiple Pregnancies Following Assisted Conception. BJOG: An International Journal of Obstetrics & Gynaecology. 2018;125(5):e12–8.
38. Maheshwari A, Griffiths S, Bhattacharya S. Global variations in the uptake of single embryo transfer. Hum Reprod Update. 2011 Feb;17(1):107–20.
39. Peeraer K, Debrock S, Laenen A, De Loecker P, Spiessens C, De Neubourg D, et al. The impact of legally restricted embryo transfer and

reimbursement policy on cumulative delivery rate after treatment with assisted reproduction technology. Hum Reprod. 2014 Feb;29(2):267–75.

40. Ombelet W. Is global access to infertility care realistic? The Walking Egg Project. Reproductive BioMedicine Online. 2014 Mar;28(3):267–72.
41. Ory J. IFFS Surveillance 2013. International Federation of Fertility Societies; 2013.
42. Atzori P, Directorate for Relations with National Parliaments. Spotlight on Parliaments in Europe - In vitro fertilisation (IVF) availability in national health services' provisions [Internet]. EP Directorate for Relations with National Parliaments; 2016. Available from: http://www.epgencms.europarl.europa.eu/cmsdata/upload/b50d9310-07f1-4691-9f2f-137cbc0cf75d/No_12_September_2016_in_vitro_fertilisation.pdf
43. ESHRE. A policy audit on fertility. Analysis of 9 EU countries [Internet]. ESHRE; 2017. Available from: http://www.fertilityeurope.eu/wp-content/uploads/2018/03/EPAF_FINAL.pdf
44. Association P. Scottish government to fund three cycles of IVF. The Guardian [Internet]. 2017 Mar 24 [cited 2019 Feb 25]; Available from: https://www.theguardian.com/society/2017/mar/24/scottish-government-to-fund-three-cycles-of-ivf
45. Groen H, Tonch N, Simons AHM, van der Veen F, Hoek A, Land JA. Modified natural cycle versus controlled ovarian hyperstimulation IVF: a cost-effectiveness evaluation of three simulated treatment scenarios. Human Reproduction. 2013 Dec 1;28(12):3236–46.
46. Moolenaar LM, Broekmans FJM, van Disseldorp J, Fauser BCJM, Eijkemans MJC, Hompes PGA, et al. Cost effectiveness of ovarian reserve testing in in vitro fertilization: a Markov decision-analytic model. Fertil Steril. 2011 Oct;96(4):889–94.
47. Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PMM, Koks C, Oosterhuis GJE, Hoek A, et al. Is IVF—served two different ways—more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod. 2015 Oct 1;30(10):2331–9.
48. Vitek WS, Galárraga O, Klatsky PC, Robins JC, Carson SA, Blazar AS. Management of the first in vitro fertilization cycle for unexplained

infertility: a cost-effectiveness analysis of split in vitro fertilization-intracytoplasmic sperm injection. Fertil Steril. 2013 Nov;100(5):1381–8.

Chapter 2: Insights from the population ethics literature

Part 1: To be or not to be: future lives in economic evaluation

Abstract

Sometimes healthcare will affect the health of people in the future, or their chance of coming into existence. Should such outcomes be valued in health-economic evaluation? Guidelines implicitly recommend their inclusion but this rule has counterintuitive implications and is not consistently applied in practice. We suggest making a distinction between 'necessary' and 'potential' future lives in Health Technology Assessment. Necessary lives will exist independent of our healthcare choices and should be included. Potential lives are choice-dependent and should be excluded. This rule offers intuitive solutions within the HTA framework, and it changes the cost-effectiveness of several interventions where necessary future lives are affected.

This subchapter has been published as:

Luyten, J., Verbeke, E., Schokkaert, E. (2021). To be or not to be: Future lives in economic evaluation. *Health economics, 31* (1), 258-265. doi: 10.1002/hec.4454

2.1.1 Introduction

Besides their effects on people living here and now, some healthcare programs will also affect people in the future. As illustrated in Table 2 (and more extensively in the Appendix), future people's lives can become affected in their quality-of-life, their life expectancy or through a changed probability of being born. Expressed in Quality-Adjusted Life Years (QALYs), these effects can be substantial but their status in the standard economic evaluation framework is unclear. Our goal is to suggest a guideline for their in- or exclusion, based on the non-impartial concept of of necessitarianism in population ethics.

2.1.2 Guidelines and practice

Economic evaluation guidelines do not specifically discuss when future lives should be included. However, they do prescribe inclusion of *all* economic or health consequences of an intervention, and this for *all* parties affected by it, including offspring. For instance, the second 'US Panel on Cost-Effectiveness in Health and Medicine' guidelines published in 2017, recommend "*that analysts seek to identify all consequences of a decision, whether positive or negative and regardless of whether they occur within or outside of the formal healthcare sector*." Furthermore, "*the scope of a study should be defined broadly enough to encompass the full range of groups of people affected by the intervention and all important consequences.*" The panel prescribes that "*if an intervention affects both mother and child, the analysts should include the health impacts and costs to each*." This prescription not only relates to born children as the guideline mentions smoking cessation programs for pregnant women and folic acid fortification explicitly as examples.(1) Similar prescriptions can be found in other texts.(1–4)

Two options are usually offered to exclude costs and effects but neither of these are relevant to future lives. Those that are trivially small can be excluded based on the 'rule of reason'. Yet, effects on future generations

are, at least potentially, substantial. Indirect, further-fetched consequences of healthcare can - although this remains matter of dispute - be excluded on grounds of ‘unrelatedness’.(5) Yet effects on future generations are often immediately connected to the program and therefore related. Hence, we interpret the guidelines as a prescription to include all health effects on future lives, without further qualification.

In practice, as documented in a literature review about economic evaluation of interventions affecting fertility (6), studies usually exclude future lives from cost-effectiveness ratios, and when they are included, it is done in an inconsistent manner. The authors concluded that more specific guidelines are needed. Studies of disease eradication programs or other programs with a large intergenerational stake (e.g. climate change policies), often include benefits for future lives through projecting future population growth for several decades.(7–12)

2.1.3 Non-impartial views in population ethics

Which future lives to include depends on their moral status, and on the extent to which their interests should be considered in today's policy-making, e.g. in resource allocation decisions. The field of population ethics has a long tradition of discussing policy trade-offs between present and future generations, usually in the context of environmental protection. But these insights are applicable to HTA as well.

The so-called "non-impartial" theories make a distinction between future lives that are relevant for the evaluation and future lives that are not. Table 1 summarizes three broad approaches to identifying relevant future lives.(13,14) From a 'presentist' perspective, there are only 'existing' and 'non-existing' people and only the former's interests should be taken into account by policy makers. Future people do not (yet) exist, have therefore no moral status and can be ignored. *Actualists* argue that existence in itself is not a necessary condition to be relevant and that all future lives that will 'actually' live at some point in time have interests that need to be respected. According to *necessitarians*, we need to make a different distinction: between those future people who will always exist, irrespective of the choice we make, and those people whose existence is purely contingent upon those choices. Whereas the former type of people exist in a 'necessary' way, the latter's existence is merely 'potential' as it is the choice itself that will determine whether the future person will effectively come into existence. A presently existing person is by definition also a 'necessary', and an 'actual' individual, but these latter types do not always presently exist. A necessary future live is always actual but an actual future live is not always a necessary life because its actual existence may be the result of choice.
Apart from deciding which future lives should be taken into account, the literature in population ethics deals with the question *how* these lives should be taken into account exactly. Three main options are to take the sum total of the outcomes of future people (a position in line with total utilitarianism), to take the average across people (average utilitarianism)

or only to consider those outcomes that exceed a certain threshold (critical-level utilitarianism). Moreover, some authors have also presented prioritarian approaches (see, e.g., (13,14) for more elaborate discussion). The present practice in HTA implements some form of "totalism" and in order to focus on the question of which (future) lives should be considered we will follow that assumption in this short paper.

Can these distinctions guide us in how to incorporate future lives in the current practice of health-economic evaluation? Systematic inclusion of all possible future lives that become affected by funding decisions in the present (the guidelines view, prescribing inclusion of all affected parties, a position in line with actualism) has counterintuitive implications when strictly followed in HTA. For instance, interventions that purposely reduce fertility (e.g. contraception or carrier screening) would be discouraged because they eliminate future QALYs while interventions that increase fertility would be stimulated. In the case of Assisted Reproductive Technologies (ART), inclusion of a lifetime of additional QALYs of the future child would make these programs highly cost-effective. Researchers who study economic evaluation of ART therefore suggest avoiding the use of QALYs in this context. (15–17)

The view of presentism is counterintuitive as well as it would deny the obvious relevance of future people's health outcomes for the value of certain healthcare interventions. For instance, programs that protect future babies (e.g. through substance abuse prevention) would be valued only based upon their maternal effects, not on their ability to avoid disabled future children. Also, the interests of future generations in long-term public health programs (e.g. disease eradication) would be excluded and hence a key reason for undertaking these programs would be eliminated. Intuitively, at least *some* future lives are relevant to assess an intervention's value.

A necessitarian perspective seems to offer promising guidance to HTA. Introducing a distinction between necessary and potential lives, where potential lives must be excluded while necessary lives must be included, would avoid that policy makers favour creating new people over increasing the quality of life of the 'necessarily existing population'. It would also avoid that the health interests of future lives that would be born anyway and that would be affected by today's decisions are ignored.

Which healthcare interventions would affect 'necessary' and which ones 'potential' lives? Interventions done for reproductive reasons, such as ART, carrier screening, anticonception or abortion will all affect 'potential' future lives, as the coming into existence of these lives is dependent on the intervention. Also interventions that affect fertility more indirectly such as hysterectomy or testicular cancer treatment will affect the coming into existence of 'potential' lives. But interventions that affect future people that would come to life anyhow but whose Quality-Adjusted Life Expectancy (QALE) would be different because of the intervention, affect 'necessary' lives and these should be considered. In the first place, this means that as soon as a fetus becomes a 'person' (and hence a present life), (s)he should be taken into account, as a present life is a necessary one. Necessitarianism makes a difference, however, for interventions that affect an early fetus before it is treated as a 'person' (e.g. before 12 weeks of pregnancy). If the intervention affects the fetus in terms of its chance of coming into existence (e.g. stillbirth prevention) it would affect potential lives, whereas an intervention that does not affect viability of the fetus but that does affect its later QALE would affect a necessary life. Future lives should therefore be considered in economic evaluations of maternal vaccinations, nutritional fortification and supplementation, psychotherapy during pregnancy, chlamydia treatment or other prenatal interventions. Arguably, this makes a large difference to the cost-effectiveness of these programs, as both maternal and child health are influenced, with the latter effects

potentially stretching out during the lifetime of the baby. As these effects are not always incorporated in existing economic evaluations (see Table 2), the cost-effectiveness of these programs is underestimated.

Public health interventions (like e.g. disease eradication, environmental safety programs, action on antimicrobial resistance or substance bans), where the health of a population is affected rather than that of specific individuals, are somewhat more difficult. These programs will affect necessary and potential lives simultaneously. They will affect future generations that will exist anyhow and who will benefit from the intervention, either through reduced future risks or through healthier conception and fetal development. But increased public health can also increase fertility, leading to an increase in the size of future generations. Whereas the former future lives are necessary, the latter lives are potential lives that must be excluded. Evidently, it will be difficult, but not impossible to distinguish between the two.

Table 1: Three population ethics theories

Theory	Core	Arguments in favour	Arguments against
Presentism	Only lives that presently exist have interests that should be respected.	- Future lives do not yet exist, have no identity yet and can therefore not have any rights. - Our relationship with future people is asymmetric and cannot be the same as our relationship with contemporaries (where there is *direct* reciprocity as a grounding of ethics). - Future lives are too uncertain and their interests too unknown to take into account in today's decisions.	- Outcomes in which future people would live a blissful or a hellish life would be considered equally good. - The welfare of future lives is attributed zero value and can be sacrificed at any expense. - The rights of not-yet-existing future lives can still be grounded elsewhere than in a particular identity, for instance in our common humanity.
Actualism	All present or future lives that will actually exist in one of the possible worlds under evaluation have interests that should be respected.	- An action can only be good or bad for a person if that person ever actually lives. People who exist merely potentially can have no experiences. - Our relationship with future lives is based on *indirect, descending* reciprocity. Current generations have to take care of future generations, just as they were taken care of by previous generations.	- Aggregating the welfare of actual lives in each alternative implies that outcomes in which more future lives will exist might be preferred, even if individual future welfare would be lower than in the alternatives. - The coming into existence of uncertain and abstract future lives could be prioritized at the expense of present people's vital interests.
Necessi-tarianism	All lives that will necessarily exist have interests that should be respected. Existing necessarily means that	- We have obligations to future generations not to harm their vital interests but not to bring lives-worth-living into existence. - The welfare of a life itself is not a reason to bring that life into existence.	- The reasons to exclude potential future lives are weaker in cases where their life would not be worth living. It seems unjustifiable that the suffering of a potential child does not provide us a reason to prevent its existence.

	existence is not dependent upon the decision under consideration.		- Decisions become choice-set dependent: whether a future life is considered 'necessary' will depend on the alternatives considered in the decision context. - Only the quality of future lives is considered but not the 'quantity'. As a consequence, even extreme implications on future population growth cannot be taken into account.(8,9)

Table 2: Healthcare interventions affecting future lives, in/exclusion in published studies and implications of using a necessitarian rule for in/exclusion of future lives

Intervention	Effects of intervention on future lives		Outcomes included in published studies (Reproduced from Goldhaber-Fiebert and Brandeau, 2015)	Necessitarian prescription					Expected impact on cost-effect-iveness of the inter-vention
	Effect on coming into existence?	Effect on QALE?		QALYs of Parents	QALYs of future offspring				
					'affected' before conception	affected during first 12 weeks of pregnancy	affected after 12 weeks of pregnancy	affected after birth	
Contraception	Yes	No	Cost per pregnancy averted or cost per woman's QALY gained. No studies consider child QALY losses from intentionally prevented pregnancy nor maternal QALY losses for complications such as infertility due to	All parental QALYs gained through contraception should be counted.	Offspring QALYs lost through decreased fertility of parents should not be counted.	None affected	None affected	None affected	None

			ectopic pregnancy						
Therapeutic abortion (in first 12 weeks)	Yes	No	Studies evaluate different abortion methods so all strategies will equally result in pregnancy termination. Strategies may result in higher risks of future infertility but these future lost QALYs are not considered.	All parental QALYs gained/lost through abortion should be counted.	Offspring QALYs lost through potential future maternal infertility should not be counted.	Offspring QALYs lost through aborting fetuses before 12 weeks should not be counted.	None affected	None affected	None
Genetic testing of pregnant women and their fetuses, plus option to terminate (first 12 weeks)	Yes	No	Studies generally include the cost of having a child with disabilities. Some studies include maternal QALY gains/losses due to false positive screening result,	All parental QALYs gained through testing and termination should be counted.	None affected	Offspring QALYs lost through aborting fetuses before 12 weeks should not be counted.	None affected	None affected	None

			miscarriage, pregnancy termination, and giving birth to an affected child. Some studies assume the possibility of a "replacement child" after pregnancy termination						
Preconceptiona l genetic testing	Yes	No	Cost per future affected birth avoided. One study also reports cost per carrier detected. No studies consider QALYs.	All parental QALYs gained through testing should be counted.	Offspring QALYs lost through decreased fertility of parents should not be counted.	None affected	None affected	None affected	None
Contraception for HIV-infected women	Yes	No	One study considers births averted; one considers HIV-infected births averted. Neither considers	All parental QALYs gained through contraception should be counted.	Offspring QALYs lost through decreased fertility of parents	None affected	None affected	None affected	None

			antiretroviral therapy to attenuate newborn HIV infection, nor benefits to mother of ongoing treatment nor of having a child without HIV.		should not be counted.				
IVF vs. no IVF	Yes	No	Cost per intermediate outcome (live birth or live healthy birth) but no loss of QALYs for fetal loss or children with birth complications (low birth weight). Sometimes maternal QALYs considered. One study that considers	All parental QALYs gained through IVF should be counted.	Offspring QALYs gained through increased fertility of parents should not be counted.	None affected	None affected	None affected	None

			maternal QALYs presents philosophical reason for not including offspring QALYs.						
Vasectomy	Yes	No	Costs per pregnancy averted or couple-years of protection. Studies do not consider potential demand for future vasectomy reversal/inability to reverse in all cases/loss of unwanted future infertility. Studies do not consider loss of child QALYs	All parental QALYs gained through contraception should be counted.	Offspring QALYs lost through decreased fertility of parents should not be counted.	None affected	None affected	None affected	None
Hysterectomy	Yes	No	Women's QALYs gained due to treatment	All parental QALYs gained through	Offspring QALYs lost	None affected	None affected	None affected	None

			of the benign conditions. No fertility issues considered even though some of the comparators are less invasive and more fertility sparing	hysterectomy should be counted	through infertility of parents should not be counted.				
Chemotherapy for cancer in women of childbearing age	Yes	No	Studies consider woman's QALYs gained due to treatment but do not consider fertility losses due to treatment.	All parental QALYs gained through chemotherapy should be counted	Offspring QALYs lost through infertility of parents should not be counted.	None affected	None affected	None affected	None
Chemotherapy for testicular cancer	Yes	No	Studies consider gains in men's health due to cancer treatment, but do not consider potential health losses due to future infertility	All parental QALYs gained through chemotherapy should be counted	Offspring QALYs lost through infertility of parents should not be counted.	None affected	None affected	None affected	None
Chlamydia treatment	Yes	Yes	Infertility and ectopic	All parental QALYs gained	Offspring QALYs	Offspring	Offspring QALYs	Offspring QALYs	Limited effect

			pregnancy included as a loss of maternal QALYs. QALY gains for infants born if they avoid health harms. Male QALYs gained due to chlamydia treatment; no male QALY losses for fetal loss or caring for sick child.	from chlamydia treatment should be counted.	gained through increased fertility of parents should not be counted.	QALYs gained through less miscarriages in first 12 weeks should not be counted.	gained through less miscarriages after 12 weeks should be counted.	gained through increased QALE should be counted.	
HPV vaccination	Yes	Yes	Girls: Female QALY gains including those from avoiding hysterectomy but no fertility-specific effect of this nor child QALYs. Boys: Cost per male or female QALY gained. Female QALYs	All parental QALYs gained through HPV prevention should be counted	Offspring QALYs gained through increased fertility of parents should not be counted.	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Limited effect

			gained when males do not transmit HPV to them.						
Smoking cessation for women of childbearing age	Yes	Yes	Considers net gain/loss of value. Benefit is reduced short-term cost of low birth weight children.	All parental QALYs gained through smoking cessation should be counted.	Offspring QALYs gained through increased fertility of parents should not be counted.	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Limited effect
Antiretroviral therapy for HIV-infected women	Yes	Yes	Infant QALYs gained due to avoiding HIV infection. No maternal QALYs gained due to child avoiding HIV infection.	All parental QALYs gained through ART should be counted.	Offspring QALYs gained through increased fertility of parents should not be counted.	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Substantial effect
Preeclampsia treatment	Yes	Yes	One study included offspring	All parental QALYs gained through	None affected	Offspring QALYs	Offspring QALYs gained	Offspring QALYs gained	Substantial effect

			QALYs; one did not include these QALYs due to lack of data; one reported intermediate outcome only.	preeclampsia treatment should be counted.		gained through less miscarriages in first 12 weeks should not be counted	through less miscarriages after 12 weeks should be counted	through increased QALE should be counted.	
Treatment of diabetes in pregnant women	Yes	Yes	Studies generally include offspring QALYs and/or maternal QALYs lost from loss of fetus/baby and maternal QALYs gained from improved maternal health	All parental QALYs gained through diabetes treatment should be counted.	None affected	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Substantial effect
Smoking cessation for pregnant women	Yes	Yes	Often cost-benefit analysis (costs of outcomes such as preterm newborns but not necessarily lost QALYs if	All parental QALYs gained through smoking cessation should be counted.	None affected	Offspring QALYs gained through less miscarriages in first 12 weeks	Offspring QALYs gained through less miscarriages after 12 weeks	Offspring QALYs gained through increased QALE should be counted.	Substantial effect

			they die) or intermediate outcomes considered. Much more focused on maternal QALYs than potential offspring QALYs.			should not be counted	should be counted		
Genetic testing of pregnant women and their fetuses, plus option to treat (e.g. CRISPR)*	Yes	Yes	NA	All parental QALYs gained through testing and treatment should be counted.	None affected	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Substantial effect
IVF: one technique vs. another*	Yes	Yes	Studies generally include the cost of having a child with disabilities. Some studies include maternal QALY	All parental QALYs gained through IVF should be counted.	Offspring QALYs gained through increased fertility of parents	Offspring QALYs gained through less miscarriages in first 12 weeks	Offspring QALYs gained through less miscarriages after 12 weeks	Offspring QALYs gained through increased QALE should be counted.	Substantial effect

			gains/losses due to false positive screening result, miscarriage, pregnancy termination, and giving birth to an affected child. Some studies assume the possibility of a “replacement child” after pregnancy terminatio		should not be counted.	should not be counted	should be counted		
Pre-implantation genetic screening of embryo plus option to replace*	Yes	Yes	Cost per healthy infant. No loss for termination or problems after birth	All parental QALYs gained through testing should be counted.	None affected	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Substantial effect
Follate supplementation for pregnant women	Yes	Yes	Generally QALYs gained from children born healthy and	All parental QALYs gained through supplementatio	None affected	Offspring QALYs gained	Offspring QALYs gained through less	Offspring QALYs gained through	Substantial effect

			avoiding neural tube defects because of the intervention. Studies do not all count fetuses lost before birth. No studies count maternal/paternal QALYs from having such a loss nor for having children with or without a neural tube defect.	n should be counted		through less miscarriages in first 12 weeks should not be counted	miscarriages after 12 weeks should be counted	increased QALE should be counted.	
Imaging in pregnant women	Yes	Yes	Maternal QALYs from direct health benefits. Newborn QALYs conditional on birth. No maternal QALY loss for leukemia, preterm birth or miscarriage	All parental QALYs gained through imaging should be counted	None affected	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	Offspring QALYs gained through increased QALE should be counted.	Substantial effect

			caused by imaging.						
Occupational health and safety*	Yes	Yes	Studies consider benefits in present population only.	All QALYs gained in current workers through health and safety measures should be counted	QALYs gained in offspring through increased fertility of current workers should not be counted.	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	QALYs gained in offspring and future workers through increased QALE should be counted.	Limited effect
Environmental protection*	Yes	Yes	Studies also consider benefits in future population.	All QALYs gained in current generations should be counted	Offspring QALYs gained through increased fertility of parents should not be counted.	Offspring QALYs gained through less miscarriages in first 12 weeks should not be counted	Offspring QALYs gained through less miscarriages after 12 weeks should be counted	QALYs gained in future generations through increased QALE should be counted.	Substantial effect
Disease eradication*	Yes	Yes	Studies mostly consider benefits in present population but often also in	All parental QALYs gained through disease prevention should be counted	Offspring QALYs gained through increased	Offspring QALYs gained through less miscarriage	Offspring QALYs gained through less miscarriages after 12	QALYs gained in future generations through increased	Substantial effect

			projected future population for several decades.		fertility of parents should not be counted.	s in first 12 weeks should not be counted	weeks should be counted	QALE should be counted	

* Not in original review by Goldhaber-Fiebert and Brandeau, 2015

2.1.4 Implications and criticisms

First, a threshold needs to be defined from when a foetus becomes a person. In Table 2 we set the threshold at 12 weeks after conception. Healthcare that affects the viability of a mere foetus would affect a potential life, whereas from the point that the foetus becomes a person it would change into a necessary (present) life. This demarcation line will change the moral status of the foetus from one day to another. This seems arbitrary but it is nonetheless a standard approach in bioethics and regulation of interventions such as abortion.

