KCE REPORT 272

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

MULTI-CRITERIA DECISION ANALYSIS FOR THE APPRAISAL OF MEDICAL NEEDS: A PILOT STUDY

KCE REPORT 272
HEALTH SERVICES RESEARCH

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

MULTI-CRITERIA DECISION ANALYSIS FOR THE APPRAISAL OF MEDICAL NEEDS: A PILOT STUDY

IRINA CLEEMPUT, STEPHAN DEVRIESE, WENDY CHRISTIAENS, LAURENCE KOHN

COLOPHON

Title:	Multi-criteria decision analysis for the appraisal of medical needs: a pilot study
Authors:	Irina Cleemput (KCE), Stephan Devriese (KCE), Wendy Christiaens (KCE), Laurence Kohn (KCE)
Project coordinator:	Nathalie Swartenbroekx (KCE)
Acting project coordinator:	Laurence Kohn (KCE)
External experts:	Francis Arickx (RIZIV – INAMI), Lode De Bot (Vlaams Patiëntenplatform vzw), Ri De Ridder (RIZIV – INAMI), Micky Fierens (Ligue des Usagers des Services de Santé asbl), Viviane Gendreike (INAMI – RIZIV), Viviane Knappenberg (INAMI – RIZIV), Isabelle Kniebs (Patienten Rat & Treff), Sophie Lannoy (Ligue des Usagers des Services de Santé asbl), Florence Leveque (INAMI – RIZIV), Greet Musch (FAGG – AFMPS), Maarten Plessers (Vlaams Patiëntenplatform vzw), Tinne Vandensande (Koning Boudewijnstichting), Ilse Weeghmans (Vlaams Patiëntenplatform vzw), Rene Westhovens (UZ Leuven)
External validators:	Bruno Flamion (Université de Namur), Marc Koopmanschap (Erasmus Universiteit Rotterdam, BMG, Nederland), Mireille Goetghebeur (University of Montreal, Canada)
Acknowledgements:	Jef Adriaenssens (KCE), Members of de commission for advice on temporary compensation for the costs a pharmaceutical product of the INAMI – RIZIV
Other reported interests:	All experts and stakeholders consulted within this report were selected because of their involvement in the topic of this study. Therefore, by definition, each of them might have a certain degree of conflict of interest to the main topic of this report
Infographics:	Julien Ligot (Fedopress)
Layout:	Filip Coppens (Smals), Sophie Vaes (KCE), Ine Verhulst (KCE)

Publication date:	28 June 2016
Domain:	Health Services Research (HSR)
MeSH:	Decision Making, Decision Support Techniques, Needs Assessment
NLM Classification:	WA 525
Language:	English
Format:	Adobe® PDF™ (A4)
Legal depot:	D/2016/10.273/68
ISSN:	2466-6459

How to refer to this document?

Cleemput I, Devriese S, Christiaens W, Kohn L. Multi-criteria decision analysis for the appraisal of medical needs: a pilot study. Health Services Research (HSR) Brussels: Belgian Health Care Knowledge Centre (KCE). 2016. KCE Reports 272. D/2016/10.273/68.

This document is available on the website of the Belgian Health Care Knowledge Centre.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
ALS	Amyotrophic lateral sclerosis
CAT	Computerized adaptive testing
CATT / CAIT	“Commissie voor advies in geval van tijdelijke tegemoetkoming voor het gebruik van een geneesmiddel” / “Commission d’avis en cas d’intervention temporaire pour l’utilisation d’un medicament”
CMD	College of Medical Directors
ETR	Early Temporary Reimbursement
FAMHP	Federal Agency for Medicines and Health Products
HTAi	Health Technology Assessment International
ICER	Incremental cost-effectiveness ratio
KBF	King Baudouin Foundation
LUSS	Ligue des Usagers des Services de Santé
MCDA	Multi-criteria decision analysis
NIH	National Institute of Health
PCIG	Patient and Citizen Interest Group
P.H.	Public Health
PROMIS	Patient Reported Outcomes Measurement Information System
PRT	Patienten Rat & Treff
S.A.	Social Affairs
VPP	Vlaams Patiëntenplatform