Second, programs for women of childbearing age could be drastically different in terms of cost-effectiveness compared to the exact same program but executed in similarly-aged men or in older women, as only the former women have the potential to affect necessary future lives. This is a logical consequence of the premise that certain future lives should matter in a fair policy evaluation.

Third, the reasons to exclude potential lives seem weaker in cases where their quality of life would be low. While it is intuitive that the QALYs of a child do not provide a reason to bring it into existence, it seems less justifiable that the suffering of a potential child does not provide a reason to prevent its existence.(18,19) For example, if a fetus is aborted after genetic screening, we can consider its future suffering a reason to prevent its existence. From our suggested approach, the value of the screentest would solely be based on the outcomes for the parents, not on the value of the avoided suffering by the child itself.

Fourth, when three alternative states of the world need to be compared in necessitarianism, this can lead to a problem of ‘cyclicity’, where state A is judged better than B, B better than C and C better than A because lives considered necessary in one alternative can be merely potential in another.(20) This cyclicity reflects the deeper problem that the outcomes

in a necessitarian approach are dependent on the alternatives that are available in the choice set. It is an interesting issue whether such context-dependency is problematic (as argued, e.g., in (13)) or rather offers a potential opening for a richer theory. A deeper discussion of this issue goes beyond the purpose of this paper. However, from a more pragmatic perspective, it can be noted that, in the context of HTA, where only two alternatives are compared in ICERs, the problem of cyclicity does not occur. At the same time, drawing the demarcation line between "necessary" and "potential" lives based on how inevitable they are, may depend on the specifics of the evaluation problem and is therefore inherently context-specific. Consider, e.g., the comparative evaluation of different ART treatment options in a situation where ART treatment will start anyhow: since an ART-conceived child will be born anyhow, all the considered treatments will then necessarily lead to the conception of future offspring. Therefore, although the decision whether to fund ART in principle involves only potential lives (as described in the previous section), the choice between different ART treatment options implies consequences for necessary lives. There would be no reason to disregard the interest of an ART-child that will be born in any of the situations under evaluation.

2.1.5 Conclusion

There is a lack of coherent guidance for analysts aiming to estimate the cost-effectiveness of healthcare interventions that affect future lives. We suggest a necessitarian perspective in which only necessary future lives, which will exist independent of reimbursement decisions, are to be included. Some deeper issues, such as the one of choice-set dependency, should be further analysed but in general our proposal provides an intuitive distinction between which future lives to count and which ones not. Strict application of this rule has the potential of changing the cost-effectiveness of several interventions where necessary lives are involved.

References

1. Neumann P, Sanders G, Russell L, Siegel J, Ganiats T. Cost-effectiveness in health and medicine. 2nd ed. New York: Oxford University Press; 2017. 496 p.
2. NICE. The guidelines manual: Process and methods - 7 Assessing cost effectiveness [Internet]. The guidelines manual. 2012 [cited 2018 Jun 20]. Available from: https://www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness
3. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
4. Swedish Council on Health Technology Assessment (SBU), EUnetHTA. Methods for health economic evaluations - A guideline based on current practices in Europe. 2015;100.
5. Morton A, Adler AI, Bell D, Briggs A, Brouwer W, Claxton K, et al. Unrelated Future Costs and Unrelated Future Benefits: Reflections on NICE Guide to the Methods of Technology Appraisal. Health Econ. 2016;25(8):933–8.
6. Goldhaber-Fiebert JD, Brandeau ML. Evaluating Cost-Effectiveness of Interventions that Affect Fertility and Childbearing: How Health Effects are Measured Matters. Med Decis Making. 2015 Oct;35(7):818–46.
7. Levin A, Burgess C, Garrison LP, Bauch C, Babigumira J, Simons E, et al. Global Eradication of Measles: An Epidemiologic and Economic Evaluation. Journal of Infectious Diseases. 2011 Jul 1;204(Supplement 1):S98–106.
8. Fitzpatrick C, Sankara DP, Agua JF, Jonnalagedda L, Rumi F, Weiss A, et al. The cost-effectiveness of an eradication programme in the end game: Evidence from guinea worm disease. PLOS Neglected Tropical Diseases. 2017 Oct 5;11(10):e0005922.
9. Sicuri E, Evans DB, Tediosi F. Can Economic Analysis Contribute to Disease Elimination and Eradication? A Systematic Review. PLOS ONE. 2015 Jun 12;10(6):e0130603.

10. World Bank. Economic Evaluation of Climate Change Adaptation Projects : Approaches for the Agricultural Sector and Beyond [Internet]. 2010 [cited 2022 Apr 5]. Available from: https://openknowledge.worldbank.org/handle/10986/27752
11. USAIDS. Methods for economic analysis of climate change adaptation interventions. 2013. (African and Latin American Resilience to Climate Change Project).
12. Chambwera M, Heal G, Dubeux C, Hallegatte S, Lecrec L, Markandya A, et al. Economics of adaptation. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In New York, NY, USA: Cambridge University Press; 2014. p. 945–77.
13. Greaves H. Population Axiology. Philosophy Compass. 2017;12(11):e12442.
14. Arrhenius G. Population Ethics. The Challenge of Future Generations. (Unpublished version May 2016).
15. Devlin N, Parkin D. Funding fertility: issues in the allocation and distribution of resources to assisted reproduction technologies. Hum Fertil (Camb). 2003 May;6 Suppl 1:S2-6.
16. Chambers GM, Adamson GD, Eijkemans MJC. Acceptable cost for the patient and society. Fertility and Sterility. 2013 Aug 1;100(2):319–27.
17. Collins JA. An international survey of the health economics of IVF and ICSI. Hum Reprod Update. 2002 May 1;8(3):265–77.
18. Broome J, editor. The value of a person. In: Ethics out of Economics [Internet]. Cambridge: Cambridge University Press; 1999 [cited 2020 Jun 18]. p. 228–42. Available from: https://www.cambridge.org/core/books/ethics-out-of-economics/value-of-a-person/D2464EC9B0C856C762386ED41C1EAF6B
19. Parsons J. Axiological Actualism. Australasian Journal of Philosophy [Internet]. 2010 Apr 8 [cited 2020 Jun 18]; Available from: https://aap.tandfonline.com/doi/pdf/10.1093/ajp/80.2.137?needAccess=true#.Xus5_EUzZPY
20. Broome J. The Value of Life and the Value of Population. Journal of Population Economics. 1996;9(1):3–18.

Part 2: Future offspring costs in economic evaluation

Abstract

Economic evaluation guidelines increasingly prescribe inclusion of all future costs. We point at an important dimension of future costs that is systematically neglected. Healthcare can affect future offspring, either through affecting the patient's fertility or through determining future offspring's health. As we show, the future costs associated to these changes can be substantial and they will vary across interventions and demographic groups. However, a systematic inclusion of these future offspring costs would raise many problems on its own. Based on the population ethics concept of necessitarianism, we suggest that only those future costs that spring from 'necessary' future lives should be included in future cost calculations, while all costs associated to 'potential' future lives can be ignored. This approach allows excluding most future offspring costs and avoids skewed cost-effectiveness outcomes of interventions with fertility effects, while taking into account the economic implications of preventing disease in future generations that will exist by necessity. Overall, future generations expose a substantial gap in today's Health Technology Assessment (HTA) methodology and further discussion of the issues they raise is needed.

This subchapter has been published as:

Verbeke, E., Luyten, J. (2021). Future Offspring Costs in Economic Evaluation. *Pharmacoeconomics*. doi: 10.1007/s40273-021-01102-w

2.2.1 Future costs

Although the in- or exclusion of future costs in health economic evaluation has been the subject of debate for decades, attention for this topic has recently increased.(1–7) Currently, future costs *related* to an intervention are typically recommended to be included in cost-effectiveness estimates. For instance, in the short term stents may be a less expensive treatment for artery disease than a bypass, but when the future costs of revascularization are accounted for, this may be different.(8) The more contentious issue is the inclusion of so-called '*unrelated*' future costs: the future costs that are not the direct ('related') consequence of the intervention.(9,10) For instance, increased life expectancy of patients also implies higher costs for e.g. dementia or cancer care, or from a societal perspective, higher pension costs.

The ongoing debate and empirical research was recently summarized by de Vries et al. and Kearns.(1,4) Whereas in earlier days, consensus seemed to move in the direction of only including related future costs in cost-effectiveness ratios, more recent guideline revisions move towards the inclusion of both (related and unrelated future costs). For instance, Israel, the Netherlands, Sweden, Thailand and the United States now recommend including unrelated future costs as well.(1,4) The second Panel on Cost-effectiveness in health and medicine also updated recommendations to include *all* future costs: "*We recommend that all healthcare costs, related and unrelated, should be considered either when survival under alternative interventions is not the same or when cost components cannot be readily identified as related to the target condition."*(11) NICE currently still recommends the exclusion of unrelated future costs but acknowledges that this is a complex debate.(1,12)

Standardized calculation methods for future costs are by now available (13–15) and have been used in a variety of applications such as

immunization, diabetes treatment and interventions for chronic heart diseases.(16–19) A frequently cited model is that of Meltzer, 1997. Overall per capita medical spending (specific for age and sex) is linked to survival curves to estimate average future medical costs.(15) More recently, the tool PAID (Practical Application to Include Future Disease Costs) was developed in the Netherlands, departing from per capita costs of 99 different diseases. In contrast to the 'Meltzer method', PAID excludes related future costs and takes into account increased costs (associated with increased hospital expenditure) in the last year of life.(13) In 2020, the tool was updated with more recent cost-data and the introduction of future non-medical consumption costs.(20) Alternatively, Briggs et al. developed a top-down approach based on the overall NHS budget. In this approach, related disease costs are subtracted off the overall NHS budget and the remaining unrelated expenditure is then allocated to five categories of care.(14)

Our intent is to draw attention to one dimension of future costs, which has, to our knowledge, received no attention in the current debate, and is not considered in these developed calculation tools, namely the future costs generated by a patient's (potential) future offspring. While these costs are typically ignored, we argue that this dimension can nevertheless be impactful and requires consideration in the development of a coherent, valid framework for future cost calculations. Subsequently, we suggest a rule regarding when to in- or exclude these costs.

2.2.2 Future offspring costs

Besides the effects of healthcare on the patient undergoing treatment, healthcare can also affect a patient's future offspring. This can happen when an intervention influences someone's ability to have children: through impacting his or her reproductive health during fertile ages (e.g. hysterectomy) or by affecting survival (e.g. heart surgery). It can also occur when an intervention has effects on the later quality-adjusted life

expectancy (QALE) of a patient's prospective child(ren). Examples of such programs are maternal vaccination, supplementation during pregnancy, the prevention of transmittable diseases and public health programs with intergenerational effects such as substance abuse prevention, immunization or disease eradication. For interventions where these future offspring effects are indeed plausible, incremental future costs will not only occur within treated patient's lives but also within the lives of their future offspring. Using a smoking cessation program for young women as an example, Table 1 illustrates how the total future medical costs of this program (both related and unrelated) can be affected by future offspring.

Table 1: The effects of a smoking cessation program for women of child-bearing age and their future cost implications (from a healthcare payer perspective).

Effects of a smoking cessation program for women of child-bearing age	**Impact on future medical costs**	**Why?**	**Type of future medical costs**	**Type of life affected**
Women experience increased quality of life	↓	Less future disease related to smoking behaviour	Future related costs	Present life
Women live longer	↑	By living longer, the women will require more healthcare in general: e.g. dementia care, treatment of hearing problems, heart diseases…	Future unrelated costs	Present life
Higher probability of women becoming pregnant[2]	↑	A child will require healthcare throughout its life, more future children implies more future medical costs in general.	**Future offspring costs:** 1) Related costs: fertility is affected by the intervention 2) Unrelated costs: fertility is affected	Potential future life

[2] Increased fertility is also captured by an increase in quality of life for the woman herself. Here, we consider the separate, additional effect that more children will be born.

			by changes in life expectancy of the women	
Lower probability of adverse birth outcomes (e.g. low birth weight)	1) ↓ 2) ↑	1) Less future children will require healthcare because of adverse birth outcomes 2) Children will live longer in the future, requiring more healthcare in general.	**Future offspring costs:** 1) Related 2) Unrelated	Necessary future life

In the ongoing debate and developed methods, only the future costs associated with presently existing generations have been considered. However, economic evaluation guidelines prescribe the inclusion of *all* relevant consequences (i.e. both costs and health effects), for *all parties affected* by an intervention. Even when these parties were not part of the target population or have a different degree of proximity to the intervention.(11,21–23) Neumann et al. consider as an example the inclusion of the important consequences for infants related to folic acid fortification in pregnant women. When interventions have a "long tail" of consequences that are realized among multiple generations, it is argued that a time horizon should be considered that allows capturing the salient consequences of every alternative. As a consequence, aside from the debate on related versus unrelated future costs, it is unclear why future offspring as a whole are systematically neglected in this context. We argue that 'future offspring costs' should be included for at least the following two reasons.

First, the implications of future offspring on future cost calculations will not be constant across interventions. A broad range of interventions affects future offspring but for many other interventions this reproductive dimension is completely absent (e.g. dementia or dental treatments). Furthermore, for those interventions where offspring is affected, age-dependent fertility implies that the expected future offspring costs (or savings) will vary by patient age and will have a bigger impact the younger the patient is. Gender will be an additional determining factor. While paternal health also affects fertility and can lead to specific epigenetic changes in descendants (24–27), the probability and hence impact of these consequences will be lower as the health of a future child is not only determined during conception but also during pregnancy. In contrast, for treatments specifically focused on pregnant women or women with a child wish, fertility rates will be above average and hence, the impact of including future offspring costs will be larger. Future offspring costs might also vary significantly within a

specific disease area and for otherwise identical patients. For instance, in the case of cancer treatment, chemotherapy can have a different impact than immunotherapy on a patient's fertility and hence it will affect future offspring differently.(28) Because the impact of future offspring costs is different across interventions and patient groups, their systematic exclusion will misrepresent the resource implications of some interventions more than others.

Second, when future offspring is indeed affected, the associated costs will not be trivial but can have a large impact on the incremental cost-effectiveness ratio (ICER). In Table 2 we illustrate this with a 'back-of-the-envelope' cost-effectiveness calculation of antiretroviral therapy for HIV in Europe. An HIV-positive woman at the age of 18, 25 and 40 is considered, receiving antiretroviral therapy because of which she moves from a state of symptomatic HIV to asymptomatic HIV. One of the effects of antiretroviral therapy is that asymptomatic HIV also prevents transmission from mother to child during pregnancy and hence, a child's lifetime amount of future treatment costs and QALY losses could be prevented as well. (28) Assuming that the woman will have future children according to the average European fertility rate, we consider how including the future related costs saved (and QALYs gained) of preventing HIV transmission on this expected future offspring impacts the ICER. The effect on the calculated cost-effectiveness is substantial: for an 18-year-old woman, the ICER decreases with 80% without discounting, 58% at a 3% discount rate and 33% at a 10% discount rate. Also for other transmittable diseases (29) and in different health domains such as perinatal mental health (30), substance abuse (31) or diabetes (32), interventions can generate significant long-term health gains for future offspring accompanied by substantial related future medical savings. Of course, including unrelated future costs can offset these calculated savings, as higher birth rates or longer life expectancies will in contrast increase total costs again. Fig. 1 illustrates how the birth of a child also gives rise to a new trajectory of unrelated future medical costs.

Furthermore, many other societal costs can become relevant from a broader, societal perspective. Future lives will for example also consume non-medical goods and require education and pensions, but they will also generate tax revenues through labour income.(33)[3] The total net effect of including future offspring on cost-effectiveness is therefore ambiguous and will depend on the applied perspective, as well as on the difference between the incremental cost/QALY ratio generated for the parent compared to that for the future child. Nevertheless, we have shown that this effect can be substantial and should not be ignored without further discussion.

[3] Interestingly, from a societal perspective, including future offspring costs might reverse (to a certain degree) the finding that younger, fertile patients are the most cost-effective group to save. While they themselves generate productivity for society, early life medical costs of their children can outweigh these children's discounted later productivity revenue (as for example illustrated by Kruse et al.).(34)

Table 2: 'Back-of-the-envelope' cost-effectiveness calculation of antiretroviral therapy treatment for HIV-positive women in Europe, with and without the costs saved and QALYs gained by avoiding HIV-transmission on (average expected) future offspring. *We assume that* antiretroviral therapy *treatment per patient costs €1000 per year* (36) *and that this treatment generates 1) a utility gain for women moving from symptomatic HIV to asymptomatic HIV of 0.12/year, and 2) a utility gain for future offspring whose HIV infection is prevented of 0.185/year per future child.*(37) *In case the infection of future offspring is prevented, we include future related medical savings because the child will not require treatment (hence €1000/year per child is avoided). We assume a reduction in average mother-to-child transmission rate from 30% to 0%* (38) *and an average life expectancy of 80 for all cases. The women receiving* antiretroviral therapy *are expected to have future children according to the European fertility rate by age (2019).*(39) *Equal discount rates (0%, 3% and 10%) are applied for both costs and effects.*

	Woman at age 18			**Woman at age 25**			**Woman age 40**		
	0%	3%	10%	0%	3%	10%	0%	3%	10%
Sum future related medical costs woman	63.000 €	29.000 €	10.973 €	56.000 €	27.774 €	10.947 €	41.000 €	24.115 €	10.779 €
Sum future related medical costs (here savings) by preventing antiretroviral therapy treatment of future offspring[4]	- 37.622 €	- 10.180 €	- 1.793 €	- 31.918 €	- 10.104 €	- 2.431 €	- 1.753 €	- 647 €	- 209 €
Total sum future related costs	**25.378 €**	**18.820 €**	**9.180 €**	**24.082 €**	**17.670 €**	**8.516 €**	**39.247 €**	**23.467 €**	**10.570 €**
Sum future QALYs gained woman	7,6	3,5	1,3	6,7	3,3	1,3	4,9	2,9	1,3
Sum future QALYs gained future offspring[5]	7,8	1,9	0,3	6,6	1,9	0,4	0,4	0,1	0,0
Total sum future QALYs	**15,4**	**5,4**	**1,6**	**13,4**	**5,3**	**1,8**	**5,3**	**3,0**	**1,3**

[4] For each life year, we multiply the woman's expected offspring (based on the average EU fertility rate) by the expected savings of avoiding annual antiretroviral therapy treatment. For example, at the age of 18 the expected yearly savings of future offspring equal: 0.01294 (fertility rate at 18y)*0.3 (reduction in HIV transmission rate)*€1000 (treatment costs per patient avoided) = €3.88. At the age of 19, estimated yearly savings equal: €3.88 (related to the expected offspring at age 18) + (0.02006 (fertility rate at 19y)*0.3*€1000) = €9.90. Per subsequent year, savings are summed and discounted.

[5] The yearly QALY gain is calculated in the same manner, at the age of 18 the estimated QALY gain equals: 0.01294 (fertility rate at 18y)*0.3 (reduction in HIV transmission rate)*0.185 (yearly QALY gain because of avoided infection) = 0.0007 QALY. Per subsequent year, the QALYs gained are summed and discounted.

	Woman at age 18			Woman at age 25			Woman age 40		
	0%	3%	10%	0%	3%	10%	0%	3%	10%
ICER excl. future offspring [€/QALY]	8333	8333	8333	8333	8333	8333	8333	8333	8333
ICER incl. future offspring [€/QALY]	1650	3478	5568	1803	3366	4829	7427	7779	7934
% difference	**-80%**	**-58%**	**-33%**	**-78%**	**-60%**	**-42%**	**-11%**	**-7%**	**-5%**

Figure 1: Illustration of how having a child (e.g. at the age of 25) gives rise to a new lifetime trajectory of future unrelated medical costs. ***These yearly future unrelated costs are estimated using the PAID 3.0 model[6], using the following parameters: we assume no related diseases, a life expectancy of 80 years, the consideration of all healthcare providers and we take the average between male and female costs for the newborn.***



[6] This tool is freely accessible via https://imta.shinyapps.io/PAID3/ (calculations were made in June 2021).

2.2.3 A proposal

In view of the foregoing, we can conclude that future offspring can have a large, potentially decision-changing impact on cost-effectiveness results and therefore, a systematic approach on how to consider them in future cost calculations is required. However, systematically including all future offspring seems to open a 'Pandora's box'. Besides practical difficulties, it raises many equity issues, one being that it would affect women differently from men, making the same intervention much more/less cost-effective depending on the sex and age of the targeted patient group. On the other hand, systematic exclusion of these costs would be inaccurate and without justification it would arbitrarily favour certain interventions and disadvantage others.

In the first part of this chapter, we have discussed what to do with the QALYs of future offspring. We have suggested to differentiate future lives depending on the decision-context in which they are considered, based on the population ethics concept of necessitarianism. We distinguish *necessary future lives*, the lives that will come into existence independent of the intervention under evaluation, from *potential future lives*, which would only occur as a direct consequence of the intervention. Necessary lives would be present in both the intervention and the comparative scenario, whereas potential lives would only emerge in one of both. Necessary lives are for instance those affected by maternal vaccination or supplementation programs for pregnant women, as these children would be born anyway but their QALE would be different. However, potential lives would be those that are 'brought into existence', e.g. as a result from in vitro fertilization, preventing miscarriages or lifesaving interventions of potential parents. They are 'potential' as their existence depends purely on whether the intervention is carried out, or not. They would be part of the intervention's consequences but remain absent from the comparator. Counting these potential lives in ICERs would heavily skew cost-effectiveness results,

in a way that seems unjustified and counterintuitive. For instance, an IVF program would generate a lifetime of QALYs on top of the health gains for parents and as such, it would outcompete most other healthcare programs in terms of health benefits generated. For necessary lives however, the impact on the ICER would be much smaller and it seems more reasonable to include these effects. These children would be born anyhow (and hence should be considered in both intervention and comparator) but their QALE would be differently affected. In line with this proposal, we would suggest to only count the future offspring costs attached to necessary future lives and to ignore those stemming from potential future lives.

What are the implications of this proposal for future cost calculations in practice? It means that if an intervention affects the probability of a patient to have future children, i.e. when fertility itself would be affected, associated future offspring costs can be excluded because this offspring is merely potential. This would be the case for e.g. contraception, abortion, fertility treatments, hysterectomies and all interventions affecting a patient's survival. In these examples, future generations would only create costs (and QALYs) in either the intervention or comparator side of the ICER, but not in both. Therefore, in the majority of cases, the practice of ignoring future offspring costs in economic evaluation is justified and the developed future cost calculation methods will provide correct estimates. However, for certain interventions, where offspring has a necessary nature, where it would emerge in both intervention and comparator and where its future QALE would be differentially affected, QALYs as well as costs of future offspring should be counted. This is the case for prevention of vertically transmittable diseases, diabetes treatment, substance abuse prevention or mental health interventions in so far as these interventions do not affect fertility as such but do affect the long-term health of future offspring.[7] Our rule also

[7] Often, these interventions can affect both fertility and QALE, in this case, only the QALYs and costs related to the latter should be included.

applies to many public health programs such as disease eradication, air quality policies or environmental programs where necessarily existing future generations will be affected in their health-related quality of life. In all these examples, future offspring will be born necessarily but the future (medical) costs occurring throughout their lives will be different in the comparator and intervention group.

Of course, this proposition has limitations as well (also see Part 1 of this chapter). One of them is that, omitting significant costs of potential future lives still entails ignoring real changes in future healthcare budgets (at least for one of the comparators). As a consequence, the actual economic implications of a decision could still be misrepresented. Specifically in cases were potential lives will be born with severe disability or are expected to live in very poor health, future medical (and societal) costs will be substantial. However, in contrast to the current practice of ignoring these costs by default, at least now such exclusions would be based on an argumentation. Second, including the costs of necessary future offspring will also increase decision-making uncertainty and therefore more evidence on the long-term consequences of healthcare on future offspring is required. A third observation is that in this paper we only considered the first, unborn future generation, because we assumed the use of typical discount rates (around 3%) leading to presumably negligible impacts of subsequent future generations. While the discussion of an appropriate discount rate for future generations is out of the scope of this paper, this factor will also significantly affect the role and impact of future offspring overall and requires further discussion.

2.2.4 Conclusion

Guidelines increasingly prescribe to include all future costs in order to provide a (fairly) complete estimate of the resource implications of funding decisions on the health system and beyond. As we have argued, future costs of offspring are neglected in these methods, although they can be significant and can have a differential impact across different patient groups and interventions. This raises the question whether correct calculation methods should include future offspring costs or not. Based on the concept of necessitarianism, which distinguishes the relevance of future lives to policy decision-making based on the 'necessity' of their existence, we have suggested to only include future costs in case they belong to future offspring whose existence is not dependent on the evaluated intervention. This rule justifies the current practice of excluding future generations in most cases. However, for these interventions where offspring will be born anyway but with a different QALE, the incremental future costs generated by this future offspring need to be considered. This is mostly relevant for ante- or perinatal health interventions, or for broad public health programs that can have intergenerational effects. On the whole, future generations expose a substantial gap in today's methodology and require more attention from the HTA community.

References

1. Kearns B. NICE DSU Report. The relevance of future, unrelated health costs in economic evaluation in NICE appraisals. [Internet]. Decision Support Unit, ScHARR, University of Sheffield; 2020 p. 48. Available from: http://nicedsu.org.uk/wp-content/uploads/2020/11/Future_unrelated_costs_Final.pdf
2. Morton A, Adler AI, Bell D, Briggs A, Brouwer W, Claxton K, et al. Unrelated Future Costs and Unrelated Future Benefits: Reflections on NICE Guide to the Methods of Technology Appraisal. Health Econ. 2016;25(8):933–8.
3. van Baal P, Morton A, Meltzer D, Brouwer W. Future unrelated medical costs need to be considered in cost effectiveness analysis. Eur J Health Econ. 2019 Feb 1;20(1):1–5.
4. de Vries LM, van Baal PHM, Brouwer WBF. Future Costs in Cost-Effectiveness Analyses: Past, Present, Future. Pharmacoeconomics. 2019;37(2):119–30.
5. van Baal P, Morton A, Brouwer W, Meltzer D, Davis S. Should cost effectiveness analyses for NICE always consider future unrelated medical costs? BMJ. 2017 Nov 10;j5096.
6. Nord E, Lamøy C. Including Future Consumption and Production in Economic Evaluation of Interventions that Save Life-Years: Commentary. PharmacoEconomics Open. 2018 Dec 1;2(4):357–8.
7. McCabe C. Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. PharmacoEconomics. 2019 Apr 1;37(4):457–60.
8. Yock CA, Boothroyd DB, Owens DK, Garber AM, Hlatky MA. Cost-effectiveness of bypass surgery versus stenting in patients with multivessel coronary artery disease. The American Journal of Medicine. 2003 Oct 1;115(5):382–9.
9. Garber A, Phelps C. Economic foundations of cost-effectiveness analysis. Journal of Health Economics. 1997 Feb 1;16(1):1–31.
10. Weinstein MC, Fineberg HV. Clinical Decision Analysis. Saunders; 1980. 376 p.

11. Neumann P, Sanders G, Russell L, Siegel J, Ganiats T. Cost-effectiveness in health and medicine. 2nd ed. New York: Oxford University Press; 2017. 496 p.
12. NICE. 5 The reference case | Guide to the methods of technology appraisal 2013 | Guidance | NICE [Internet]. 2013 [cited 2019 Jun 11]. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case
13. van Baal PHM, Wong A, Slobbe LCJ, Polder JJ, Brouwer WBF, de Wit GA. Standardizing the Inclusion of Indirect Medical Costs in Economic Evaluations. Pharmacoeconomics. 2011 Mar 1;29(3):175–87.
14. Briggs ADM, Scarborough P, Wolstenholme J. Estimating comparable English healthcare costs for multiple diseases and unrelated future costs for use in health and public health economic modelling. PLOS ONE. 2018 May 24;13(5):e0197257.
15. Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. Journal of Health Economics. 1997 Feb 1;16(1):33–64.
16. de Boer PT, Backer JA, van Hoek AJ, Wallinga J. Vaccinating children against influenza: overall cost-effective with potential for undesirable outcomes. BMC Med [Internet]. 2020 Jan 14 [cited 2021 Jan 28];18. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6958762/
17. Ramos IC, Versteegh MM, Boer RA de, Koenders JMA, Linssen GCM, Meeder JG, et al. Cost Effectiveness of the Angiotensin Receptor Neprilysin Inhibitor Sacubitril/Valsartan for Patients with Chronic Heart Failure and Reduced Ejection Fraction in the Netherlands: A Country Adaptation Analysis Under the Former and Current Dutch Pharmacoeconomic Guidelines. Value in Health. 2017 Dec 1;20(10):1260–9.
18. Walker S, Asaria M, Manca A, Palmer S, Gale CP, Shah AD, et al. Long-term healthcare use and costs in patients with stable coronary artery disease: a population-based cohort using linked health records (CALIBER). Eur Heart J Qual Care Clin Outcomes. 2016 Apr;2(2):125–40.
19. Meltzer D, Egleston B, Stoffel D, Dasbach E. Effect of future costs on cost-effectiveness of medical interventions among young adults: The example of intensive therapy for type 1 diabetes mellitus. Medical Care. 2000;38(6):679–85.

20. Kellerborg K, Perry-Duxbury M, de Vries L, van Baal P. Practical Guidance for Including Future Costs in Economic Evaluations in The Netherlands: Introducing and Applying PAID 3.0. Value in Health. 2020 Nov 1;23(11):1453–61.
21. NICE. The guidelines manual: Process and methods - 7 Assessing cost effectiveness [Internet]. The guidelines manual. 2012 [cited 2018 Jun 20]. Available from: https://www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness
22. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
23. Swedish Council on Health Technology Assessment (SBU), EUnetHTA. Methods for health economic evaluations - A guideline based on current practices in Europe. 2015;100.
24. Hoek J, Koster MPH, Schoenmakers S, Willemsen SP, Koning AHJ, Steegers EAP, et al. Does the father matter? The association between the periconceptional paternal folate status and embryonic growth. Fertility and Sterility. 2019 Feb 1;111(2):270–9.
25. Dimofski P, Meyre D, Dreumont N, Leininger-Muller B. Consequences of Paternal Nutrition on Offspring Health and Disease. Nutrients. 2021 Aug;13(8):2818.
26. Schagdarsurengin U, Steger K. Epigenetics in male reproduction: effect of paternal diet on sperm quality and offspring health. Nat Rev Urol. 2016 Oct;13(10):584–95.
27. Beal MA, Yauk CL, Marchetti F. From sperm to offspring: Assessing the heritable genetic consequences of paternal smoking and potential public health impacts. Mutation Research/Reviews in Mutation Research. 2017 Jul 1;773:26–50.
28. Duma N, Lambertini M. It Is Time to Talk About Fertility and Immunotherapy. Oncologist. 2020 Apr;25(4):277–8.
29. Terris-Prestholt F, Watson-Jones D, Mugeye K, Kumaranayake L, Ndeki L, Weiss H, et al. Is antenatal syphilis screening still cost effective in sub-Saharan Africa. Sex Transm Infect. 2003 Oct;79(5):375–81.

30. Bauer A, Knapp M, Parsonage M. Lifetime costs of perinatal anxiety and depression. Journal of Affective Disorders. 2016 Mar 1;192:83–90.
31. Viteri OA, Soto EE, Bahado-Singh RO, Christensen CW, Chauhan SP, Sibai BM. Fetal Anomalies and Long-Term Effects Associated with Substance Abuse in Pregnancy: A Literature Review. Am J Perinatol. 2015 Apr;32(5):405–16.
32. Peterson C, Grosse SD, Li R, Sharma AJ, Razzaghi H, Herman WH, et al. Preventable health and cost burden of adverse birth outcomes associated with pregestational diabetes in the United States. American Journal of Obstetrics and Gynecology. 2015 Jan 1;212(1):74.e1-74.e9.
33. Petrou S, Sach T, Davidson L. The long-term costs of preterm birth and low birth weight: results of a systematic review. Child: Care, Health and Development. 2001;27(2):97–115.
34. Kruse M, Sørensen J, Gyrd-Hansen D. Future costs in cost-effectiveness analysis: an empirical assessment. Eur J Health Econ. 2012 Feb;13(1):63–70.
35. Luyten J, Verbeke E, Schokkaert E. To be or not to be: future lives in economic evaluation. Health Economics. 2021;
36. European Centre for Disease Prevention and Control. HIV treatment, care and support: monitoring implementation of the Dublin Declaration on Partnership to Fight HIV/AIDS in Europe and Central Asia : 2017 progress report. [Internet]. LU: Publications Office; 2017 [cited 2021 Jun 8] p. 13. Available from: https://data.europa.eu/doi/10.2900/94617
37. Tengs TO, Lin TH. A Meta-Analysis of Utility Estimates for HIV/AIDS. Med Decis Making. 2002 Dec 1;22(6):475–81.
38. WHO | Mother-to-child transmission of HIV [Internet]. WHO. [cited 2021 Jun 8]. Available from: http://www.who.int/hiv/topics/mtct/about/en/
39. Eurostat. Fertility rates by age [Internet]. Eurostat Data Browser. 2021 [cited 2021 Sep 13]. Available from: https://ec.europa.eu/eurostat/databrowser/view/demo_frate/default/table?lang=en

Chapter 3: Valuing future lives in healthcare priority-setting: a population survey

Abstract

Objectives: Resource allocation decisions made today not only affect presently existing people, but at times also people that will be born in the future. Inclusion of such future life effects can have important consequences for the cost-effectiveness assessment of health interventions, but it remains an open question whether they should be taken into account in economic evaluation studies. We investigate the public's opinion about this value judgement.

Methods: A survey was executed in a representative sample of the Dutch (n=930) and Flemish (n=934) population. Respondents were asked to prioritize treatment between two identical women in equal need of healthcare, but who differed in terms of their affected future offspring. Questions varied in terms of the development stage of the affected future offspring and the type of intervention effect (affecting the coming into existence or the Quality of Life of the future offspring).

Results: Up to 70% of respondents did not seem to value the affected future offspring when setting healthcare priorities. When one of the women was already pregnant or just gave birth, 40% of respondents still ignored her affected child. Based on individual choice patterns, we identify 60% of respondents as 'presentists' (only valuing present lives) and 20% as 'actualists' (always taking into account the future offspring),

20% made diverging choices that were more difficult to interpret. Our results do not suggest a clear support for a ‘necessitarian’ view.

Conclusion: We find that, in a healthcare priority-setting context, most people focus on present lives. People justified their indifference with regard to the future offspring by their refusal to discriminate between present patients in equal need of treatment. Further research is required to explore the robustness of our findings in different contexts.

This chapter will be submitted as:

Verbeke, E., Schokkaert, E., Luyten, J., Valuing future lives in healthcare priority-setting: a population survey.

3.1 Introduction

In many policy domains such as environmental protection or climate change mitigation, resource allocation decision-making in the present also affects people that will live in the future. Also in healthcare these resource allocation decisions can have effects on future lives, referred to in this chapter as *'future life effects'* (FLEs). Examples are fertility treatments, prenatal genetic screening, disease eradication programs or public health programs with a distant time horizon, where the health of present people is affected, but also that of future generations. It remains an open question whether the cost-effectiveness of these programs should be based only on the consequences for present people or also on those for future generations. As argued in the previous chapter, including the latter can have a decision-changing impact for the cost-effectiveness of a range of programs.(1,2) Currently, future lives are mostly ignored or treated in heterogeneous manners which hampers comparability of cost-effectiveness results.(3) Also theoretically there is no straightforward solution regarding how to fairly incorporate future generations.(4,5) Table 1 summarizes the previously discussed mutually exclusive viewpoints on intergenerational fairness and indicates how they could be implemented in Health Technology Assessment (HTA).

Whether or not to include future generations in cost-effectiveness ratios is an important social value judgement that needs to be made by HTA researchers. A key element in this debate that has, to our knowledge not been studied before, is how the general public thinks about this issue. This is however important for several reasons. First, policy-makers should be able to take into account the public opinion to ensure that priority-setting decisions are considered acceptable and justifiable by the general public.(6–8) Second, as every person is a potential future patient that might require healthcare resources, the general public is personally involved in healthcare priority-setting decisions. It seems therefore only fair to take into account the opinion of the people who (or whose future offspring) might bare the opportunity costs related to budget allocation

decisions.(8,9) Third, studying public preferences can reveal important dimensions that people take into account when setting healthcare priorities, but that might be overlooked in the theoretical literature. Last, the way in which people's socio-demographic characteristics are related to their ethical view can provide valuable insights on how decision-making needs to be adjusted in specific environments.(8)

Previous empirical evidence has investigated related areas, with the largest share of studies exploring societal time preferences. In these studies, the 'discount rate' that people apply to trade-off lives saved today, versus lives saved at different points in the future, is estimated. (10–16) For example, Cropper et al. found that respondents valued one presently existing life equally valuable as 45 lives within 100 years in the future.(12) Considering the same time horizon, Johansson et al. found a ratio of 1 present for 243 future lives.(15) While these examples suggest that people value future lives less, in a context of environmental policies it has been shown that most people care (equally or more) about future generations and would even be willing to reduce their standard of living so that people can enjoy better lives in the future.(13,17) However, this does not mean that in a public policy context, respondents will typically prioritize policies with long-term benefits over short-term programs, as uncertainty and doubts about long-term political commitments can conversely lead to negative time-preferences.(16,17) Furthermore, these studies typically do not distinguish between interventions that save lives in the future, and interventions affecting the (probability of) coming into existence, or the QoL of future lives. In the field of population ethics, this distinction has been considered in studies that explore how people rank future populations that vary in terms of size and welfare levels.(18–20) Because these studies only study different future populations, they do however not inform us about public preferences in situations where both present and future people are affected by a program. Finally, in the specific context of HTA, studies have explored the relative weight that people attribute to people of

different ages, including children.(14,15,21–23) Here, evidence shows that the perceived value of (present) lives decreases with age, implying that younger children should be prioritized. One study considered young fetuses and found that they were valued less than children.(21) Lives before conception were however not considered and also here no distinction was made between morbidity, mortality or coming into existence. This evidence is therefore found to be insufficient to conclude if, and when, the general public takes into account the different types of effects of healthcare on future lives in a healthcare priority-setting context.

Our aim is to address this knowledge gap and to provide evidence on how laypersons intuitively value future lives when they are asked to allocate scarce healthcare resources. We developed a survey to study if, and in which circumstances, people think that future lives should be accounted for in healthcare priority setting. We explore how the public's choice behaviour relates to the ethical views summarized in Table 1. And we analyse whether respondents with different views significantly differ in terms of socio-demographic characteristics.

Table 1: Key-ethical views about the value of future lives and their implications in an HTA context.

View	Practical implications for HTA	Main arguments in favor	Main arguments against
Presentism (4,5) Only presently existing people have value.	All future life effects (FLEs) are excluded.	- Future lives do not yet exist and their coming into existence is too uncertain to take them into account - Present patients should not be discriminated based on their potential future offspring	- Implies a complete indifference with regard to future people's existence and health - Requires a consensual boundary that determines when a life becomes present and its lifetime QALYs suddenly become valuable
Actualism (4,5) All lives that can become actual now, or in the future, matter.	All FLEs are included.	- All QALYs have equal value and future lives should therefore be treated in the same impartial manner - An evaluation should be based on complete information in order to ensure efficient	- Including the long tail of consequences related to e.g. fertility treatment or disease eradication is likely to outweigh health gains for presently existing patients, even with discounting - Life-saving programs or the treatment of vertically

		decision-making throughout time	transmittable diseases become more attractive for young women, compared to older women or men
Necessitarianism (4,5,24,25) A future life matters when it will exist necessarily: this means that it will exist irrespective of our choices (e.g. independent of the policy-decision under consideration). There is no value in creating additional lives itself (these lives are considered merely potential lives).	Only QALYs related to changes in future people's quality-adjusted life expectancy are taken into account. When a future life comes into existence (or does not come into existence) because of an intervention (e.g. fertility treatment or abortion), the related QALYs are not included.	- We have the obligation to protect the interests of the people that will exist in the future - Avoids that every potential future life should be taken into account - Avoids that the coming into existence of future people is considered to be extremely cost-effective (in terms of QALYs)	- Implies an indifference with regard to changes in future fertility rates or future population size - Does not allow the consideration of the suffering that might be created by bringing a life in very bad health into existence - Requires a consensual boundary to determine when a life becomes present and its existence/life expectancy suddenly becomes valuable

3.2 Methods

3.2.1 Sample and data collection

We sent untargeted invitations via email to members of an online panel of the market research agency Dynata to complete our survey between April 26, 2021 and May 1, 2021. This panel consisted of 8000 Dutch members and 5500 Belgian members who mirror the national population (18-80 years) of both countries as well as possible. Dynata continuously evaluates their panel members and eliminates low-quality responders. Participants were sampled to be representative for the adult Flemish and Dutch population in terms of gender, age and geographical spread. With regard to Belgium, only Flemish citizens were included as they represent the Dutch-speaking part of the country. Including two (Dutch-speaking) nationalities avoided the risk of misinterpretations related to the translation of the complex choice tasks in the survey. On the other hand, it allowed us to account for the effect of national differences related to for example cultural values or local reimbursement policies. The Brussels region was excluded from data collection as only 9% of inhabitants is Dutch-speaking. We aimed to target 1000 respondents for each country, to obtain a comparable or larger sample as empirical studies in related fields.(10–15,17–20,26) Participation was rewarded by incentive options, which could be redeemed as gift cards, charitable contributions or partner products and services. For the initial pilot launch, 10% of this sample was recruited.

3.2.2 Survey

The online survey consisted of three parts. The first, introductory part comprised background information on our research and ethical approval, followed by an informed consent page. Respondents were asked about their general socio-demographic characteristics and whether they have children, or would like to have children in the future. We also elicited respondents' opinions about the degree to which they believe we should care about the wellbeing of future generations in general. After this part,

respondents were introduced to the issue of scarce healthcare budgets, the necessity to set funding priorities and the focus of the study on interventions affecting future lives and how to value them. Respondents were asked to adopt the perspective of an 'advisor to the national Minister of Health' who decides on which programs should be funded. The second part consisted of a specific choice task discussed below. As a final part of the survey, we asked respondents about the comprehensibility of the questions and we presented them a simple sum (15+27=?) as a control question. Respondents also had the opportunity to provide any further remarks about the survey.

The choice task aimed to investigate 1) *if* respondents take into account the FLEs of an intervention when prioritizing presently existing patients, 2) whether this depends on the development stage during which the future life is affected by the intervention, and 3) whether the nature of the FLE itself (changes in conception/birth rates versus future QoL) matters. We defined two scenario's in which respondents were asked to choose between two identical women of childbearing age in equal need of treatment, of which only one will have a future child that also benefits from the treatment. Therefore, if respondents value future lives, we expect them to prioritize the woman with future child.

In scenario one, two female patients were competing for a hypothetical medicine called 'Rinutek', curing a disease that otherwise leads to death within one month. The first woman (A) will remain childless, the other woman (B) will have a child (now, in the near future or within a decade). Allocating the treatment to woman A would save her life, whereas allocating the treatment to woman B would save both her life and that of her future child. Figure 1 shows how we presented this scenario, including a graphical depiction to aid understanding.
In order to study whether the type of intervention effect matters, the second scenario considered the same women but introduced a different medicine 'Jerutek': curing a vertically transmittable disease that causes

lifelong disability[8] (Figure 2). When the treatment is given to woman A, her illness would be cured. By giving it to woman B, she would be cured and would also no longer transmit the disease to her future child.

Per scenario, respondents were asked to make five choices between the childless woman (A) and the future mother (B). The variation in the questions was in the development stage of the future child of woman B, namely: 1) woman B just gave birth[9], 2) was 7 months pregnant, 3) was 1 week pregnant, 4) would have a child within 1 year, or 5) within 10 years. Respondents had to indicate whether they prioritized the childless woman, the woman with a future child, or they could be indifferent between both. In addition, respondents were forced to explain every choice in a comment box, allowing us to understand their reasoning. The Rinutek and Jerutek scenario were presented to respondents in a random order to avoid any order effects. However, within every scenario, the same sequence of questions (each considering a subsequent and more abstract development stage of the future child) was presented to ease interpretation.

[8] We defined disability as a 50% loss in quality of life compared to perfectly healthy people.
[9] In this case, the treatment also cured her newborn child via antibodies in the mother's breast milk.

Figure 1: Sample question for scenario 1 in which mortality is prevented. This question was repeated 5 times, considering the child of woman B during 5 different development stages.

RINUTEK ®

In the following scenarios we assume that a new medicine (called RINUTEK) was developed to cure a deadly disease that affects young women. Because of the disease, ill women will **die within one month. RINUTEK is completely safe and cures the women with 100% certainty.**

The number of available RINUTEK treatments is however limited. **If you are forced to choose, which of the following two women do you think should be prioritized to receive a RINUTEK treatment?**



An 	An is 25 years old and **will not have any kids** in the future.
Bea 	Bea is 25 years old and **7 months pregnant. The child she is carrying is also saved by the treatment.** Afterwards, Bea will not have any more children.

a) To who do you give the life-saving Rinutek treatment?