SCIENTIFIC REPORT

1 INTRODUCTION

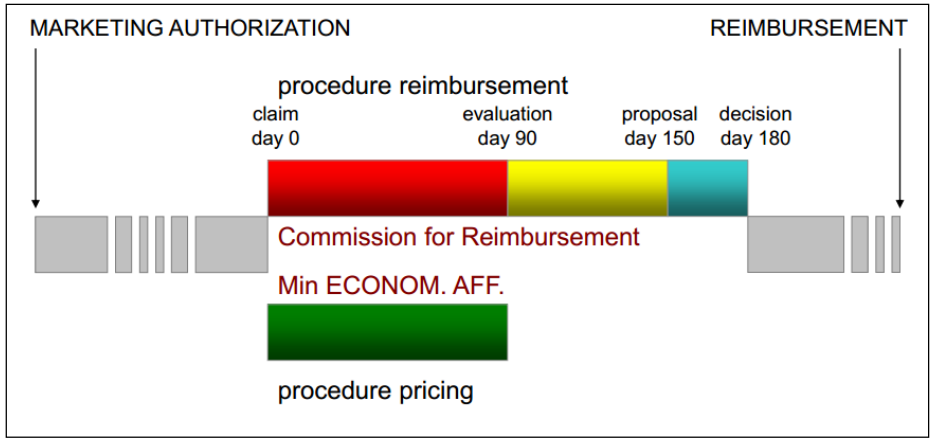
This scientific report provides more background material to the synthesis. It is not intended to be readable as a stand-alone product, but is an extension to the synthesis for readers who wish to have more details about specific elements.

1.1 Compassionate use and medical needs programmes as defined by law

The medical needs programme and the compassionate use programme are both meant to fill a gap for high medical needs. High medical needs are defined by law as **chronic**, **seriously debilitating or life-threatening** conditions for which **no satisfactory alternative authorized treatment** exists. Because of the time lag between the conclusions of the phase III clinical trials and the marketing authorization (MA) on the one hand and between the MA and the reimbursement decision (see Figure 1) on the other hand, the decision maker decided to install programmes to give quicker access to treatments for patients with a high medical need. The programmes regulate **the use** of medicinal products that have not yet received marketing authorization for a particular unmet medical need. They do **not** automatically imply **reimbursement** by the National Institute for Health and Disability Insurance (RIZIV/INAMI).

Coverage of the costs associated with products approved for the compassionate use or medical need programme is taken care of either by the pharmaceutical company or by the RIZIV/INAMI under particular conditions (described in paragraph 1.2). Approval of a compassionate use or medical need programme is the responsibility of the Federal Agency for Medicines and Health Products (FAMHP) in Belgium.

Figure 1 – Standard reimbursement decision procedure



Source: Veerle Van de Velde & Florence Levêque, RIZIV / INAMI, Coopami presentation. Reimbursement of medicines in Belgium. 22 October 2011 (http://www.coopami.org/en/countries/countries_partners/south_korea/projects/2011/pdf/2011112203.pdf)

The Law of 25 March 1965 installed the compassionate use and medical need programme in Belgium. The features of the two programmes are presented in Table 1.

Table 1 – Compassionate use versus medical needs programmes

	Compassionate use	Medical needs
Authorized?	No	Yes, but for (an)other indication(s)
MA application in progress?	Yes *or* ongoing clinical trials (with some evidence)	Yes *or* ongoing clinical trials (with some evidence)
Law	25 March 1965	
Royal Decree	RD of 14 December 2006, modified by the RD of 25 April 2014	
	Art.106	Art.108

MA: Marketing Authorization

Compassionate use programmes are meant for medicinal products for which no marketing authorization has been granted yet in Belgium. The product is not on the market in Belgium, but the MA application has been submitted or phase III clinical trials are ongoing. This implies that at least some evidence regarding the safety and efficacy of the product must already be available for the product to be eligible for the compassionate use or medical needs programme.