- ○ I think An should be prioritized
- ○ I think Bea should be prioritized
- ○ I think both are equally important and I cannot choose

b) Why did you make this choice?

Figure 2: Presentation of scenario 2 in which future disability is prevented.

JERUTEK®

In the following scenarios we assume that a new medicine (called JERUTEK) was developed to cure a disease that affects young women. Because of the disease, **ill women will experience lifelong disability and only 50% of quality of life compared to a healthy person.**

In practice, this means that ill women will experience problems with walking, cleaning and dressing themselves and daily activities such as working or studying. They experience pain or discomfort and are anxious or depressed.

Without threatment they will remain **ill for the rest of their lives**. When an ill woman gives birth to a child, **this child will also experience lifelong illness. The disease is hereditary. JERUTEK is completely safe and cures the women with 100% certainty.**

The number of available JERUTEK treatments is however limited. **If you are forced to make a choice, which of the following two women do you think should be prioritized to receive a JERUTEK treatment?**



An 	An is 25 years old and **will not have any kids** in the future.
Kaat 	Kaat is 25 years old and **1 week pregnant,** if she is cured, **the transmission of the disease to her future child is also avoided.** Afterwards, Kaat will not have any more children.

3.2.3 Data analysis

We start by summarizing the sample's choices in terms of proportions. Using Mc Nemar tests, we analyse whether these proportions significantly change when the development stage, intervention effect or the order of the scenarios varied.

Subsequently, we study individual choice patterns to identify the underlying ethical view of respondents. First, we define the choice pattern that we expect to correspond to the discussed ethical views. As shown in Table 2, we consider different possible boundaries that respondents might adopt to determine when a future life becomes a present one. Respondents are allocated to a perspective, based on the lowest number of dissimilarities between their choices and these pre-defined choice patterns. We consider the comments of respondents that represent 'a perfect match' with the pre-defined profiles, to verify whether their motivations indeed correspond to the assumed ethical view.

Table 2: Expected choice profiles. ✓: priority is given to the women with future offspring, 0: respondent is indifferent between both women. ∃: the intervention affects future existence, QoL: the intervention affects future people's quality of life.

Choice profile	**Newborn**		**7 months pregnant**		**1 week pregnant**		**Child within 1y**		**Child within 10y**	
	∃	QoL	∃	QoL	∃	QoL	∃	QoL	∃	QoL
Actualists	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Presentists										
Boundary = conception	✓	✓	✓	✓	✓	✓	0	0	0	0
Boundary= late pregnancy	✓	✓	✓	✓	0	0	0	0	0	0
Boundary = after birth	✓	✓	0	0	0	0	0	0	0	0
Necessitarians										
Boundary = conception	✓	✓	✓	✓	✓	✓	0	✓	0	✓
Boundary= late pregnancy	✓	✓	✓	✓	0	✓	0	✓	0	✓
Boundary = after birth	✓	✓	0	✓	0	✓	0	✓	0	✓

Second, we analyse respondent's choice patterns using a k-medoids cluster analysis to study whether we overlooked choice profiles (and underlying views) that might correspond better with people's prioritization behaviour. Rather than comparing respondents to pre-defined profiles, this method categorizes respondents into groups (clusters) by minimizing the sum of dissimilarities between the individual responses themselves. The identified clusters therefore represent the most common and similar response patterns within the sample. In case the cluster analysis identifies unexpected choice patterns, we study respondent's comments in order to better understand their underlying view.
We assess the accuracy of these two allocation methods by reporting for each question the proportion of respondents within a profile whose choice deviates from the expected/cluster answer.

Last, using multinomial logistic regression, we investigate whether socio-demographic variables can predict respondent's profile membership. Answers to the introductory questions about future generations' wellbeing are summarized into one overall score, which was considered as an additional variable in the estimated models. We report odds ratios, indicating the odds that a respondent with a certain characteristic will belong to a specific response profile.
Analyses were conducted using Excel and R studio version 1.4.1717.

3.3 Results

3.3.1 Sample

A total of 1090 (FL) and 1125 (NL) responses was collected. To ensure high quality data, we screened comment boxes during collection and excluded respondents that provided nonsensical answers (consisting of non-existing or completely irrelevant words) immediately from the sample. Also respondents indicating that they did not understand the assignment were excluded. Data collection ended once we obtained a group of more than 1000 'serious' respondents for each region. We excluded participants that completed the survey in less than 4,5 minutes (n=71 FL and n=75 NL) (which equals one third of the median completion time for the entire sample – a standard criterion used by the market research company). Last, we conducted an additional quality check based on the comments provided by respondents that systematically indicated 'I don't know' in the introductory opinion questions and respondents that answered the control question wrong. This resulted in the additional exclusion of 101 (FL) and 123 (NL) respondents. We obtain a remaining sample of 934 Flemish and 930 Dutch respondents, whose characteristics are summarized in Table 3. We compare these to national population data of both countries (see Appendix) and find an underrepresentation of respondents with lower education as well as a smaller share of religious respondents in the Dutch sample (39%) compared to national numbers (47%). We however conclude that our sample is overall sufficiently representative.

Table 3: Sample characteristics, FL= respondents living in Flanders, NL= respondents living in the Netherlands.

CHARACTERISTIC		FL (BE) (n=934) %	NL (n=930) %
GENDER	Male	48%	46%
	Female	52%	54%
	Other	0%	0%
AGE	18-24y	10%	8%
	25-34y	13%	14%
	35-44y	17%	18%
	45-64y	37%	38%
	65+y	23%	22%
EDUCATION	None or primary	4%	2%
	Secondary	47%	52%
	Higher education	49%	46%
HOW OFTEN DO YOU HAVE PROBLEMS PAYING FOR MONTHLY HOUSEHOLD EXPENSES?	Never	58%	60%
	Once a year	17%	19%
	Once in 3 months	13%	12%
	Every month	12%	9%
MARITAL STATUS	Single	27%	28%
	Cohabiting/married	60%	62%
	Divorced	10%	7%
	Widow(er)	3%	3%
SELF-PERCEIVED HEALTH STATE	Very good/good	52%	52%
	Reasonably good	34%	35%
	Reasonably bad	6%	4%
	Very bad/bad	7%	8%
	I don't know/ do not wish to answer	1%	1%
RELIGIOUS	Yes	49%	39%
	No	51%	61%
RELIGION	Christian	42%	32%
	Muslim	3%	3%
	Hindu	0%	1%
	Atheist, Buddhist or Jewish	1%	0%
	Other	2%	2%
	Not religious	51%	61%
	I do not wish to answer	1%	1%

DO YOU HAVE CHILDREN?	No	37%	40%
	Yes	63%	60%
TO WHICH CATEGORY DO YOUR CHILDREN BELONG?	1 (biological) child	18%	14%
	2 (biological) children	28%	31%
	3+ (biological) children	15%	12%
	Adopted children	0%	0%
	Stepchildren	7%	5%
	Grandchildren	28%	26%
DO YOU HAVE A CHILDWISH?	No kids and no childwish	11%	16%
	Kids and/or childwish	79%	74%
	No kids + I don't know if I want children/ having children is not an option due to my age	10%	10%
EMPLOYMENT	Active on the labour market	51%	51%
	Retired	27%	19%
	Housewife/husband	4%	7%
	Disabled	6%	4%
	Student	7%	6%
	Unemployed	3%	6%
	Other	2%	7%
OPINION ABOUT FUTURE GENERATIONS	I'm worried about the wellbeing of people that will be born in the future.		
	Fully agree/ agree	84%	73%
	Fully disagree/disagree	12%	20%
	I don't know	4%	6%
	People that will be born in the future will be better off than us.		
	Fully agree/ agree	21%	29%
	Fully disagree/disagree	62%	51%
	I don't know	17%	20%
	We need to make sacrifices to protect the wellbeing of future people.		
	Fully agree/ agree	72%	72%
	Fully disagree/disagree	18%	18%
	I don't know	10%	10%
	Overpopulation is a big societal problem.		
	Fully agree/ agree	83%	81%
	Fully disagree/disagree	10%	12%
	I don't know	7%	7%

COMPREHENSIBILITY OF THE SURVEY	I did not understand the assignment and my answers are invalid (= excluded from analysis).	0%	0%
	I thought the questions were clear and easy to answer.	28%	30%
	I thought the questions were clear, but difficult to answer.	63%	62%
	I though the questions were unclear and difficult to answer.	3%	2%
	Other	6%	6%

3.3.2 Choice task

Aggregated choices

Figure 2: Choice frequencies: % of respondents that prioritized the female patient without future offspring, with future offspring or that were indifferent between both patients.

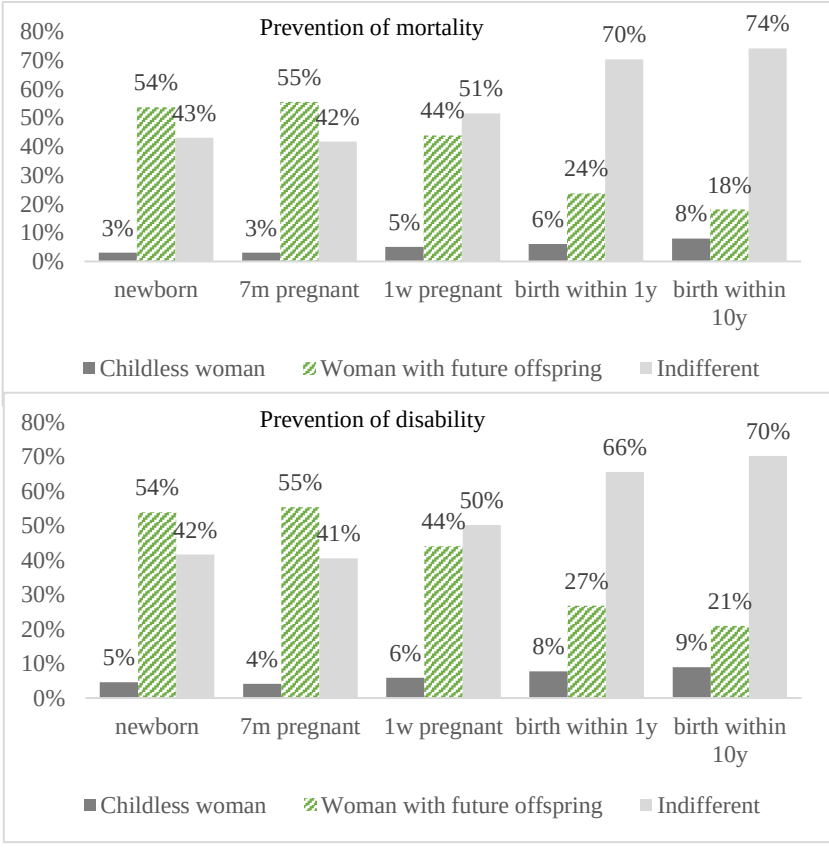


Figure 2 shows the percentages of respondents that prioritized the childless woman, the woman with future offspring or were indifferent between both. The majority of respondents (~70%) was indifferent towards the future offspring before conception. Once the women was pregnant, an increasing proportion (44%) prioritized the woman with child. In contrast, 18-21% of respondents valued the future life in all development stages and systematically prioritized the woman with future offspring. Only 3% of respondents prioritized the childless woman in the first questions, which increased up to 9% when the future offspring became more abstract.

The impact of the intervention effect on choice behaviour seems limited as we find largely similar results in each scenario. Only when the future life effect occurred before conception (i.e. the last two questions in each scenario), we find that a slightly larger proportion of respondents prioritized the woman with offspring in the case of disability prevention compared to the case of mortality prevention (McNemar test, $p<0,001$).

Individual choice profiles

Comparing individual choices with the pre-defined profiles, we find the lowest number of dissimilarities when we assume that respondents consider the child to be a present life once the woman is 7 months pregnant. Table 4 shows that, based on this boundary, 64% of respondents is identified as a presentist, even though only 6% of this group had zero dissimilarities with the pre-defined profile. 26% of respondents is considered to be an actualist (of which 41% had zero dissimilarities). Last, 9% of the sample is identified as a necessitarian (of which 5% with zero dissimilarities).

Table 4 also reports the obtained cluster profiles, we identify actualists (18%) and presentists (57%). Interestingly, we do not find a cluster in which QoL is systematically prioritized, following the pre-defined necessitarian profile. Rather, some respondents indeed prioritize the woman with future offspring more often in the disability prevention case,

but these respondents still become indifferent at the most abstract level (when the child only exists within 10 years). Comparison of the cluster profiles with the pre-defined profiles shows that 79% of respondents is allocated to the same three profiles (of which 17% actualists, 5% necessitarians and 57% presentists). Based on the obtained clusters, we identify two additional profiles. 'Natalists' (9%) have the opposite behaviour of the necessitarian-like category and prioritized the woman with future offspring more often when the intervention affected existence. 'Antinatalists' (3%), mostly prioritized the woman without future offspring.

We find substantially higher within-variations (+50%) for the pre-defined presentist and necessitarian profiles related to the questions that consider the most developed stages of the future child. This illustrates the discrepancy between our expectation that for these questions, all respondents would prioritize the women with her child. In contrast, we find that more than 40% of respondents remains indifferent between both women, even when the child is already born. The variation in the cluster-based profiles is lower as we obtain a separate cluster that includes these systematically indifferent respondents (containing 34% of the total sample).

Overall, our results seem to indicate support for a presentist view according to which only present lives should be taken into account in healthcare priority-setting. As presentist respondents might also have chosen indifference because they simply did not want to answer the question, we verify our interpretation by studying the comments of systematically indifferent respondents (n=527). We find that 98% of these respondents argues that they do not wish to discriminate between both women or find this choice unethical. This confirms that the presentist respondents made the deliberate choice to ignore the additional FLE to avoid problematic (deontological) consequences.

Table 4: Allocation of respondents to the pre-defined and cluster-based profiles for choice task 1. ∃: the intervention affects future existence, QoL: the intervention affects future people's quality of life. ✓: priority is given to the women with future offspring, 0: respondent is indifferent between both woman, ✘: priority is given to the woman without future offspring.

Profiles (% of respondents allocated to this profile)	**Choice per development stage and intervention effect** (within profile variation= % of respondents within cluster with a different response)									
	Newborn		**7 months pregnant**		**1 week pregnant**		**Child within 1y**		**Child within 10y**	
	∃	QoL	∃	QoL	∃	QoL	∃	QoL	∃	QoL
PRE-DEFINED PROFILES (boundary present life = late pregnancy)										
Actualists (26%)	✓ 9%	✓ 16%	✓ 6%	✓ 12%	✓ 5%	✓ 7%	✓ 10%	✓ 10%	✓ 14%	✓ 13%
Presentists (64%)	✓ 30%	✓ 60%	✓ 58%	✓ 60%	0 2%	0 3%	0 2%	0 3%	0 3%	0 4%
Necessitarians (9%)	✓ 56%	✓ 32%	✓ 55%	✓ 30%	0 23%	✓ 24%	0 24%	✓ 30%	0 27%	✓ 35%
CLUSTER-BASED PROFILES										
Actualists (18%)	✓ 8%	✓ 10%	✓ 6%	✓ 6%	✓ 6%	✓ 5%	✓ 5%	✓ 9%	✓ 14%	✓ 11%
Presentists (57 %)										
Boundary = conception (13%)	✓ 8%	✓ 7%	✓ 0%	✓ 1%	✓ 0%	✓ 0%	0 7%	0 0%	0 3%	0 3%
Late pregnancy (11%)	✓ 12%	✓ 8%	✓ 9%	✓ 9%	0 14%	0 16%	0 9%	0 9%	0 9%	0 7%
After birth (34%)	0 4%	0 4%	0 1%	0 1%	0 2%	0 2%	0 2%	0 3%	0 3%	0 3%
'Necessitarians' (14%)										
Only QoL matters after conception (7%)	0 9%	✓ 20%	0 10%	✓ 11%	0 7%	✓ 31%	0 9%	0 38%	0 7%	0 30%
QoL matters earlier than existence (7%)	✓ 7%	✓ 8%	✓ 3%	✓ 5%	✓ 10%	✓ 6%	0 29%	✓ 7%	0 3%	0 7%
'Natalists' (9%)										
Only existence matters after conception (5%)	✓ 11%	0 10%	✓ 4%	0 7%	✓ 32%	0 5%	0 29%	0 8%	0 17%	0 13%
Existence matters earlier than QoL (4%)	✓ 20%	✓ 32%	✓ 12%	✓ 30%	✓ 38%	✘ 45%	✘ 32%	✘ 18%	✘ 30%	✘ 11%
'Antinatalists' (3%)	✘ 19%	✘ 23%	✘ 28%	✘ 23%	✘ 17%	✘ 21%	✘ 26%	✘ 28%	✘ 15%	✘ 15%

Predictors of profile membership

Table 5 presents the multinomial logistic regression estimates of odds ratios for the selected socio-demographic variables and order effects. Respondents that are male ($p<0.001$), religious ($p<0.05$), Dutch ($p<0.01$) and have children or a childwish ($p<0.05$) are found more likely to be categorized as an actualist compared to a presentist.

Antinatalist respondents (identified by the cluster analysis), were more likely to be male (compared to presentists) ($p<0.05$) and to have no children nor a childwish ($p<0.05$). Overall, only 3% of respondents were categorized as 'antinatalist'. However, almost 9% of respondents prioritized the childless woman in the last question when the future child would be born within 10 years. Interestingly, the provided comments show that the motives of these respondents diverge per intervention scenario. For both interventions, respondents often argued that the childless woman should be prioritized because of overpopulation (indeed attributing negative value to procreation), or a too high uncertainty about the other woman's future offspring. However, with regard to the scenario in which future disability was prevented, comments showed that respondents held the woman responsible for having a child in the future, while being aware of her transmittable disease.

Finally, the order in which the intervention scenarios were presented is found to have a significant effect on respondent's choice behaviour (chisq test $p<0.001$). Respondents that were first confronted with the disability prevention case were significantly more likely to value the future life and to be labelled an actualist ($p<0.001$), a necessitarian ($p<0.05$) or a natalist ($p<0.05$) compared to being a presentist. One possible interpretation is that this implicitly illustrates a necessitarian intuition of respondents. Prevention of future disability might be perceived as more important and respondents might therefore retain their choice to value the future life more often.

Table 5: Multinomial logistic regression model estimates based on the identified response profiles in choice task 1.[10]

	Pre-defined actualists (vs pre-defined presentists)			**Cluster actualists (vs cluster presentists)**			**Pre-defined necessitarians (vs pre-defined presentists)**			**Cluster necessitarians (vs cluster presentists)**		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.15	0.11 – 0.20	**<0.001**	0.19	0.14 – 0.25	**<0.001**	0.11	0.08 – 0.17	**<0.001**	0.20	0.15 – 0.28	**<0.001**
Gender												
Female	1			1								
Male	1.84	1.47 – 2.31	**<0.001**	1.86	1.45 – 2.39	**<0.001**	1.15	0.83 – 1.58	0.407	1.10	0.83 – 1.45	0.520
Religious												
No	1			1								
Yes	1.45	1.15 – 1.82	**0.002**	1.37	1.07 – 1.76	**0.014**	0.98	0.70 – 1.36	0.901	0.83	0.62 – 1.11	0.217
Country												
BE	1			1								
NL	1.53	1.21 – 1.92	**<0.001**	1.41	1.10 – 1.81	**0.007**	1.19	0.86 – 1.65	0.281	1.10	0.83 – 1.45	0.515
Children/ childwish												
Kids and/or childwish	1			1								
No kids + I don't know / children are not an option due to age	0.46	0.30 – 0.72	**0.001**	0.56	0.35 – 0.89	**0.014**	0.53	0.28 – 1.01	0.053	0.58	0.34 – 0.98	**0.044**
No kids + no childwish	0.49	0.33 – 0.71	**<0.001**	0.58	0.38 – 0.87	**0.009**	1.16	0.76 – 1.77	0.501	1.03	0.70 – 1.53	0.867
Order effect												
Rinutek first	1			1								
Jerutek first	1.62	1.29 – 2.04	**<0.001**	1.62	1.26 – 2.08	**<0.001**	1.10	0.80 – 1.52	0.544	1.39	1.05 – 1.84	**0.022**

[10] We cannot report an R^2 because of the nature of the model (with the dependent variable being the respondent's categories). Our sample size is considered to be sufficiently large as guidelines for multinomial logistic regression indicate a minimum of 10 observations per independent variables. (31)

Predictors	**Pre-defined necessitarians (vs pre-defined actualists)**			**Cluster natalists (vs cluster presentists)**			**Cluster antinatalists (vs cluster presentists)**		
	Odds Ratios	*CI*	*p*	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	0.77	0.50 – 1.18	0.224	0.13	0.09 – 0.19	**<0.001**	0.02	0.01 – 0.05	**<0.001**
Gender									
Female									
Male	0.62	0.43 – 0.89	**0.010**	1.06	0.76 – 1.48	0.745	1.90	1.05 – 3.46	**0.035**
Religious									
No									
Yes	0.68	0.47 – 0.98	**0.036**	1.00	0.72 – 1.41	0.980	0.83	0.45 – 1.54	0.561
Country									
BE									
NL	0.78	0.54 – 1.12	0.183	0.90	0.64 – 1.26	0.541	1.67	0.91 – 3.07	0.097
Children/ childwish									
Kids and/or childwish									
No kids + I don't know if I want children/ having children is not an option due to my age	1.15	0.55 – 2.40	0.719	0.94	0.54 – 1.61	0.812	0.65	0.19 – 2.16	0.478
No kids + no childwish	2.38	1.40 – 4.02	**0.001**	1.12	0.70 – 1.79	0.642	2.01	1.02 – 3.97	**0.043**
Order effect									
Rinutek first									
Jerutek first	0.68	0.47– 0.98	**0.036**	1.47	1.05 – 2.04	**0.024**	0.96	0.53 – 1.73	0.883

3.4 Discussion

The objective of this study was to explore how laypeople make choices in a healthcare priority-setting context in which not only existing patients, but also future lives are affected. We study if people take into account the consequences of healthcare on future lives, and whether this depends on the development stage of the future life or the type of intervention effect. Based on these outcomes, we study how the public's perspective corresponds to the ethical views explored in the theoretical literature.

Overall, our results suggest that the majority of people do not take into account future life effects when setting healthcare priorities. Approximately 70% of respondents was indifferent between both female patients when the future offspring was not yet conceived. When the woman with future offspring was already pregnant or just gave birth, 40% of respondents remained indifferent between both patients. The type of intervention effect (affecting conception/birth rates versus the QoL of future lives) only had a limited impact on choice behaviour.

We allocated respondents to pre-defined and cluster based choice profiles, which overlapped for 80%. Roughly 60% of respondents was generally indifferent between both women and is therefore considered a presentist, only attributing value to present people. People's tendency to ignore the affected future offspring however contrasts to previous findings in related areas (and our introductory questions) indicating that people do value future generations (albeit at a discount rate) and are willing to make sacrifices for them.(10–17) Our results therefore seem to support the finding that people's value judgements about *outcomes* per se might differ from their judgement about the specific *actions* that are required to obtain these outcomes.(19,20) For example, most people might agree that saving a woman and a future child is a better outcome than only saving a woman. These same people might however not agree

with the action that this outcome requires: refusing another woman access to treatment, simply because she will not have children in the future.
20% of respondents was identified as an actualist because they typically prioritized the woman with future offspring and hence valued the affected future life in every situation. This group was found to be the most consistent in their choices and significantly differed from the presentist respondents in terms of gender, religion, nationality and the fact that they have children or a childwish. The last 20% of respondents made less consistent choices that were more difficult to interpret. While we did identify 9% of the sample as necessitarians based on our expected choice profiles, the cluster analysis shows that a similar percentage of respondents adopted an opposite 'natalist' perspective.

The main strength of this study is that, for the first time, prioritization behaviour was studied in a context that distinguishes interventions affecting the coming into existence versus the QoL of future lives, during different stages before and after conception. This unique context allowed us to explore public support for different views proposed in the population ethics literature. Complex ethical dilemmas were reduced to simplified choice questions in which only one dimension related to FLEs was varied per question. The obligation to justify each choice in a comment box allowed us to apply strict quality criteria and provided meaningful insights in respondents' underlying motives.

However, we are aware that our survey comes with limitations as well. Because of the complex matter addressed in this study, preliminary tests showed that person trade-offs were very abstract and too difficult to interpret. Therefore, we decided not to estimate a marginal rate of substitution. As a consequence however, we cannot make any claims about the degree to which future lives were valued (or discounted) compared to presently existing people, or compared to future lives in different situations.

Second, it must be kept in mind that our results were obtained considering one specific context. In order to keep questions simple and transparent, we presented a choice between two female patients in an *absolute* scarcity context (the number of treatments was limited). Respondents might therefore focus more on the concrete actions and consequences related to these 'bedside' decisions. HTA however typically aims to inform *macro* budget allocation decisions concerning a *relative* scarcity context in which the reimbursement of interventions is limited. Presenting the same type of choices in a different context (e.g. choosing between the reimbursement of an intervention for two groups of 100.000 patients) might have resulted in different outcomes.
Third, to ease interpretation and limit the number of choice options we only included the choice 'I think both are equally important and I cannot choose'. This formulation might have influenced respondents' comments and did not allow other interpretations of choice refusal such as incommensurability of the alternative outcomes presented.(19)
Furthermore, the topic and choices presented to respondents still required substantial commitment and the willingness to face ethically challenging dilemmas. Even though we applied very strict selection criteria, there is still a significant risk that questions were misinterpreted. Nevertheless, 91% of respondents did indicate that the questions were clear, but 63% found them difficult to answer.
Last, even though the study was conducted in a substantial, representative sample, online panel membership might be associated with particular characteristics (e.g. access to internet, willingness to participate in research), possibly affecting representativeness.

Our results create a platform for future research. This includes studying the robustness of our findings in a context that considers reimbursement decisions on a population level and exploring in more detail the boundary at which people consider a life to become present. Further research could explore a marginal rate of substitution between future and present people (and their QALYs), as well as between potential and necessary future

lives. Last, a context in which health interventions have an (unintended or indirect) negative effect on future people's QoL or future fertility rates could be studied, as people might experience stronger moral obligations to prevent future suffering.

References

1. Luyten J, Verbeke E, Schokkaert E. To be or not to be: Future lives in economic evaluation. Health Economics. 2022;31(1):258–65.
2. Verbeke E, Luyten J. Future Offspring Costs in Economic Evaluation. PharmacoEconomics. 2022 Feb 1;40(2):141–7.
3. Goldhaber-Fiebert JD, Brandeau ML. Evaluating Cost-Effectiveness of Interventions that Affect Fertility and Childbearing: How Health Effects are Measured Matters. Med Decis Making. 2015 Oct;35(7):818–46.
4. Arrhenius G. Population Ethics. The Challenge of Future Generations. (Unpublished version May 2016).
5. Greaves H. Population Axiology. Philosophy Compass. 2017;12(11):e12442.
6. Daniels N, Sabin JE. Setting limits fairly. 2nd ed. Oxford University Press; 2008.
7. Powers M, Faden R. Social Justice: The Moral Foundations of Public Health and Health Policy. 1st edition. Oxford New York: Oxford University Press; 2006. 248 p.
8. Gaertner W, Schokkaert E. Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice [Internet]. Cambridge Core. 2011 [cited 2022 Mar 2]. Available from: http://www.cambridge.org/core/books/empirical-social-choice/E8B35622AA5DA2E91EA410A46A3DF751
9. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
10. Johannesson M. The willingness to pay for health changes, the human-capital approach and the external costs. Health Policy. 1996 Jun 1;36(3):231–44.
11. Cairns J. Discounting in economic evaluation. In: Economic Evaluation in Health Care. Oxford: Oxford University Press; 2006.
12. Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs: how the public discounts time and age. J Risk Uncertainty. 1994 May 1;8(3):243–65.
13. Graham AL, Papandonatos GD, Bock BC, Cobb NK, Baskin-Sommers A, Niaura R, et al. Internet- vs. telephone-administered questionnaires in a randomized trial of smoking cessation. Nicotine Tob Res. 2006 Dec;8 Suppl 1(Suppl 1):S49-57.
14. Scarborough H, Bennett J. Estimating intergenerational distribution preferences. Ecological Economics. 2008 Jul;66(4):575–83.

15. Johansson-Stenman O, Martinsson P. Are some lives more valuable? An ethical preferences approach. Journal of Health Economics. 2008 May 1;27(3):739–52.
16. Frederick S. Measuring Intergenerational Time Preference: Are Future Lives Valued Less? Journal of Risk and Uncertainty. 2003 Jan 1;26(1):39–53.
17. Fairbrother M, Arrhenius G, Bykvist K, Campbell T. How Much Do We Value Future Generations? Climate Change, Debt, and Attitudes towards Policies for Improving Future Lives [Internet]. The Institute for Future Studies; 2022. Available from: https://osf.io/preprints/socarxiv/kd2fp/
18. Spears D. Making people happy or making happy people? Questionnaire-experimental studies of population ethics and policy. Soc Choice Welf. 2017 Jun;49(1):145–69.
19. Caviola L, Althaus D, Mogensen AL, Goodwin GP. Population ethical intuitions. Cognition. 2022 Jan;218:104941.
20. Schoenegger P, Grodeck B. Concrete over abstract: Experimental Evidence of Reflective Equilibrium in Population Ethics (Forthcoming). In: Behavioral Ethics and Experimental Philosophy. Springer;
21. Jamison JC. Perceptions Regarding the Value of Life Before and After Birth. Reproductive System & Sexual Disorders [Internet]. 2016 [cited 2019 Jan 9];05(04). Available from: https://www.omicsonline.org/open-access/perceptions-regarding-the-value-of-life-before-and-after-birth-2161-038X-1000195.php?aid=82198
22. Dolan P, Tsuchiya A. The person trade-off method and the transitivity principle: an example from preferences over age weighting. Health Economics. 2003;12(6):505–10.
23. Lewis PA, Charny M. Which of two individuals do you treat when only their ages are different and you can't treat both? J Med Ethics. 1989 Mar;15(1):28–34.
24. Heyd D. Genethics: Moral Issues in the Creation of People. University of California Press; 1992.
25. Singer P. Practical Ethics. Cambridge University Press; 2011. 353 p.
26. Frederick S. Measuring Intergenerational Time Preference: Are Future Lives Valued Less? Journal of Risk and Uncertainty. 2003 Jan 1;26(1):39–53.
27. Statbel. Statbel, het Belgische statistiekbureau | Statbel [Internet]. [cited 2021 Oct 27]. Available from: https://statbel.fgov.be/nl
28. Statistiek Vlaanderen. Vlaanderen in cijfers 2021 [Internet]. Available from: https://publicaties.vlaanderen.be/view-file/44796

29. CBS. StatLine - Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari [Internet]. [cited 2021 Oct 24]. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461bev/table?ts=1635070359651
30. Hosmer D, Lemeshow S. Applied Logistic Regression. John Wiley & Sons; 2004.

Appendix: Sample characteristics relative to those of the Flemish and Dutch population. Source: Statbel and Statistiek Vlaanderen for Flemish and Belgian data (31,32), Statline for Dutch data (33).

CHARACTERISTIC		FL (BE) (n=934)	General population FL	NL (n=930)	General population NL
		%		%	
GENDER	Male	48%	49.5%	46%	49%
	Female	52%	50.5%	54%	51%
AGE	18-24y	10%	7%	8%	11%
	25-34y	13%	16%	14%	16%
	35-44y	16%	16%	18%	15%
	45-64y	37%	35%	38%	34%
	65+y	23%	26%	22%	24%
EDUCATION[11]	None or primary	4%	18%	2%	9%
	Secondary	47%	40%	52%	56%
	Higher education	49%	42%	46%	35%
HOW OFTEN DO YOU HAVE PROBLEMS PAYING FOR MONTHLY HOUSEHOLD EXPENSES?					
	Never	58%		60%	
	Once a year	17%		19%	
	Once in 3 months	13%		12%	
	Every month	12%	12%	9%	12%
MARITAL STATUS					
	Single	27%	23%	28%	37%
	Cohabiting / married	60%	59%	62%	47%
	Divorced	10%	11%	7%	10%
	Widow(er)	3%	7%	3%	6%
RELIGIOUS	Yes	49%	52%	39%	47%
	No	51%	48%	61%	53%
DO YOU HAVE CHILDREN?					

[11] National percentages are reported for the age group 15-64y in FL and 15-75y in NL.

	No kids + no childwish[12]	11%	6%	16%	8%
	Kids and/or childwish	79%		74%	
	No kids + I don't know if I want children/ having children is not an option due to my age	10%		10%	
EMPLOYMENT	Active on the labour market	51%	50%	51%	60%
	Education sector	*1%*	*3%*	*4%*	*4%*
	Healthcare sector	*1%*	*5%*	*7%*	*12%*
	Retired	27%	23%	19%	22%
	Housewife/husband	4%		7%	2%
	Disabled	6%	5%	4%	5%
	Student	7%	7%	6%	9%
	Unemployed	3%	4%	6%	2%
	Other	2%		7%	

[12] National percentages represent the proportion of voluntarily childless people. We are not able to draw specific conclusions of this comparison as it is possible that a respondents' child wish evolves over time.

Chapter 4: Long-term CUA of NIPT in Belgium
Screening alternative perspectives on neonatal QALYs

Abstract

Background: The cost-effectiveness of prenatal genetic screening programs is typically calculated in terms of a cost per case detected or per affected birth averted. Cost-utility analyses in which a cost per QALY is reported would allow the adoption of a broader evaluation perspective as well as increased uniformity in study outcomes. The use of QALYs is however typically avoided in this context, as it remains unclear if, and how neonatal QALYs should be included in an evaluation.

Objective: Our aim is to conduct a cost-utility analysis of non-invasive prenatal testing (NIPT) in which both long-term maternal and neonatal QALYs are included. We study how cost-effectiveness varies according to the ethical perspective adopted for the inclusion of neonatal QALYs.

Methods: We extend the time horizon of a recent study that evaluated NIPT screening for Down Syndrome in Belgium, by developing a Markov model that follows the mother and future child throughout the rest of their lives. Cost-effectiveness is calculated according to four plausible views on the value of the unborn children: presentism, actualism, placeholder necessitarianism and asymmetric actualism.

Results: We find that first line NIPT screening generates a positive Net Monetary Benefit (compared to a no screening scenario) at a willingness to pay threshold of €50.000/QALY according to every perspective. An

actualist view, according to which all unborn children should be taken into account, generates the lowest benefit because substantial neonatal QALY losses related to pregnancy termination are included. NIPT generates the most value-for-money according to an asymmetric actualist perspective that only includes neonatal QALYs related to the suffering avoided in future children.

Conclusion: We show that there are different ways to consider neonatal costs and QALYs in a consistent manner, based on different plausible views on the value of future lives, and that these lead to largely diverging cost-effectiveness outcomes. Even though long-term cost-utility analyses require data that are currently lacking, only a consensual view on the value of unborn, future lives will enable the generation of robust cost-effectiveness evidence.

This chapter will be submitted as:

Verbeke, E., Schokkaert, E., Luyten, J., Long-term cost-utility analysis of NIPT in Belgium: screening alternative perspectives on neonatal QALYs.

4.1 Introduction

Economic evaluations of prenatal genetic screening programs typically calculate cost-effectiveness in terms of a cost per case of chromosomal abnormality detected, or per affected birth averted. As a consequence, these studies fail to incorporate long-term costs and health outcomes, nor do they allow the evaluation of allocative efficiency (in which cost-effectiveness of genetic screening is compared to other health domains).(1,2) In response, a limited number of more recent studies has adopted a broader time horizon, including lifetime costs related to the chromosomal abnormality (3–7), as well as maternal QALYs.(8–11) In theory, QALYs are considered the preferred outcome measure to evaluate prenatal screening programs as they allow the consideration of different health dimensions for different individuals affected by the intervention.(12) However, the methodology used in these long-term studies is found to be highly inconsistent. While lifetime savings of preventing the birth of an affected child are typically included, these studies systematically omit any forgone benefit related to the unborn child in terms of both its costs and QALYs.(13–16) A recent systematic review of economic evaluations of non-invasive prenatal testing (NIPT) therefore concludes that no study currently satisfies the condition of including 'all major short-term and long-term maternal, foetal, as well as negative outcomes and costs'.(1)

The reason why this methodological inconsistency arises is that the inclusion of neonatal QALYs (or other types of long-term benefits) implies a concrete value judgement of all available options after screening. Specifically the option of pregnancy termination after a positive screening result raises fundamental questions to which current guidelines do not provide any answers. Including the neonatal QALYs lost after pregnancy termination has been argued to disadvantage screening and to compromise parental autonomy.(10,17,18) Furthermore, this would imply that more accurate screening tools generate higher QALY losses, which also raises the question whether

future replacement children should be included as well.(1,10,17) Because of the lack of answers to these questions, short term evaluations that calculate a cost per natural unit (such as the number of cases detected) have been preferred in the past. As, in these studies, costs and outcomes are at least considered in a symmetric manner.(12)

These morally uncomfortable questions do however not justify an exclusion of all long-term costs and health outcomes, nor do they provide a reason to only include potential future savings while ignoring all other consequences. Furthermore, what has previously been overlooked in this discussion, is that these problematic implications arise under the assumption that *all* QALYs of *every* unborn child should be included in an evaluation. The future QALYs of these future children can however be considered in different manners, according to alternative ethical perspectives on the value of unborn, future lives.

Therefore, the goal of this study is to conduct an explorative long-term cost-utility analysis of non-invasive prenatal testing (NIPT) that includes both maternal and neonatal QALYs according to different plausible perspectives. We extend the time horizon of a previous Belgian study and explore how each perspective gives rise to diverging cost-effectiveness results. The goal of our analysis is not to enter into the debate of which of the ethical perspectives that we discussed in the previous chapters should be preferred. Rather, it is to illustrate the significant variation in outcomes that arises because of these different perspectives. As a result, we show the importance of a consensual and consistent view with regard to the status of future lives in the evaluation of prenatal genetic testing, but also more generally, in all evaluations where healthcare affects both presently existing and future lives.

4.2 Methods

4.2.1 Initial model

A cost consequence model that was developed in 2014 to assess the introduction of first or second line NIPT for the detection of T21 (Down syndrome) in Belgium was used as the starting point of our analysis. The population considered in this model consisted of all pregnant women in the Belgian population, excluding twin pregnancies. Input variables and costs were based on Belgian data, supplemented by data from the literature. Considered costs included direct healthcare costs from a healthcare payer perspective, payments out of the public healthcare budget and patients' co-payments. The study considered a short-term perspective, including costs and effects before birth. First line NIPT was identified as the preferred scenario as it reduces procedure related-miscarriages and misses fewer cases of T21. However, it was considered to be only realisable at a substantially lower price of NIPT. Further details of the design, conduct and results of the original model are described in the related paper and scientific report.(19,20)

4.2.2 Extended model

We develop a Markov model in Excel that builds upon the initial decision-tree model to calculate the long-term cost-effectiveness of first line NIPT screening compared to a no-screening scenario. Even though second line screening or an invasive testing scenario would be a more realistic comparator, we chose a no screening scenario to keep the model and result interpretation transparent, as we make our calculations for methodological purposes only. We consider the same cohort size of 129.199 Belgian pregnant women as in the initial study, and assume that these women are 31 years old with a life expectancy of 83 years (based on Belgian averages). As shown in Figure 1, the model departs from four possible screening-related birth outcomes that are based on the calculated probabilities in the initial study and follows the screened mother and potential child throughout their lives. In addition, we assume

that 75% of women who experienced a procedure related miscarriage or terminated pregnancy, will have a replacement child within 2 years.(8) The no-screening scenario does not include replacement children, as lost pregnancies are in this scenario only caused by spontaneous miscarriages. (Table 1)

Figure 1: The long-term cost-utility model departs from the probabilities calculated in the initial study and considers the mother and screened children throughout their lives.

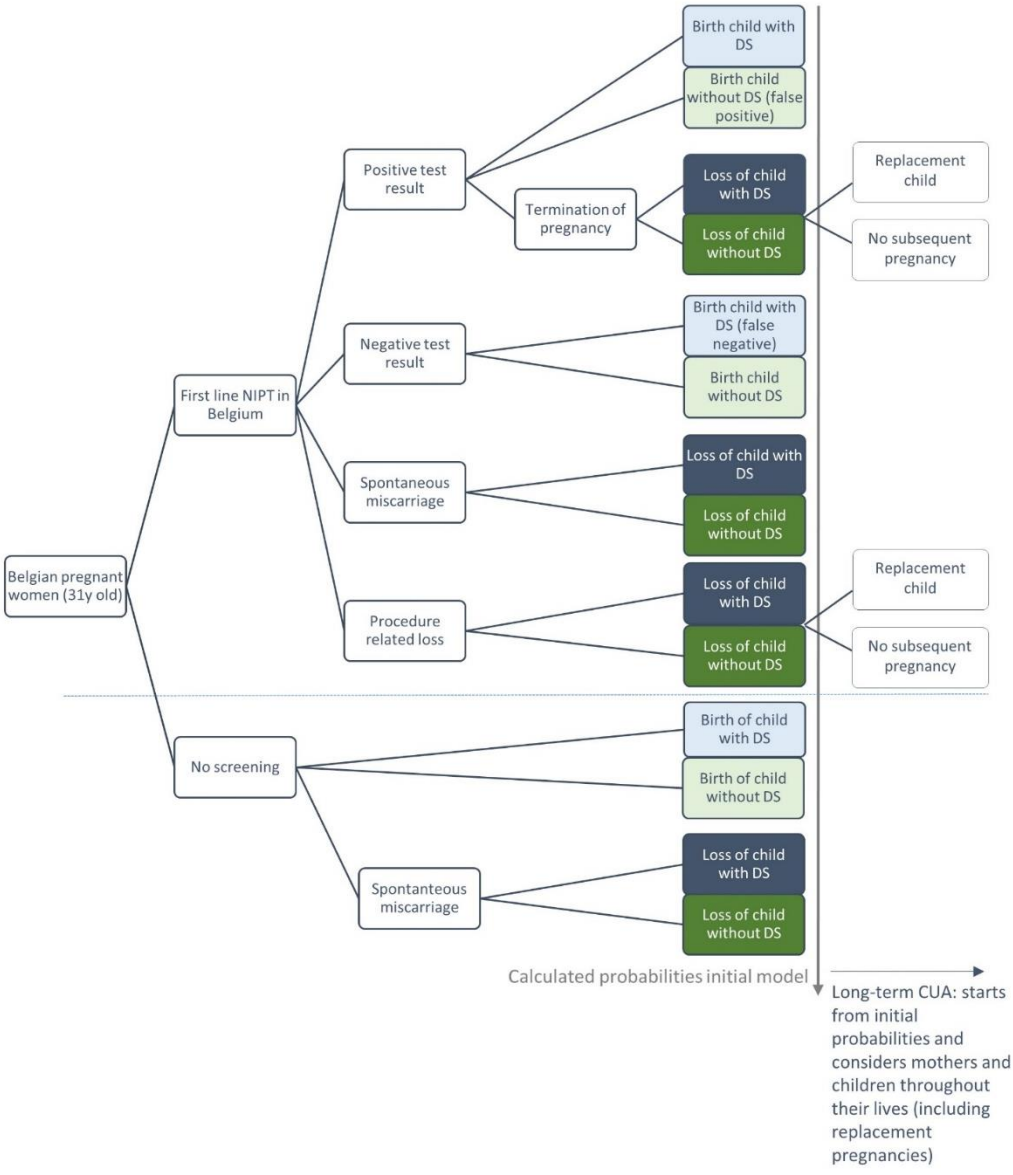


Table 1: Input variables retrieved from initial model: mean probabilities

Variable	**Mean (%)** (20)
Probabilities of input variables	
Screening uptake	78,87
NIPT sensitivity	99,3
NIPT specificity	99,84
No result rate first test (week 12)	4,00%
No result rate first test (week 13)	2%
Probability of having an invasive test (after positive NIPT)	87,50%
Invasive testing sensitivity	100%
Invasive testing specificity	100%
Procedure related fetal loss after invasive test	1,00%
Pregnancy termination after T21 diagnosis	95,45%
Miscarriage rate overall (after week 14)	1,50%
miscarriage rate non T-21 after w14	1,44%
T21 miscarriage rate after w14	25,00%
*Calculated probabilities of outcomes **after first line NIPT** (100%= the initial population of pregnant women)*	
Women giving birth	94,861%
Birth of child with Down syndrome (DS)	0,050%
True positive test result	0,049%
False negative test result	0,001%
Birth of child without DS	94,811%
True negative test result	94,787%
False positive test result	0,024%
Procedure related fetal losses	0,020%
T21 loss	0,006%
Non-T21 fetal loss	0,014%
Spontaneous miscarriages	4,960%
T21 miscarriages (week 10)	0,095%
non-T21 miscarriages	4,905%
Pregnancy terminations	0,159%
Probabilities of outcomes in a no-screening scenario (100%= the initial population of pregnant women)	
Women giving birth	95,000%
Birth of child with DS	0,170%
Birth of child without DS	94,830%
Spontaneous miscarriages	5,000%

T21 miscarriages (week 10)	0,095%
non-T21 miscarriages	4,905%
Probabilities included in extended model	
Pregnancy replacement rate after pregnancy termination or procedure related loss	75%

4.2.3 Costs

Short-term costs related to the screening pathway are taken from the initial model and are based on data from the National Institute for Health and Disability Insurance and expert opinions, an updated price for NIPT is included as this decreased significantly in recent years. With regard to long-term costs, we include cost categories that are similar to the costs included in the existing long-term studies (Table 2).(3–8,10,11) A healthcare payer perspective was adopted (following the initial model) and costs were discounted at 3%.