In contrasts to the compassionate use programme, the medical needs programme is meant for products that already have marketing authorization for one or more indications, however, not for the target indication. The target indication is one for which the medical need is high because it is a severe condition for which no satisfactory alternative treatment is available.

Key points

Compassionate use:

- **Drugs that are non-authorized in Belgium**
- **For patients with a chronically or seriously debilitating disease or whose disease is considered to be life-threatening, and who cannot be treated satisfactorily by an authorized medicinal product.**
- **The medicinal product concerned must either be the subject of an application for a MA or must be undergoing clinical trials**

Medical needs programme:

- **Products authorized in Belgium**
- **For chronical disease, disease with a serious impact or life threatening disease that cannot be treated satisfactorily by a product that is licensed for this indication (and commercially available) in Belgium**
- **Demand to obtain MA for the indication in question needs to be in process or clinical trials are ongoing in this indication or the indication has been obtained but the product is not commercially available**

1.2 Early temporary reimbursement for unmet medical needs

In 2014 a new law established the process for the early temporary reimbursement (ETR) of products targeting an unmet medical need (Law of 7 February 2014). The modalities of the ETR process were defined in the Royal Decree (RD) of 12 May 2014. The aim of the law was to give quicker access to promising innovative products for unmet medical needs (UMN) for which no MA has been obtained yet. While the compassionate use and medical need programmes were established to give approval for the use of these products under specific conditions, the unmet medical needs procedure regulates early temporary reimbursement to give patients quicker access to these products. Indeed, as long as these new innovative products are not covered (either by the company or the national health insurance), patients *de facto* often do not have access. National governments have to ensure the sustainability of the health system and therefore have to consider carefully what can and what cannot be reimbursed. This applies in particular, to orphan drugs, which are most often very expensive and unaffordable for individual patients.

In the past, early temporary reimbursement was possible in the following ways:

- A company could ask for the early reimbursement of products for which they claimed an added therapeutic value (i.e. class 1 drugs) or which had an orphan designation, and for which they had received a positive advice from the CHMP at EMAa.
- Individual patients could ask for reimbursement of their treatment costs by the Special Solidarity Fund (SSF) if their treatment was not reimbursed through the regular procedure. This applies for instance to products for which MA has not been granted yet. Reimbursement by the SSF is usually provided after the start of the treatment. Medical Need Program Art 56: individual demand, reimbursement after start of treatment, (MA not yet obtained, e.g. herceptin)
- The Minister can negotiate a contract with a manufacturer for innovative products that meet a societal or therapeutic need and have a clinical value, according to Article 56 §2 second 1994 of the health insurance law. The temporary reimbursement is in this case subject to conditions relating to, for instance, scientific research and reporting.[b] Article 81 and 81 bis of the Royal Decree of 21 december 2001 define the practical modalities.