Table 2: Input variables (costs)

Variable	Mean	Source
Screening related costs		
Price NIPT	€ 260	(21)
Invasive test	€ 934	(20)
Hospitalization for leakage	€ 3.515	(20)
Pregnancy termination	€ 914	(20)
Long term costs		
Discounted lifetime hospitalization costs of child with DS	€ 50.000	(20)
Yearly institutionalisation costs of adult person with DS (publically financed)	€ 12.000 (age 18-55y = average life expectancy)	(20)
Assisted Reproductive Technology cost per replacement pregnancy	€ 1730	(20)

4.2.4 Utilities

Utility values, reflecting the health-related Quality of Life (QoL) experienced by the mother and child, are retrieved from the literature (Table 3). Maternal utilities related to testing outcomes are applied for 1 year, assuming that quality of life is only affected during pregnancy and the first months after birth.(8) Because no recent data were available related to false negative outcomes, we assumed a utility value equal to the long-term utility experienced by mothers of a child with Down Syndrome (DS) (0,48). For the proportion of women who have a replacement child within 2 years, we assume a utility of 0,880 during the second year after testing. Utilities related to the long-term outcomes after testing, such as being a parent of a child with DS or being a child with DS, are applied for the remainder of the mother or child's expected lifespan (based on average life expectancies: 83y for a Belgian woman, 80y for an unaffected child and 60y for a child with Down).(8) Calculated QALY values are discounted at 1,5%, as recommended by Belgian guidelines.(22)

Table 3: Input variables (utilities)

Variable	Mean	Duration	Source
Utilities mother			
Negative test result (true negative)	1	1y	(own assumption-NIPT is considered to reassure parents)
Unexpected birth with DS (false negative)	0,48	1y	(own assumption-equal to utility of parent of child with DS)
Positive test result + continuation of pregnancy (true positive & false positive)	0,655	1y	(23)
Positive test result + pregnancy termination	0,771	1y	(23)
Spontaneous miscarriage	0,59	1y	(23) (pregnancy loss)

Procedure related miscarriage	0,59	1y	(23) (pregnancy loss)
Pregnancy loss + replacement child within 2 years	0,880	1y (after 1y)	(23)
Postpregnancy: child with DS	0,480	52y	(23)
Postpregnancy: child without DS	1	52y	(reference QoL level)
Utilities child			
Child with DS	0,55	60y	(24)
Child without DS	1	80y	(reference QoL level)

4.2.5 Ethical perspectives

Prenatal genetic screening is not the only intervention that raises ethical questions about the moral status of future lives in economic evaluation. In the field of population ethics, these questions have been extensively studied and alternative (mutually-exclusive) perspectives have been elaborated.(25,26) We select four of these perspectives, each offering a different plausible view on the value generated by prenatal genetic screening programs (Table 4).

The limited number of economic evaluations that have adopted a long-term perspective, have typically only included future savings related to cases prevented, while neonatal QALYs have been ignored. This approach, which we consider the 'common practice' for long-term cost-utility analyses, seems therefore inconsistent as costs and effects are normally included simultaneously for the same parties, over the same time horizon. We start by considering two obvious alternative approaches.

First, it can be argued that prenatal genetic screening should be evaluated solely based on how the program affects the screened parents. As the main goal of genetic screening is to inform parents and give them the opportunity to terminate pregnancy in case of a positive test result. In

this respect, a *presentist* view can be adopted according to which only presently existing people have value and all neonatal costs and QALYs should be ignored. The downside of this view is that the substantial long-term costs related to a genetic disorder would be ignored, as the cost-effectiveness of NIPT will solely be based on the comparison of the screening costs versus the generated benefits for the parents. In a broader (healthcare) context, this perspective would require a complete indifference with regard to the wellbeing of future people in general. Investments in disease eradication or even climate change would therefore become difficult to justify even though many subsequent future generations could benefit from these programs.
In contrast, it can be argued that unborn, future people's QALYs should be treated in the same manner as present people's QALYs, as all lives (and their QALYs) are considered to have equal value in HTA. An *actualist* view can therefore be defended, according to which the QALYs of all children that could be born after screening or in the no screening scenario, should be taken into account in the same, impartial manner. Therefore, if less children are born after screening, this loss of future children ('s QALYs) should be taken into account as well. As previously discussed, this consequence is the reason why the inclusion of neonatal QALYs has been rejected in the past. As we show in our model, the included QALY losses can however also be compensated by replacement children without a genetic disorder, who will typically experience 'more' QALYs than the initial children that are lost. Nevertheless, in a general HTA context an actualist view remains difficult to defend, as it gives rise to a disproportionate impact of future lives on cost-effectiveness outcomes. Disease eradication programs or fertility treatments would generate lifetime amounts of QALYs, becoming extremely attractive investments.(25,26)

We also adopt two alternative views that offer a compromise between the systematic in- or exclusion of future lives. A first alternative approach is to evaluate genetic screening in terms of the QoL gains it

generates for future children by preventing genetic disorders. As discussed in the previous chapters, a necessitarian view allows to take into account the wellbeing (QoL) of the future people that will exist, without having to include all potential future people whose existence depends on the intervention. In the context of genetic screening it can similarly be argued that the number of pregnancy terminations should not matter, rather, the quality of life of the children that *will* be born, matters. Interpreting necessitarianism in a narrow sense would however not provide additional insights, as the children whose genetic disorder is 'avoided' by pregnancy termination are actually potential lives that should not be included in the evaluation. This is why we adopt a broader interpretation of necessitarianism and consider 'necessary placeholders' in society. This *necessitarian placeholder view* only takes into account the future lives that have a necessary 'place' in future society. Because of this necessary place, an intervention can only affect their quality of life, but not their existence. (25–29) For example, parents might decide to have 'a third child', independent of the genetic screening outcome. In case of pregnancy termination after a positive test result, they will still have a 'different' third child in the near future, but now possibly without genetic disorder. The *placeholder* interpretation therefore does not distinguish between e.g. a child 'Anna' that could be born today or a child Rosa that could be born tomorrow. This 'third child' (independent of its identity) is hence a necessary placeholder (as its existence does not depend on the screening program) whose QoL has improved because of the screening. In our study, this translates into the inclusion of neonatal costs and QALYs related to the children that are born (with or without DS) and the children that are lost but also replaced. The children who are lost but not replaced are considered merely *potential*, the screening program only affects their existence but not their QoL and they should therefore not be taken into account.

One could however question whether every possible improvement in future children's QoL is worth investing in, taking into account the opportunity costs of these investments that are borne by presently

existing people in need of healthcare. Therefore, a last option is to discriminate future lives, not based on their potentiality, but based on their QoL starting level, so that only the unborn children whose future suffering is avoided by genetic screening are taken into account. According to an *asymmetric actualist* view, *all*[13] future lives in good health have *neutral value* (there utility does not make an outcome better nor worse) but future lives in poor health have negative value. This negative value reflects future people's suffering that we, as present people, should try to avoid.(25,26,30,31) In our model, a child with DS is considered to have a utility value equal to 0,55. As a utility of 1 equals perfect health, considering every life with a utility below 1 to experience 'suffering' seems an extreme assumption. We therefore make the (arbitrary) assumption that lives with a utility below 0,6 suffer. Of course, as we later discuss in this paper, this boundary could also be lower which would imply that a child with DS has neutral value. In the base case however, the birth of a child with DS is hence included as a negative value equal to -0,05 (=0,6-0,55). In contrast, the loss or termination of every child with DS is included as a positive QALY value of 0.05, because this future suffering is avoided. Children without DS are not included in terms of QALYs, because they do not suffer and hence have a neutral utility value.

[13] An asymmetric approach can also be combined with other views. An asymmetric necessitarian would for example take into account 'necessary placeholders' as well as potential lives in poor health.

Table 4: Alternative ethical perspectives that can be adopted for the valuation of neonatal QALYs and their practical implications for the economic evaluation of NIPT.

Ethical perspective	The value of prenatal genetic screening is based on	Implications for neonatal costs and QALYs in economic evaluation
Perspective adopted in previous studies ('common practice') Future lives are only valued in terms of the future costs they will generate. Future benefits produced, or utility/wellbeing experienced by these lives does not have any value. (1,13)	the amount of future healthcare costs avoided by preventing affected births	Only the future costs related to a child with disability are taken into account (no other neonatal costs or QALYs are considered).
Presentism Only presently existing people matter.(25,26)	only the consequences for the parents, as the main goal of genetic screening is to inform parents and offer them the autonomy to decide whether or not to terminate pregnancy in case of a positive test result.	Neonatal costs and QALYs are systematically excluded.
Actualism All lives that can become actual in the future matter.(25,26)	how the program affects all people that exist or can exist in the future, because every life (and their QALYs) has equal value.	Neonatal costs and QALYs are systematically included when a child is born, lost or replaced.

Placeholder necessitarianism Only the lives that *will* exist in the future have value. We interpret this view from a broad perspective, by valuing the lives whose 'place' in future society does not depend on the intervention. (29)	the degree to which the QoL of the future 'newborn population' is improved by preventing affected births.	Neonatal costs and QALYs are only included when a child is born, or when a child is lost but replaced by another one.
Asymmetric actualism The coming into existence of lives in good health has neutral value (it does not make an outcome better or worse) but the coming into existence of lives in poor health has negative value (because we have the obligation to avoid future suffering).(25,31)	the degree to which future suffering is avoided.	QALYs gains are calculated asymmetrically: the birth of a child with DS has a negative value (because the child experiences 0,05 less utility than a life without suffering). The loss or termination of a child with DS has a positive QALY value, as a utility decrement of 0,05 is avoided. Children without DS are not considered in the evaluation, because they have neutral value. (We here assume that a life with utility below 0,6 suffers.)

Table 5 provides an overview of the key-variables that each perspective includes in an economic evaluation. We calculate the cost-effectiveness of NIPT screening in terms of a net monetary benefit[14] (NMB) that is based on an assumed threshold of €50.000/QALY. We report NMBs

[14] The NMB is calculated as: (incremental benefit [QALYs] x threshold [€/QALY]) – (incremental cost [€]).

instead of incremental cost-effectiveness ratios, as the NMB allows an unambiguous interpretation of positive results (in case of cost-effectiveness) and negative results (in case the intervention is not cost-effective). We start with a base case analysis in which we compare the long-term cost-effectiveness of NIPT screening according to every perspective. Subsequently, we explore different scenarios in which one (or more) of the key variables in the model is altered and we study how this affects the initial NMB. Based on these results, we identify the key-drivers of cost-effectiveness for each perspective.

Table 5: Variables that are included in the evaluation according to each perspective.

Included variables	Common practice	Presentism	Actualism	Place-holder necessi-tarianism	Asym-metric actualism
QALYs mother (& father)	✓	✓	✓	✓	✓
QALYS child without DS					
Born			✓	✓	
Lost			✓	(✓)*(replacement rate)	
Replaced			✓	✓	
QALYs child with DS					
Born			✓	✓	✓ (-)
Lost			✓	(✓) (if replaced)	✓ (+)
Replaced			✓	✓	
Costs mother (screening)	✓	✓	✓	✓	✓
Costs child (future healthcare child with DS)	✓		✓	✓	✓

4.3 Results

4.3.1 Base case scenario

Extending the initial model to a time horizon that considers the mother and child throughout the rest of their lives, we find that first line NIPT screening in Belgium generates a net monetary benefit of more than €318 million (table 6). This result is based on the 'common practice' approach in which only the long-term healthcare savings of averted births with DS are included, but no other potential forgone benefits related to the unborn child are taken into account.

Table 6: Costs, effects and cost-effectiveness in terms of net monetary benefit (NMB) according to each perspective (threshold NMB= €50.000/QALY).

Perspective	Incremental costs	Incremental QALYs	NMB	% difference with common practice	Undiscounted NMB
Common practice	-€ 81.770.274	4728	€ 318.151.639		€564.347.609
Presentism	€ 488.475	4728	€ 235.892.891	-35%	€345.535.372
Actualism	-€ 81.770.274	2635	€ 213.504.335	-49%	€474.099.775
Place holder necessitarianism	-€ 69.829.066	4011	€ 270.380.501	-18%	€547.482.756
Asymmetric actualism	-€ 81.770.274	5459	€ 354.725.930	+10%	€620.426.910

The first alternative way to evaluate NIPT is to also exclude the costs related to the child and calculate cost-effectiveness solely based on the long-term consequences for the mother (adopting a presentist perspective). Because the long-term savings related to the avoided births with DS are no longer included in the model, the NMB decreases with 35%.

Alternatively, both long-term costs *and* QALYs related to the child can be taken into account according to an actualist perspective. This approach has previously been avoided, as it requires the inclusion of the lifetime amounts of neonatal QALYs lost due to pregnancy termination. We indeed find that the inclusion of these neonatal QALY losses has a larger impact on the NMB compared to the exclusion of the future (neonatal) savings in the presentist approach. Even though the included QALY losses are partly compensated by the replacement children, this perspective leads to the lowest NMB (-49% compared to the common practice outcome).

Third, a necessitarian view only includes the QALY losses that are compensated by (healthy) replacement children's QALYs, but not those of the unreplaced pregnancy terminations. The negative impact of including neonatal QALYs on the NMB is therefore smaller (-18% compared to the common practice result).

Last, according to an asymmetric actualist perspective, only the QALYs related to the suffering of children with DS should be included. While the birth of a child without DS has a neutral utility value, the birth of child with DS is considered a QALY loss (and consequently the loss of a child with DS is considered a QALY gain). First line NIPT screening therefore generates a larger (neonatal) QALY gain compared to the other perspectives, resulting in the smallest NML (+10% compared to the common approach).

4.3.2 Scenario analysis: which factors drive cost-effectiveness?

Because each perspective takes into account different costs and QALYs related to NIPT, changes in the model's variables will have different effects on cost-effectiveness. We illustrate these dynamics by studying the impact of different (extreme) scenarios on the net monetary benefit as shown in Table 7 (a detailed table including costs, effects and the NMBs can be found in appendix 1). Table 8 reports for each perspective the critical value of the analysed variables that determines whether NIPT leads to a net monetary benefit or a net monetary loss. (We discuss our findings on the subsequent pages.)

Table 7: Scenario analysis: we report the absolute difference in NMB according to each perspective, expressed in millions of Euros.

	Common practice	**Presentism**	**Actualism**	**Placeholder necessitarianism**	**Asymmetric actualism**
Base case NMB	€318,2	€235,9	€213,5	€270,4	€354,7
UTILITY PARENTS					
SHORT TERM UTILITY: experienced during first year after screening					
Utility of a negative test result= 0.931 (base case= 1)	-€ 416,2	- €416,2	- €416,2	- €416,2	- €416,2
Giving birth to a child with DS, if DS was expected=1, if DS was unexpected=0	+ €1,0	+ €1,0	+ €1,0	+ €1,0	+ €1,0
LONG TERM UTILITY: experienced during the rest of their life					
Utility of a mother of a child with DS = 0, utility mother of a child without DS = 1	+ €136,1	+ €136,1	+ €136,1	+ €136,1	+ €136,1
Utility father included (equal utility as mother)	+ €236,4	+ €236,4	+ €236,4	+ €236,4	+ €236,4
UTILITY OF CHILD					
Utility child with DS= 0,3	Utility child with DS= 0,3	Utility child with DS= 0,3	Utility child with DS= 0,3	Utility child with DS= 0,3	Utility child with DS= 0,3

	Common practice	Presentism	Actualism	Placeholder necessitarianism	Asymmetric actualism
Utility child with DS= utility child without DS =1	€0	€0	- €329,2	- €279,4	- €36,6
FINANCIAL CONSEQUENCES CHILD					
Net financial implication of an adult with DS compared to adult without DS: generates an additional €10.000 net benefit	- €117,4	€0	- €117,4	- €100,3	- €117,4
REPLACEMENT & TERMINATION					
Replacement rate= 0%	+ €1,3	+ €1,0	- €398,4	- €224,8	+ €1,3
Replacement rate= 100%	- €0,4	- €0,3	+ €132,8	+ €74,9	- €0,4
Termination rate= 100%	+ €12,6	+ €9,2	+ €9	+ €10,9	+ €14,1
Termination rate= 100%, replacement rate= 0%	+ €13,9	+ €10,3	- €405,7	- €223,8	+ €15,5
SENSITIVITY/SPECIFICITY					
100% sensitivity & specificity (0 false negatives or false positives	+ €6,5	+ €6,1	+ €7,5	+ €7,7	+ €6,7

Table 8 Critical values that determine when NIPT is no longer found to be cost-effective and the NMB turns negative (keeping all other variables constant). (N.a. indicates that this variable does not have any impact on cost-effectiveness.)

NIPT screening generates a positive NMB when	**Common practice**	**Presentism**	**Actualism**	**Placeholder necess.**	**Asymmetric actualism**
The maternal utility related to a negative test result (screening related distress)	≥ 0,95	≥ 0,97	≥ 0,97	≥ 0,96	≥ 0,95
The maternal utility related to parenting a child with DS (long-term)	remains positive when utility= 0	remains positive when utility= 0	remains positive when utility= 0	remains positive when utility= 0	remains positive when utility= 0
The utility of a child with DS	n.a.	n.a.	≥ 0,85	≥ 0,99	remains positive when utility= 1
The yearly healthcare related costs of an adult with DS	≥ € 48.000 yearly additional **benefit** compared to an adult without DS	n.a.	≤ €8000	≥ € 16.000 yearly additional **benefit** compared to an adult without DS	≥ € 55.000 yearly additional **benefit** compared to an adult without DS
The replacement rate	remains positive at 0%	remains positive at 0%	≤ 34%	remains positive at 0%	remains positive at 0%
The termination rate	remains positive at 0%	remains positive at 0%	≤ 20%	≤ 15%	remains positive at 0%

Parental utility

As the mother (and father) are present people, they are not discriminated by the alternative views explored in our study. As a consequence, changes in parental utility generate the same absolute difference in NMB for each perspective. We illustrate that overall screening related distress (included as the utility related to a negative test result) has a substantial impact on cost-effectiveness because it affects almost the entire screened population in comparison with other test results, such as the utility related to an unexpected birth of a child with DS. As shown in table 8, in order to generate a positive NMB this utility value should be higher than 0,95 to 0,97 (depending on the perspective). Decreasing the long-term utility related to being a parent of a child with DS is found to have a substantial positive impact on cost-effectiveness as NIPT now avoids larger parental QoL losses.

Neonatal utility

Neonatal utility is only taken into account by an actualist, necessitarian and asymmetric actualist view. Decreasing the utility related to DS increases the cost-effectiveness of NIPT as the prevention of DS now generates a larger QALY gain. Similarly, increasing the utility of a child with DS decreases the cost-effectiveness of NIPT. The common practice approach, presentist and asymmetric perspective however still give rise to a positive NMB, even when we assume that a child with and without DS both experience a utility equal to one. In this case, the asymmetric actualist view would consider the child with DS to also have neutral value. Adopting an actualist view, the NMB turns negative when a child with DS would experience a utility of 0,85 or higher. The placeholder necessitarian view only generates negative results when this utility would be 0,99 or higher.

The considered scenario in Table 7 has the same absolute (positive) effect on the NMB for the actualist and asymmetric actualist view, because both perspectives include all potential future children with DS.

The positive effect on the necessitarian NMB is more limited, as this perspective does not consider the children with DS that are lost but not replaced.
The actualist and asymmetric actualist view however diverge when a child with DS experiences a utility higher than 0,6, or when the boundary of suffering is set lower than 0,55, implying that the future children with DS would not suffer. In this case, the actualist view still takes into account all future children that can exist, while the asymmetric view considers both children with and without DS to have a neutral QALY value. The asymmetric approach therefore implies that only the future costs related to the unborn children (but not their QALYs) will have an impact on cost-effectiveness. Interestingly, this view hence generates the same outcomes as the common practice approach, even though it is now justified by an alternative (yet consistent) view.

Future costs related to DS

In the base case analysis we assumed that a person with DS requires a yearly institutionalization cost of €12.000 between the age of 18 and 55. We now consider a scenario in which adults with DS generate financial revenues that might compensate for these costs and result in a yearly net financial benefit of € 10.000. (This is an extreme assumption as this implies that adults with DS would actually generate an additional net benefit of €10.000 compared to adults without DS.)
These financial implications are included by all views other than presentism. Only for the necessitarian view the absolute difference differs, because only future savings related to the necessary future children are taken into account. The critical values again illustrate how an actualist and asymmetric actualist view respectively generate the 'worst' and 'best' cost-effectiveness outcomes. While an actualist view considers NIPT to be no longer cost-effective when the yearly costs of an adult with DS are lower than €8000, an asymmetric actualist view only generates a negative NMB when an adult with DS would generate €55.000 additional revenue compared to an adult without DS.

Replacement and termination rate

Varying the replacement and termination rate has an effect on both maternal and neonatal costs and QALYs and therefore leads to different effects on the NMB. A decrease in replacement rate avoids pregnancy related costs and maternal distress and leads to a (limited) positive effect on the cost-effectiveness, according to the common practice, presentist and asymmetric actualist approach. Less or zero replacement children also imply a significant decrease in future neonatal QALYs for the actualist and necessitarian view (according to an asymmetric view these replacement children have neutral value). In this case, the no-screening scenario now generates more neonatal QALYs because more children (with DS) are born, which has a negative impact on the NMB for both perspectives. This effect is smaller for the necessitarian view, as only the QALYs related to the *necessary* future children are included. Increasing the replacement rate to 100% has an opposite, positive effect.

Increasing the termination rate has a similar positive effect on all perspectives. The combination of a 100% termination rate with a zero replacement rate again illustrates the difference between the actualist and necessitarian approach. Even a termination rate equal to zero still generates a positive NMB according the common practice approach, presentist and asymmetric actualist view, because NIPT screening still generates QALYs for the mother who will still be informed whether she will give birth to a child with or without DS.

Sensitivity

Last, we explore the criticism that an actualist view would bias against more accurate screening tools. As current levels of sensitivity and specificity are already above 99%, we find that the effect of an increase in these variables is limited for all perspectives. In fact, the screening program becomes slightly more cost-effective, as the savings and QALY gains of avoiding false negatives outweigh the costs and QALY losses related to the increase in true positives.

4.4 Discussion

In this chapter, we illustrate that different, plausible perspectives can be adopted to estimate the value-for-money generated by prenatal genetic screening and that these perspectives correspond to different ethical views on the moral value of future, unborn lives. We conduct a long-term cost-utility analysis of NIPT screening according to four perspectives and compare the obtained net monetary benefits. The diverging cost-effectiveness results, ranging from a decrease of 50% to an increase of 10% compared to common practice, illustrate the need to make ethical presuppositions explicit. Our findings are not only relevant for prenatal genetic screening programs but for the evaluation of all health interventions that affect future lives, such as fertility treatments, maternal vaccination or the prevention of vertically transmittable diseases.

Besides a presentist view that values NIPT in terms of the generated value for the parents undergoing screening, we explore three alternative views that take into account the unborn children in different manners. While an actualist perspective considers how NIPT affects all the children that *could* be born, a (broad) necessitarian perspective only takes into account how NIPT improves the QoL of the children that *will* be born. Following an asymmetric actualist perspective, all future children have a neutral QALY value and only the future suffering avoided by NIPT should be taken into account. Even though these three approaches all attribute value to the unborn children, the asymmetric actualist view leads to the smallest NML while the actualist view results in the largest net monetary loss, which is twice as large as the common practice result.

Up to now, it has typically been assumed that, *if* neonatal QALYs are consistently included in HTA, this can only be done according to an actualist perspective that takes into account all possible future children. We find that this perspective indeed has a negative impact on cost-

effectiveness because of the QALY losses included after pregnancy termination. The critical values (Table 8) however illustrate that this perspective can also lead to a positive NMB for reasonable parameter estimations. Nevertheless, in a more general HTA context, an actualist perspective seems to give rise to some fundamental challenges. The value attributed to every potential future child seems to provide a reason to bring these lives into existence. Increasing live births would become a very cost-effective manner to generate QALYs and hence, (in theory) to increase population health. Furthermore, the specific QoL reference points on which QALYs are based seem to imply that every life with a utility that is perceived better than death would be considered a QALY gain.

In this paper, we however illustrate that there are alternative, yet consistent approaches that avoid these challenges while still taking into account unborn, future children. A necessitarian placeholder view avoids the valuation of the coming into existence of future lives, by only considering changes in the QoL of necessary future placeholders. An asymmetric actualist view considers all future children to have *neutral* value, unless they would suffer in the future. In this case, only the 'QoL shortfall' to have a future life worth living is taken into account. The latter approach therefore requires a consensual boundary in terms of QoL and life expectancy, defining whether or not a future life has a right to healthcare. If a child with DS would fall within the neutral utility range, the asymmetric view implies that all unborn children have a neutral QALY value and only their costs will have an impact on the cost-effectiveness of NIPT. As a consequence, we show that the 'common practice' approach can be justified by a consistent asymmetric actualist perspective and might reflect a general intuition that the screened children (with and without DS) have a neutral value.
Some recent empirical studies explored this intuition of neutrality. Caviola et al. did however not find evidence for an asymmetric view, as respondents were not neutral but attributed a positive value to the

creation of happy lives and symmetrically, a negative value to the creation of unhappy lives. This study however considered a very broad theoretical context in which many other dimensions that make a live birth valuable could be taken into account by respondents.(32,33) In the previous chapter, we investigated public preferences in a specific healthcare priority-setting context and found that most people were indifferent with regard to future lives. We were however unable to distinguish whether respondents adopted a presentist view (excluding all future lives), or an asymmetric view (being neutral about healthy future lives), as we did not consider cases in which lives were unquestionably not worth living.

More practically, this study also revealed a current lack of reliable data to conduct long-term cost-utility analyses of prenatal genetic screening programs. First, even though it is expected that utility values are country- or context specific, we only found one study that reported a utility decrement related to DS which was based on a proxy-assessed survey in China.(24) The adoption of an actualist or necessitarian perspective furthermore requires reliable data about replacement rates and their related costs, which are currently unavailable. Also these data are typically context specific. Choi et al. for example found that the decision after a diagnosis of DS is dependent on demographic factors such as religion, maternal age, current household size and history of abortion but also psychosocial factors including perceived parenting burden or the support of others.(34)

Second, we showed how screening-related distress and the long-term QoL experienced by parents of a child with DS can have a substantial impact on cost-effectiveness but evidence on these effects is very limited. The only available utility values were collected in one country (US) and were measured during one specific period of pregnancy, which does not allow the consideration of changes in QoL throughout the screening and decision-making process. (8,23,35,36) Furthermore, utility data related to false positives or false negatives had to be taken

from a study in 1999 and are therefore likely to be no longer accurate. We are aware that our outcomes are significantly more sensitive to these utility values compared to existing evaluations, because first line NIPT screening is compared to a no-screening scenario. Typically, cost-effectiveness studies compare alternative screening tools because of which the effect of parental QoL will be more limited. Nevertheless, this does not justify the complete exclusion of this dimension, as we show that it is an important driver of the overall value-for money consideration of prenatal genetic testing.

Finally, we emphasize that the reported outcomes are based on an illustrative model that was developed for methodological purposes only. The field of prenatal genetic screening is rapidly evolving and programs are detecting an increasing number of chromosomal abnormalities that we did not consider in this model. An accurate economic evaluation of these programs therefore requires more detailed data on the short-term consequences of screening and the long-term costs and QoL related to each of these detected genetic abnormalities. However, with this study we hope to have shown that up-to-date data are insufficient to obtain robust cost-effectiveness evidence, as long as a consensual view on the value of future lives in HTA is lacking.

References

1. Nshimyumukiza L, Beaumont J-A, Duplantie J, Langlois S, Little J, Audibert F, et al. Cell-Free DNA–Based Non-invasive Prenatal Screening for Common Aneuploidies in a Canadian Province: A Cost-Effectiveness Analysis. Journal of Obstetrics and Gynaecology Canada. 2018 Jan 1;40(1):48–60.
2. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
3. Walker BS, Nelson RE, Jackson BR, Grenache DG, Ashwood ER, Schmidt RL. A Cost-Effectiveness Analysis of First Trimester Non-Invasive Prenatal Screening for Fetal Trisomies in the United States. PLOS ONE. 2015 Jul 2;10(7):e0131402.
4. Fairbrother G, Burigo J, Sharon T, Song K. Prenatal screening for fetal aneuploidies with cell-free DNA in the general pregnancy population: a cost-effectiveness analysis. J Matern Fetal Neonatal Med. 2016 Apr 2;29(7):1160–4.
5. Song K, Musci TJ, Caughey AB. Clinical utility and cost of non-invasive prenatal testing with cfDNA analysis in high-risk women based on a US population. The Journal of Maternal-Fetal & Neonatal Medicine. 2013 Aug 1;26(12):1180–5.
6. Walker BS, Jackson BR, LaGrave D, Ashwood ER, Schmidt RL. A cost-effectiveness analysis of cell free DNA as a replacement for serum screening for Down syndrome. Prenatal Diagnosis. 2015;35(5):440–6.
7. Benn P, Curnow KJ, Chapman S, Michalopoulos SN, Hornberger J, Rabinowitz M. An Economic Analysis of Cell-Free DNA Non-Invasive Prenatal Testing in the US General Pregnancy Population. PLoS One. 2015 Jul 9;10(7):e0132313.
8. Kaimal AJ, Norton ME, Kuppermann M. Prenatal Testing in the Genomic Age: Clinical Outcomes, Quality of Life, and Costs. Obstetrics & Gynecology. 2015 Oct;126(4):737–746.
9. Anh ND, Trang LDM, Anh NQ. First-trimester screening versus non-invasive prenatal testing for Down syndrome at high-risk pregnant women in Hanoi Obstetrics and Gynecology Hospital, Vietnam: A cost-utility analysis. International Journal of Healthcare Management. 2020 May 12;0(0):1–8.
10. Ohno M, Caughey A. The role of noninvasive prenatal testing as a diagnostic versus a screening tool – a cost-effectiveness analysis. Prenatal Diagnosis. 2013;33(7):630–5.
11. Zhang W, Mohammadi T, Sou J, Anis AH. Cost-effectiveness of prenatal screening and diagnostic strategies for Down syndrome: A

microsimulation modeling analysis. PLoS One [Internet]. 2019 Dec 4 [cited 2020 Aug 14];14(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6892535/
12. Brown J, Buxton M. The economic perspective. Br Med Bull. 1998;54(4):993–1009.
13. Petrou S. Methodological limitations of economic evaluations of antenatal screening. Health Economics. 2001;10(8):775–8.
14. Mooney G, Lange M. Ante-natal screening: What constitutes 'benefit'? Social Science & Medicine. 1993 Oct;37(7):873–8.
15. Trussell J. The cost of unintended pregnancy in the United States. Contraception. 2007 Mar 1;75(3):168–70.
16. Goldhaber-Fiebert JD, Brandeau ML. Evaluating Cost-Effectiveness of Interventions that Affect Fertility and Childbearing: How Health Effects are Measured Matters. Med Decis Making. 2015 Oct;35(7):818–46.
17. Little SE, Janakiraman V, Kaimal A, Musci T, Ecker J, Caughey AB. The cost-effectiveness of prenatal screening for spinal muscular atrophy. American Journal of Obstetrics & Gynecology. 2010 Mar 1;202(3):253.e1-253.e7.
18. Stoll K, Lutgendorf M, Knutzen D, Nielsen PE. Questioning the costs and benefits of non-invasive prenatal testing. The Journal of Maternal-Fetal & Neonatal Medicine. 2014 Apr 1;27(6):633–4.
19. Neyt M, Hulstaert F, Gyselaers W. Introducing the non-invasive prenatal test for trisomy 21 in Belgium: a cost-consequences analysis. BMJ Open. 2014 Nov;4(11):e005922.
20. Hulstaert F, Neyt M, Gyselaers W. The non-invasive prenatal test (NIPT) for trisomy 21 – health economic aspects [Internet]. Belgium: Belgian Health Care Knowledge Centre (KCE); 2014. Available from: http://kce.fgov.be/sites/default/files/page_documents/KCE_222_Non_invasive_prenatal_%20test_Report.pdf
21. UZ Leuven. Niet-invasieve prenatale test (NIPT) [Internet]. [cited 2022 Feb 10]. Available from: https://www.uzleuven.be/nl/niet-invasieve-prenatale-test-nipt
22. Cleemput I, Wilder PV, Vrijens F, Huybrechts M, Ramaekers D. Richtlijnen voor farmaco-economische evaluaties in België. Federaal Kenniscentrum voor de Gezondheidszorg; 2008 p. 92. Report No.: KCE report 78A.
23. Kuppermann M, Norton ME, Thao K, O'Leary A, Nseyo O, Cortez A, et al. Preferences regarding contemporary prenatal genetic tests among women desiring testing: implications for optimal testing strategies. Prenatal Diagnosis. 2016;36(5):469–75.

24. Mok WKY, Wong WH-S, Mok GTK, Chu YWY, Ho FKW, Chow CB, et al. Validation and application of health utilities index in Chinese subjects with down syndrome. Health Qual Life Outcomes. 2014 Dec;12(1):144.
25. Arrhenius G. Population Ethics. The Challenge of Future Generations. (Unpublished version May 2016).
26. Greaves H. Population Axiology. Philosophy Compass. 2017;12(11):e12442.
27. Luyten J, Verbeke E, Schokkaert E. To be or not to be: Future lives in economic evaluation. Health Economics. 2022;31(1):258–65.
28. Verbeke E, Luyten J. Future Offspring Costs in Economic Evaluation. PharmacoEconomics. 2022 Feb 1;40(2):141–7.
29. Heyd D. Genethics: Moral Issues in the Creation of People. University of California Press; 1992.
30. Broome. The Wellbeing of Future Generations. In: The Oxford Handbook of Wellbeing and Public Policy. Oxford University Press; 2016. p. 901–28.
31. Broome J. Weighing Lives. New York: Oxford University Press; 2004. 278 p.
32. Caviola L, Althaus D, Mogensen AL, Goodwin GP. Population ethical intuitions. Cognition. 2022 Jan;218:104941.
33. Spears D. Making people happy or making happy people? Questionnaire-experimental studies of population ethics and policy. Soc Choice Welf. 2017 Jun;49(1):145–69.
34. Choi H, Van Riper M, Thoyre S. Decision Making Following a Prenatal Diagnosis of Down Syndrome: An Integrative Review. Journal of Midwifery & Women's Health. 2012;57(2):156–64.
35. Kuppermann M, Feeny D, Gates E, Posner SF, Blumberg B, Washington AE. Preferences of women facing a prenatal diagnostic choice: long-term outcomes matter most. Prenatal Diagnosis. 1999;19(8):711–6.
36. Kuppermann M, Nease RF, Learman LA, Gates E, Blumberg B, Washington AE. Procedure-related miscarriages and down syndrome–affected births: implications for prenatal testing based on women's preferences. Obstetrics & Gynecology. 2000 Oct 1;96(4):511–6.

Appendix: detailed calculation of the NMBs in different scenarios, according to the different perspectives

(effects are expressed in terms of QALYs)

	BUA					Presentism				
	Δ cost	*Δ effect*	*NMB (thresh=€50.000)*	*Absolute difference w base case*	*Δ% base case*	*Δ cost*	*Δ effect*	*NMB*	*Absolute difference w base case*	*Δ% base case*
Base case	- 81.770.274 €	4728	318.151.639 €			488.475 €	4728	235.892.891 €		
UTILITY PARENTS										
SHORT TERM utility: experienced during first year after screening										
utility negative test result 0,931	- 81.770.274 €	-3597	- 98.103.295 €	- 416.254.934 €	-131%	488.475 €	-3597	- 180.362.044 €	- 416.254.934 €	-176%
giving birth to a child with DS, if DS was expected=1, if DS was unexpected=0	- 81.770.274 €	4748	319.178.296 €	1.026.657 €	0%	488.475 €	4748	236.919.548 €	1.026.657 €	0%
LONG TERM utility: experienced during the rest of their life										
utility mother of a child with DS= 0, mother of a child without DS=1	- 81.770.274 €	7450	454.277.202 €	136.125.562 €	43%	488.475 €	7450	372.018.453 €	136.125.562 €	58%
utility father included (equal utility	- 81.770.274 €	9455	554.533.004 €	236.381.365 €	74%	488.475 €	9455	472.274.256 €	236.381.365 €	100%
UTILITY OF CHILD										
child with DS= 0,3	- 81.770.274 €	4728	318.151.639 €	- €	0%	488.475 €	4728	235.892.891 €	- €	0%
utility child with DS= utility child without DS =1	- 81.770.274 €	4728	318.151.639 €	- €	0%	488.475 €	4728	235.892.891 €	- €	0%
FINANCIAL CONSEQUENCES CHILD										
Net financial implication of an adult person with DS compared to without DS: generates a yearly income of €10.000	35.613.004 €	4728	200.768.362 €	- 117.383.278 €	-37%	488.475 €	4728	235.892.891 €	- €	0%
TERMINATION & REPLACEMENT										
Replacement rate= 0%	- 82.070.814 €	4748	319.463.934 €	1.312.295 €	0%	488.475 €	4748	236.904.645 €	1.011.755 €	0%
Replacement rate= 100%	- 81.670.094 €	4721	317.714.208 €	- 437.432 €	0%	488.475 €	4721	235.555.639 €	- 337.252 €	0%
Termination rate=100%	- 85.164.398 €	4912	330.772.254 €	12.620.615 €	4%	488.475 €	4912	245.119.381 €	9.226.490 €	4%
Termination = 100%, replacement rate= 0%	- 85.477.225 €	4933	332.138.200 €	13.986.561 €	4%	488.475 €	4933	246.172.501 €	10.279.610 €	4%
SENSITIVITY										
100% sensitivity & specificity (0 false negatives or false positives)	- 82.095.848 €	4852	324.679.807 €	6.528.168 €	2%	488.475 €	4852	242.095.484 €	6.202.593 €	3%

	Actualism					Placeholder necessitarianism				
	Δ cost	*Δ effect*	*NMB*	*Absolute difference w base case*	*Δ% base case*	*Δ cost*	*Δ effect*	*NMB*	*Absolute difference w base case*	*Δ% base case*
Base case	- 81.770.274 €	2635	213.504.335 €			- 69.829.066 €	4011	270.380.501 €		
UTILITY PARENTS										
SHORT TERM utility: experienced										
utility negative test result 0,931	- 81.770.274 €	-5690	- 202.750.600 €	- 416.254.934 €	**-195%**	- 69.829.066 €	-4314	- 145.874.434 €	- 416.254.934 €	**-154%**
giving birth to a child with DS, if DS was expected=1, if DS was unexpected=0	- 81.770.274 €	2655	214.530.992 €	1.026.657 €	0%	- 69.829.066 €	4032	271.407.158 €	1.026.657 €	0%
LONG TERM utility: experienced										
utility mother of a child with DS= 0, mother of a child without DS=1	- 81.770.274 €	5357	349.629.897 €	136.125.562 €	64%	- 69.829.066 €	6734	406.506.063 €	136.125.562 €	50%
utility father included (equal utility	- 81.770.274 €	7362	449.885.700 €	236.381.365 €	**111%**	- 69.829.066 €	8739	506.761.866 €	236.381.365 €	87%
UTILITY OF CHILD										
child with DS= 0,3	- 81.770.274 €	6292	396.375.790 €	182.871.455 €	86%	- 69.829.066 €	7115	425.601.272 €	155.220.771 €	57%
utility child with DS= utility child without DS =1	- 81.770.274 €	-3949	- 115.664.285 €	- 329.168.620 €	**-154%**	- 69.829.066 €	-1577	- 9.016.888 €	- 279.397.389 €	**-103%**
FINANCIAL CONSEQUENCES CHILD										
Net financial implication of an adult person with DS compared to without DS: generates a yearly income of €10.000	35.613.004 €	2635	96.121.057 €	- 117.383.278 €	-55%	30.514.102 €	4011	170.037.333 €	- 100.343.167 €	-37%
TERMINATION & REPLACEMENT										
Replacement rate= 0%	- 82.070.814 €	-5339	- 184.888.233 €	- 398.392.568 €	**-187%**	- 34.305.981 €	227	45.637.527 €	- 224.742.973 €	-83%
Replacement rate= 100%	- 81.670.094 €	5293	346.301.857 €	132.797.523 €	62%	- 81.670.094 €	5272	345.294.825 €	74.914.324 €	28%
Termination rate=100%	- 85.164.398 €	2746	222.475.170 €	8.970.835 €	4%	- 72.693.601 €	4172	281.307.148 €	10.926.648 €	4%
Termination = 100%, replacement rate= 0%	- 85.477.225 €	-5554	- 192.205.343 €	- 405.709.678 €	**-190%**	- 35.594.036 €	219	46.562.680 €	- 223.817.821 €	-83%
SENSITIVITY										
100% sensitivity & specificity (0 false negatives or false positives)	- 82.095.848 €	2777	220.960.911 €	7.456.576 €	3%	- 70.072.434 €	4161	278.140.671 €	7.760.170 €	3%

	Asymmetric actualism				
	Δ cost	*Δ effect*	*NMB*	*Absolute difference w base case*	*Δ% base case*
Base case	- 81.770.274 €	5459	354.725.930 €		
UTILITY PARENTS					
SHORT TERM utility: experienced					
utility negative test result 0,931	- 81.770.274 €	-2866	- 61.529.004 €	- 416.254.934 €	**-117%**
giving birth to a child with DS, if DS was expected=1, if DS was unexpected=0	- 81.770.274 €	5480	355.752.587 €	1.026.657 €	0%
LONG TERM utility: experienced					
utility mother of a child with DS= 0, mother of a child without DS=1	- 81.770.274 €	8182	490.851.493 €	136.125.562 €	38%
utility father included (equal utility	- 81.770.274 €	10187	591.107.295 €	236.381.365 €	67%
UTILITY OF CHILD					
child with DS= 0,3	- 81.770.274 €	9117	537.597.386 €	182.871.455 €	52%
utility child with DS= utility child without DS =1	- 81.770.274 €	4728	318.151.639 €	- 36.574.291 €	-10%
FINANCIAL CONSEQUENCES CHILD					
Net financial implication of an adult person with DS compared to without DS: generates a yearly income of €10.000	35.613.004 €	5459	237.342.653 €	- 117.383.278 €	-33%
TERMINATION & REPLACEMENT					
Replacement rate= 0%	- 82.070.814 €	5479	356.038.225 €	1.312.295 €	0%
Replacement rate= 100%	- 81.670.094 €	5452	354.288.499 €	- 437.432 €	0%
Termination rate=100%	- 85.164.398 €	5674	368.864.970 €	14.139.040 €	4%
Termination = 100%, replacement rate= 0%	- 85.477.225 €	5695	370.230.917 €	15.504.987 €	4%
SENSITIVITY					
100% sensitivity & specificity (0 false negatives or false positives)	- 82.095.848 €	5586	361.400.302 €	6.674.372 €	2%

Discussion

1. Summary

The objective of this PhD thesis was to contribute to the development of a uniform HTA methodology by studying the questions that future lives give rise to in cost-effectiveness assessments. After illustrating the relevance of future life effects and the need for unambiguous guidelines, this thesis explored different plausible perspectives on the value of future lives that were proposed in the population ethics literature. In order to investigate the public's support for these perspectives, a survey was developed to measure if, and when, laypeople take into account future lives in a healthcare priority-setting context. Finally, the studied ethical perspectives were applied and evaluated in a long-term cost-utility analysis of NIPT screening in which the unborn, screened children are taken into account according to four different views.

The objective of **chapter one** was to explore how future lives are treated in current cost-effectiveness evidence and to identify the fundamental problems that arise when aiming to define a consensual view on how to consider future lives in HTA. We studied existing cost-effectiveness analyses in two different types of health domains in which future lives can be affected. In the first subchapter we summarized the literature on perinatal mental health interventions, an example of healthcare that, besides the effects on the parents, is also likely to affect the long term *quality of life* of their future children. Therefore, these interventions can be considered to 'make future people healthy' and we explored to which degree these *Quality of Life* (QoL) gains for future offspring are taken into account in cost-effectiveness evidence. In the second subchapter we studied the cost-effectiveness evidence on *Assisted Reproductive Technologies* (ARTs), which are used to address infertility of patients. As the goal of these interventions is to 'create' a new future life, this

subchapter explored cost-effectiveness evidence on healthcare that 'makes healthy people'. This review therefore allowed us to study how these 'new' future lives have been taken into account in current evaluations. Both literature reviews led to the same general conclusion that current methods are ill-equipped to provide a comprehensive understanding of the value-for-money generated by these interventions. Cost-effectiveness analyses were found to include different types of costs and effects for different time horizons. Even though perinatal mental healthcare is shown to generate long-term health-effects on the mother, father and child, only 6 of the 39 studies included health outcomes for the child and only one study considered paternal health. Cost-effectiveness results were reported in terms of different outcome measures, hampering the comparability of study results. The need for a consistent methodology and use of uniform outcome measures was therefore illustrated. QALYs seem to be the preferred outcome unit as they allow the consideration of different health dimensions for the different individuals affected by an intervention. Including long-term effects on the QoL of future offspring however requires access to long-term effectiveness data with regard to all parties affected, as well as appropriate utility values that allow the inclusion of these data. In case healthcare affects the *number* of people that will come into existence in the future, it is currently unclear if and how these effects should be taken into account in HTA. In the specific domain of ARTs, some authors have argued that the use of QALYs for the inclusion of live births is not appropriate, as the conception of a new life cannot be considered an improvement of health.(1,2) In other health domains this question has however not yet been raised, even though interventions such as cancer treatment, diabetes programs or lifesaving interventions are also likely to affect future birth (or conception) rates.