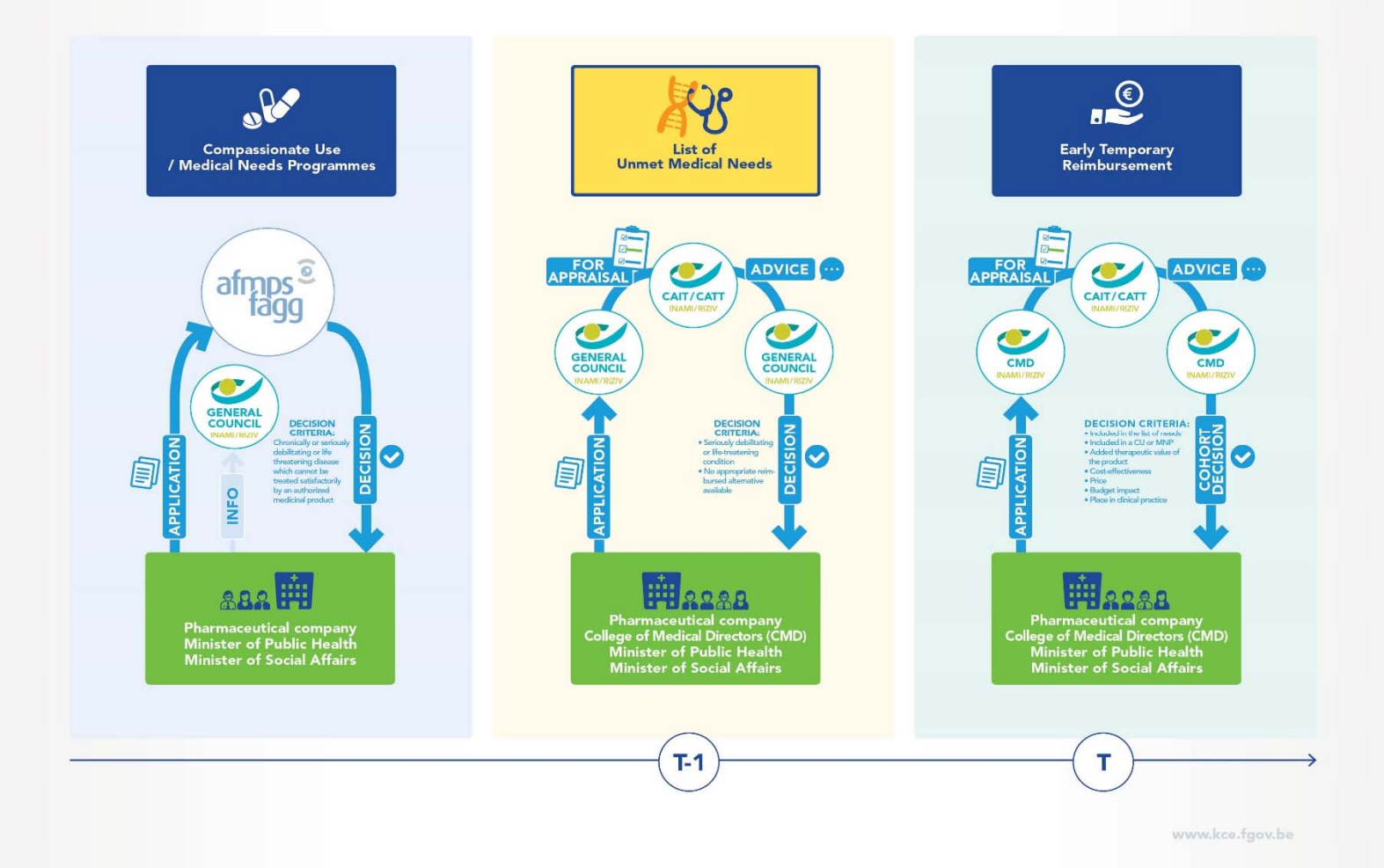
The unmet medical needs procedure is clearly distinct from the regular reimbursement procedure for pharmaceutical products. While in the regular procedure, the decision is taken by the Minister of Social Affairs and Health, after advice from the Drug Reimbursement Committee, the decision for early temporary reimbursement is taken by the College of Medical Directors (managing the Special Solidarity Fund). The Law therefore does not speak of “reimbursement” but of “cost compensation”. However, because ETR is the term commonly used in literature for these kinds of early access schemes, we will use this term in the remainder of the report.

The process of the cohort decisions is presented in Figure 2.

[a] 21 DECEMBER 2001. - Koninklijk besluit tot vaststelling van de procedures, termijnen en voorwaarden inzake de tegemoetkoming van de verplichte verzekering voor geneeskundige verzorging en uitkeringen in de kosten van farmaceutische specialiteiten, gewijzigd door KB 2007-11-20/35, art. 2, 038; Inwerkingtreding : 15-12-2007

[b] 14 JULI 1994. - Wet betreffende de verplichte verzekering voor geneeskundige verzorging en uitkeringen gecoördineerd op 14 juli 1994, art 56, § 2.