The findings of this first chapter illustrated the importance of two currently debated questions. First, as future lives are often affected through their presently existing parents, they can be considered as

(vertical) spillover effects, a topic that has received increasing attention in the literature.(3–6) Second, the consideration of broader, societal costs and effects related to future lives in HTA will be determined by the perspective that is considered to be appropriate in HTA. In this thesis, we however departed from the current HTA guidelines that typically prescribe (at least) the inclusion of all relevant health effects for all parties affected.(7–10) We aimed to show that even when a uniform consensus would be reached on these open questions, other fundamental challenges remain. The way in which future lives (in terms of both their QoL and existence) would be considered in economic evaluation will also depend on the underlying ethical view that is adopted about the value of future lives. While cost-effectiveness studies are in general based on the ethical perspective of total utilitarianism in which all QALYs are (impartially) summed, this perspective might lead to problematic conclusions in case future lives are also affected. Does the coming into existence of people in the future always makes an outcome better? Are the QALYs generated by the coming into existence of a life equally valuable as the QALYs generated by future QoL gains? Is it acceptable that the QALY gains related to higher birth rates outweigh the QALY gains related to an increase in future people's health? Hoping to empty this Pandora's box further, we turned to the field of population ethics to study the underlying challenges related to these questions and to identify alternative ethical perspectives.

This population ethics literature study resulted in the exploration and discussion of three '*person affecting'* views in **chapter two**. Each of these views defines a different subset of lives that should be taken into account when comparing alternative programs that affect present and (or) future lives. Presentism only takes into account present lives and is typically rejected as it implies a complete indifference with regard to the wellbeing and existence of all future lives, which also contradicts with the perspective adopted in other domains such as climate change policy. Second, actualism takes into account all lives that might possibly exist

in the future. The main challenges related to this view are that future life effects (in terms of QALYs) would likely have a disproportionate impact on cost-effectiveness results and that it seems to require the prioritization of young, fertile patients. As a compromise, we proposed necessitarianism as an attractive alternative view in HTA. This view distinguishes necessary future lives whose future existence is independent of the intervention under evaluation, from potential future lives, whose existence is a direct consequence of the intervention. Necessitarianism therefore captures the intuition that our moral obligation to safeguard the wellbeing of the future people that will exist, is stronger than our obligation to bring additional people into existence. By only including necessary future lives, extreme cost-effectiveness outcomes of interventions with fertility effects are avoided, while still taking into account the implications of preventing disease in future generations.

Considering these different perspectives with regard to future people's QALYs also raised the question how the costs associated with future lives should be considered. Currently, it remains an open question whether or not future related and unrelated costs should be included in economic evaluation.(11,12) In the second subchapter, we however aimed to draw attention to the fact that future lives have typically been overlooked in this future costs discussion. The relevance of 'future offspring costs' was illustrated, as they are shown to have a substantial impact on cost calculations and vary across interventions and demographic groups. Also in this context we proposed a necessitarian approach and the potential implications of this view on cost calculations were explored.

Subsequently, the objective of **chapter three** was to study the public support for the previously discussed ethical perspectives, based on a population survey. Dutch and Flemish respondents were asked to prioritize healthcare between two identical women in equal need of treatment but of which only one will have a future child that also benefits

from the treatment. Questions varied in terms of the development stage of the affected future offspring and the type of intervention effect (affecting the coming into existence or the Quality of Life of the future offspring). The majority of respondents (70%) indicated to be indifferent between both women and therefore did not seem to value the affected future offspring. Based on individual choice patterns, it was concluded that most people (60%) seem to support a presentist view in healthcare priority-setting, 20% of respondents was identified as an actualist and 20% did not seem to adopt a consistent perspective. In the introductory questions however, more than 70% of respondents indicated that they cared about future generations and would be willing to make sacrifices for them. Our findings therefore seem to illustrate how people's value judgements might differ when they evaluate outcomes per se, versus when they also consider the specific actions required to obtain these outcomes. Also more generally, our findings were very likely to be context-specific and their robustness should therefore be tested in other contexts as well.

As a final part of this thesis, we aimed to put our developed insights into practice in **chapter four** by means of an illustrative long-term cost-utility analysis of NIPT screening, an intervention of which the cost-effectiveness will be largely determined by the way in which the consequences for the unborn child are taken into account. Existing economic evaluations exclude the QALYs of the unborn, screened child in order to avoid that pregnancy termination after a positive screening result should be taken into account as a lifetime amount of QALYs lost. We studied four different plausible views that can be adopted to evaluate the neonatal QALY gains (or losses) generated by NIPT, implying four different manners of considering neonatal costs and QALYs. The presentist view valued NIPT solely based on the consequences for the parents, while an actualist perspective did not distinguish between present and future lives and therefore also included the costs and QALYs of the unborn children. Following a necessitarian view, NIPT was

evaluated based on the degree to which genetic screening improves the QoL of the children that will live in the future. Last, an asymmetric actualist perspective valued NIPT in terms of the suffering it avoids in future children. This perspective considers future lives in good health to have neutral value, while future lives with an expected QoL below a boundary that defines 'suffering' have a negative value. For this perspective, neonatal QALY gains were calculated asymmetrically so that pregnancy termination after a positive screening result was taken into account as a QALY gain, reflecting the future suffering that is avoided. The obtained cost-effectiveness outcomes illustrated the need to make ethical presuppositions explicit as each view led to diverging results. Furthermore, it was shown that the common practice approach of only including neonatal costs but not QALYs, might reflect people's intuition to consider a live birth to have neutral value, as long as the unborn child is considered to have a life worth living.

2. Strengths, limitations and avenues for further research

Next to the more specific points that were outlined in the separate papers, the following section summarizes more general strengths and limitations of the dissertation, as well as avenues for future research.

Studying if and how future lives should be taken into account in HTA, is a question that is related to many different ongoing debates in the health economics literature such as the consideration of spillover effects, future (unrelated) costs and the use of discounting in HTA.(3,4,11–15) This thesis however explored a different fundamental question, and studied for the first time the ethical challenges that arise when accounting for future lives and the different ethical perspectives that can be adopted to address these challenges. The population ethics field was found to offer (often unexplored) valuable insights, not only for HTA, but for all other domains in which priority-setting affects both present and future people.

This PhD project was however not limited to a study of the theoretical literature. To make the acquired insights practically relevant, we also studied how the public opinion corresponds to the ethical views proposed in the population ethics literature, and how these theories can in practice be applied in cost-utility analyses. Throughout the separate chapters, a wide range of different health domains and health interventions was considered, which allowed the discussion of the methodological challenges related to future lives as comprehensively as possible.

Furthermore, the effects of healthcare on future lives form an abstract research matter that gives rise to a plethora of fundamental questions. The main limitation of this dissertation is therefore that we were not able to fully empty the Pandora's box and that potentially relevant dimensions had to be disregarded to avoid overcomplexity and to maintain manageable research goals. These dimensions offer nevertheless interesting avenues for further research.

To begin with, the overall use of QALYs in current health economic evaluation was not questioned. While QALYs allow the consideration of different parties affected by healthcare in a comparable manner, they still imply the adoption of a narrow perspective that only attributes value to changes in people's quality-adjusted life expectancy. Specifically in the context of genetic screening and infertility treatment, these interventions clearly serve multiple purposes beyond health itself. The value of crucial elements such as the psychosocial consequences of a pregnancy or parental autonomy in the context of pregnancy termination are difficult to capture in one, comprehensive utility value.(16) We however considered this fundamental question to be beyond the specific context of future life effects and it was therefore considered out of scope of this research project.

Second, discounting was not extensively discussed, as this field investigates how much *weight* should be attributed to costs and effects

occurring in the future, but not *whose* future costs and effects should be considered in the first place. Nevertheless, the applied discount rate will have a substantial impact on the relevance of future life effects. As HTA guidelines currently prescribe the use of positive discount rates for QALYs (typically between 1,5 and 3,5%), effects of healthcare on subsequent generations have a rapidly decreasing impact on cost-effectiveness outcomes. This current practice has however been criticized in different domains.(14,15,17–22) Apart from the use of a zero discount rate, alternative methods have been proposed such as negative discount rates, variable discount rates that decrease throughout time, or intergenerational discount rates that differ per generation.(20,23–26) It would therefore be interesting to study how these alternative methods might provide practical answers to the methodological challenges related to future lives that were discussed in this thesis.

Third, the use of *Disability-Adjusted Life Years* (DALYs) seems to avoid some difficulties related to the inclusion of future lives in HTA. As DALYs reflect the health 'lost' due to disease, they do not seem to give rise to the same 'problematic' consequences when the conception or birth of a child is considered. In theory, the coming into existence of a future life seems to have a neutral value in terms of DALYs, unless these lives would not be in good health. Further research is however required to explore the implications of using DALYs in the context of future life effects and to evaluate whether they represent a more appropriate effectiveness measure.

Last, this dissertation mainly focused on '*person-affecting'* theories discussed in the population ethics literature. These theories hence distinguish different subsets of lives, based on how they are affected by the intervention (e.g. present lives, actual lives, necessary lives and lives that would suffer). However, the literature on *impartial* theories might, in addition, provide interesting alternatives to the total utilitarian view on which current methods are based.(27,28) More specifically, critical-

level utilitarianism seems to offer a promising perspective. Applying this view in HTA would imply that a critical QoL value is subtracted of all QALY values, implying that a health state below this boundary would have a negative value. As this critical value disappears from the equation when a person exists before and after the intervention, this view will only imply a systematic adjustment of QALY values related to the coming into existence of future people.(29–31)

3. Conclusion

Because clear guidance on the treatment of future lives in HTA is lacking, future life effects are currently included in different ways. This has led to incomparable methods and study outcomes, undermining the role of economic evaluation to facilitate consistent and transparent decision-making. This thesis thoroughly investigated and defined the unexplored challenges that future lives give rise to in an HTA context. Subsequently, it was explored whether the population ethics literature can provide answers to these fundamental challenges. It was shown that different plausible perspectives can be adopted, each having their own pros and cons.

While it can be argued that future lives do not yet have a right to healthcare resources, a complete indifference with regard to the wellbeing of future people seems unjustifiable and inconsistent with other policy domains. Including all potential future lives avoids this indifference but seems to require an extreme prioritization of future people's health and especially of population growth, as the birth of a child is considered to generate a lifetime amount of QALYs. Two perspectives were discussed that offer a compromise between the systematic in- and exclusion of future lives. Necessitarianism only takes into account future people whose QoL is affected by the evaluated intervention and therefore excludes all effects on conception or birth rates altogether. The exclusion of all potential future people becomes however counterintuitive in cases where healthcare avoids the coming into existence of lives full of suffering. An asymmetric actualist view therefore does not exclude any lives, rather, it considers future lives to have neutral value as long as they have a life worth living. In case future lives would suffer, they are considered to have a negative value, reflecting our moral obligation to prevent their future suffering.

It was shown that each view that tries to avoid the disproportionate impact of future lives on cost-effectiveness outcomes leads to trouble because certain information is always excluded. While necessitarianism and an asymmetric view seem to offer interesting compromises, these views still cannot avoid that decision-makers make a 'problematic' choice when two alternative interventions only differ in terms of the excluded (or neutral) effects. This is why population ethicists have concluded that it is impossible to define one optimal perspective that satisfies all the conditions of intuitive priority-setting. Also in an HTA context the same impossibility seems to occur. Even though intuitively, it can be argued that healthcare budgets should only be spend to make future people healthy (by safeguarding their QoL or avoiding future suffering), there are still cases in which it seems problematic to not invest in making healthy people (e.g. the prevention of stillbirths, miscarriages, or disease eradication that safeguards the coming into existence of subsequent generations). Therefore, following certain population ethicists, the only remaining option seems to be to bite the bullet and adopt one consensual perspective that leads to the least problematic conclusions. Personally, I consider an asymmetric-necessitarian perspective, which takes into account the QoL of the necessary future population as well as the future suffering of potential lives, the most attractive option. Because of the inevitable opportunity costs of healthcare priority-setting, it seems counterintuitive that ARTs would generate a lifetime amount of QALYs in addition to the health gain generated for the parents with the childwish. On the other hand, it also seems unjustifiable to not ignore the suffering that can be avoided by preventing the conception or birth of a child that would really have a life not worth living. However, the general public did not seem to agree with this view as they were found to generally ignore future life effects when they were to prioritize between patients, even though they did indicate to care about future generations.

To conclude, while the questions raised in this dissertation were sometimes considered unethical by general readers or participants of the population survey, I hope to have shown that not raising these questions would be equally unethical. The adoption of alternative ethical views on the value of future lives leads to diverging cost-effectiveness outcomes that are very likely to have a decision-changing impact. As priority-setting in healthcare and other policy domains inevitably involves opportunity costs that are borne by people in the present or the future, a uniform perspective should be adopted to at least ensure consistent priority-setting.

References

1. Devlin N, Parkin D. Funding fertility: issues in the allocation and distribution of resources to assisted reproduction technologies. Hum Fertil (Camb). 2003 May;6 Suppl 1:S2-6.
2. Collins JA. An international survey of the health economics of IVF and ICSI. Hum Reprod Update. 2002 May 1;8(3):265–77.
3. Al-Janabi H, van Exel J, Brouwer W, Coast J. A Framework for Including Family Health Spillovers in Economic Evaluation. Med Decis Making. 2016 Feb;36(2):176–86.
4. Brouwer WBF. The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. PharmacoEconomics. 2019 Apr 1;37(4):451–6.
5. Grosse SD, Pike J, Soelaeman R, Tilford JM. Quantifying Family Spillover Effects in Economic Evaluations: Measurement and Valuation of Informal Care Time. PharmacoEconomics. 2019 Apr 1;37(4):461–73.
6. McCabe C. Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. PharmacoEconomics. 2019 Apr 1;37(4):457–60.
7. Neumann P, Sanders G, Russell L, Siegel J, Ganiats T. Cost-effectiveness in health and medicine. 2nd ed. New York: Oxford University Press; 2017. 496 p.
8. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Great Britain: Oxford University Press; 2015. 379 p.
9. NICE. The guidelines manual: Process and methods - 7 Assessing cost effectiveness [Internet]. The guidelines manual. 2012 [cited 2018 Jun 20]. Available from: https://www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness
10. Swedish Council on Health Technology Assessment (SBU), EUnetHTA. Methods for health economic evaluations - A guideline based on current practices in Europe. 2015;100.
11. Kearns B. NICE DSU Report. The relevance of future, unrelated health costs in economic evaluation in NICE appraisals. [Internet]. Decision Support Unit, ScHARR, University of Sheffield; 2020 p. 48. Available from: http://nicedsu.org.uk/wp-content/uploads/2020/11/Future_unrelated_costs_Final.pdf
12. de Vries LM, van Baal PHM, Brouwer WBF. Future Costs in Cost-Effectiveness Analyses: Past, Present, Future. Pharmacoeconomics. 2019;37(2):119–30.

13. McCabe C. Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. PharmacoEconomics. 2019 Apr 1;37(4):457–60.
14. Greaves H. Discounting Future Health. In: Ruger, Norheim VOMJJE, editors. Global health priority-setting: Cost-effectiveness and beyond. Oxford, England: Oxford University Press; forthcoming.
15. Frederick S. Valuing future life and future lives: A framework for understanding discounting. Journal of Economic Psychology. 2006 Oct 1;27(5):667–80.
16. Verbeke E, Luyten J. Reproductieve effecten van gezondheidszorg. In: Nomenclatuur en tegemoetkoming in de zorg [Internet]. 1st ed. Intersentia; 2020 [cited 2022 Mar 24]. p. 488. Available from: https://www.jurisquare.be/nl/book/9789400012349/index.html
17. Arrhenius G, Bykvist K. Future Generations and Interpersonal Compensations. Moral Aspects of Energy Use. :132.
18. Beckerman W, Hepburn C. Ethics of the Discount Rate in the Stern Review on the Economics of Climate Change. 2007;8(1):25.
19. Chichilnisky G, Hammond PJ, Stern N. Should We Discount the Welfare of Future Generations? Ramsey and Suppes versus Koopmans and Arrow [Internet]. Centre for Research in Economic Theory and its Applications; 2018. Available from: https://warwick.ac.uk/fac/soc/economics/research/centres/creta/papers/manage/43_-_creta_hammond.pdf
20. Weitzman ML. Why the Far-Distant Future Should Be Discounted at Its Lowest Possible Rate. Journal of Environmental Economics and Management. 1998 Nov 1;36(3):201–8.
21. Guth JH. Resolving the Paradoxes of Discounting in Environmental Decisions. TRANSNATIONAL LAW. 18:20.
22. Brouwer WBF, Niessen LW, Postma MJ, Rutten FFH. Need for differential discounting of costs and health effects in cost effectiveness analyses. BMJ. 2005 Aug 20;331(7514):446–8.
23. Cairns J, Van der Pol M. Constant and decreasing timing aversion for saving lives. Social Science & Medicine. 1997 Dec 1;45(11):1653–9.
24. Gollier C, Weitzman ML. How should the distant future be discounted when discount rates are uncertain? Economics Letters. 2010 Jun 1;107(3):350–3.
25. Freeman MC, Groom B, Panopoulou E, Pantelidis T. Declining discount rates and the Fisher Effect: Inflated past, discounted future? Journal of Environmental Economics and Management. 2015 Sep 1;73:32–49.

26. Jamison DT, Jamison J. Characterizing the Amount and Speed of Discounting Procedures. Journal of Benefit-Cost Analysis. 2011 Apr;2(2):1–56.
27. Arrhenius G. Population Ethics. The Challenge of Future Generations. (Unpublished version May 2016).
28. Greaves H. Population Axiology. Philosophy Compass. 2017;12(11):e12442.
29. Blackorby C, Donaldson D. Social criteria for evaluating population change. Journal of Public Economics. 1984 Nov 1;25(1):13–33.
30. Dasgupta P. Lives and Well-Being. Social Choice and Welfare. 1988;5(2/3):103–26.
31. Blackorby C, Bossert W, Donaldson DJ. Population Issues in Social Choice Theory, Welfare Economics, and Ethics by Charles Blackorby [Internet]. Cambridge Core. 2005

List of publications related to this thesis

Peer-reviewed, international scientific journal articles:

- Verbeke, E., Bogaerts, A., Nuyts, T., Crombag, N., Luyten, J. (2022). Cost-effectiveness of mental health interventions during and after pregnancy: A systematic review. Birth. doi: 10.1111/birt.12623

- Luyten, J., Connolly, M., Verbeke, E., Buhler, K., Scotland, G., Lispi, M., Revelli, A., Borget, I., Cedrin-Durnerin, I., D'Hooghe, T. with D'Hooghe, T. (corresp. author) (2022). Economic evaluation of Medically Assisted Reproduction: an educational overview of methods and applications for healthcare professionals. Best Practice & Research Clinical Obstetrics & Gynaecology. doi: 10.1016/j.bpobgyn.2022.01.008

- Luyten, J., Verbeke, E., Schokkaert, E. (2021). To be or not to be: Future lives in economic evaluation. Health economics, 31 (1), 258-265. doi: 10.1002/hec.4454

- Verbeke, E., Luyten, J. (2021). Future Offspring Costs in Economic Evaluation. Pharmacoeconomics. doi: 10.1007/s40273-021-01102-w

To be submitted:

- Verbeke, E., Schokkaert, E., Luyten, J., Valuing future lives in healthcare priority-setting: a population survey.

- Verbeke, E., Schokkaert, E., Luyten, J., Long-term cost-utility analysis of NIPT in Belgium: screening alternative perspectives on neonatal QALYs.

Peer-reviewed book chapters:

- Verbeke, E., Luyten, J. (2020). De complexiteit van gezondheidsuitkomsten: het statuut van de reproductieve effecten van een interventie. In: S. Lierman, A. Vijverman, F. Louckx, J. Ghysels (Eds.), NOMENCLATUUR EN TEGEMOETKOMING IN DE ZORG, Chapt. 9, (263-263). Intersentia. ISBN: 9789400012349.

- Verbeke, E., Luyten, J., D'Hooghe, T. (2020). The economics of IVF: evaluating the necessity and value of public funding. In: Patient-Centered Assisted Reproduction: How to Integrate Exceptional Care with Cutting-Edge Technology., (106-135). Cambridge: Cambridge University Press. ISBN: 9781108796774.

Government reports:

- Nuyts, T., Crombag, N., Verbeke, E., Hermans, S., Luyten, J., Bogaerts, A. (2021). Veerkracht tijdens en tot 2 jaar na de zwangerschap. Nieuwe uitdagingen voor de geestelijke gezondheidszorg tijdens de eerste 1000 dagen van het leven. Steunpunt Welzijn, Volksgezondheid en Gezin.

Curriculum Vitae

Professional experience

Leuven Institute for Healthcare Policy (LIHP), KU Leuven
April ’18 –
PhD researcher in Biomedical Sciences (Health Economics)

Bank van Breda, Antwerp
July’16 – March’18
Financial analyst (analysis of SME's in the context of private and corporate lending)

Umicore Precious Metals Refining, Hoboken
October’15- June’16
Assistant manager Market Intelligence & Business Research

Education

Staffordshire University (UK), University of Antwerp and University of Bari (IT)
2014-2015
Master of Arts Degree in Economics of Globalization and European Integration (Erasmus Mundus)

Ekonomická univerzita v Bratislave, Slovakia
August ‘14
Erasmus Intensive Program:
New Path of Economic Growth and Social Development for Europe

University of Antwerp
2012-2014
Master of Science Degree in Business Engineering

Ecole des Hautes Etudes Commerciales de Montréal
August’13- January ‘14
Exchange program in Canada

Courses followed during PhD

Effective Altruism Netherlands	2021
High impact Fellowship	
KU Leuven	
Statistics for PhD researchers Biomedical Sciences	2021
Introduction to SAS	2020
Essential tools for R	2019
Welfare Economics	2018
Academic writing	2018
Princeton University	
Workshop on Population and Procreation Ethics	2019
University of Antwerp	
Health Economic Evaluation of Vaccines (SIMID)	2018
Health Economics in Clinical Research	2018

Acknowledgements, Personal Contribution and Conflict of Interest Statements

Scientific acknowledgements

Chapter 1 and 2: The authors wish to thank the editors and anonymous reviewers for their valuable feedback, which has greatly contributed to the quality of the included articles.

Chapter 3: The authors acknowledge that the market research company Dynata was responsible for inviting respondents to the online survey.

Personal contribution

Chapter 1: Evelyn Verbeke (EV) conducted both systematic reviews, included studies were evaluated by EV and Dr. Jeroen Luyten (JL). The first draft of both manuscripts was written by EV, with JL revising. All authors contributed, read and approved the final manuscripts.

Chapter 2: EV and JL worked together on the conceptualization of the included manuscripts. Part 1 of this chapter was written by JL and revised by Dr. Erik Schokkaert (ES) and EV. Part 2 was written by EV, with JL revising.

Chapter 3: With support of JL and ES, EV was responsible for the development of the survey, the organisation of the data collection in cooperation with Dynata and data analysis. EV wrote the first draft of the manuscript with JL and ES revising.

Chapter 4: With support of JL and ES, EV developed the long-term cost-utility model. EV wrote the first draft of the manuscript with JL and ES revising.

Conflict of interest statements

Chapter 1- part 1 (systematic review on the cost-effectiveness of perinatal mental healthcare interventions): The work included in this subchapter was supported by the Flemish Government within the framework of the Policy Research Centre Well-Being, Public Health, and Family, with ref. nr. 3M180760.

No other conflicts of interest are declared.