Figure 2 – Early temporary reimbursement process



AFMPS / FAGG = Federal Agency for Medicines and Health Products
CMD = College of Medical Directors
CAIT / CATT = Commission d'avis en cas d'intervention temporaire pour l'utilisation d'un médicament / Commissie voor advies in geval van tijdelijke tegemoetkoming voor het gebruik van een geneesmiddel
CU = compassionate use programme
MNP = medical needs programme

The first step in the process is the construction of a **list of (unmet medical) needs**. This happens in year T-1. Requests to put a condition on the list of needs can be submitted by the pharmaceutical companies, the College of Medical Directors, the Minister of Health and the Minister of Social Affairs. Requests for the list of year T have to be submitted by May 15th of the year T-1. The **General Council** of the RIZIV / INAMI is responsible for creating the list.[c] By October 3, T-1 the General Council produces the list of needs for year T, based on the advice of the "Commission for advice on temporary compensation for the costs a pharmaceutical product" ("Commissie voor advies in geval van tijdelijke tegemoetkoming voor het gebruik van een geneesmiddel" (CATT) / "Commission d'avis en cas d'intervention temporaire pour l'utilisation d'un medicament" (CAIT)) and the College of Medical Directors. The list should allow the General Council to roughly estimate the budget possibly needed to fund the possible future cohort decisions of the College of Medical Directors. The development of a methodology for defining relative priorities on this list is the focus of the current study.

In year T, pharmaceutical companies, the Minister of Health, the Minister of Social Affairs and the College of Medical Directors can submit a request for the early temporary reimbursement (ETR) of an intervention to the College of Medical Directors. The College of Medical Directors is being advised by the CAIT / CATT regarding the early temporary reimbursement (ETR) for a specified cohort of patients.

To be eligible for ETR, a product has to satisfy **all** of the following conditions:

- it targets an unmet medical need,
- it is used to treat a seriously debilitating or life-threatening condition,
- there is no appropriate reimbursed therapeutic alternative,
- the product is included in a programme for compassionate use or medical need (as described in paragraph 1.1),
- the product responds to an unmet medical need included in the list of unmet medical needs.

Therefore, the **CAIT / CATT** assesses whether the product targets an unmet medical need, based on the list of unmet medical needs defined by the General Council of the RIZIV / INAMI. It also makes an appraisal of the usual criteria used by the Drug Reimbursement Committee during the regular procedure: the products' added value, its cost-effectiveness, its price, its budget impact, its place in daily clinical practice.[d]

The **cohort decisions** of the **College of Medical Directors** specify:

- the conditions for and level of reimbursement,
- the cohort of patients eligible for ETR (i.e. inclusion- and exclusion criteria) and
- the budget needed for covering the product.

Cohort decisions are defined by law[e] as decisions **restricted in time** and based on available economic and medical data. The budget is defined yearly.

The agent who submits a request for a cohort decision is responsible for the execution of the programme, the designation of a responsible physician for handling the requests to be included in the programme, administration of a register of included patients and registration of unexpected adverse events. When the request has been submitted by the Minister of Social Affairs and Health, the RIZIV / INAMI is responsible for the organization and execution of the programme.

When an individual patient submits to the College a request for compensating the costs of a new pharmaceutical product that is not reimbursed through the regular procedure, the College will first consider whether a cohort decision exists for that patients' indication and whether the patient satisfies the eligibility criteria. Patients who are included in the ETR programme have to give their informed consent.

---

c Wet betreffende de verplichte verzekering voor geneeskundige verzorging en uitkeringen, gecoördineerd op 14 juli 1994, Art 25

d Koninklijk besluit van 12 mei 2014 tot uitvoering van de artikelen 25 en volgende van de wet betreffende de verplichte verzekering voor geneeskundige verzorging en uitkeringen, Art 8

e Art. 25quater/1, §1 of the Wet betreffende de verplichte verzekering voor geneeskundige verzorging en uitkeringen, gecoördineerd op 14 juli 1994

Key points

- **The unmet medical needs procedures foresees in the possibility to obtain early temporary reimbusement (ETR) for products which treat an unmet medical need and which are included in a medical need or compassionate use programme approved by the FAMHP.**
- **A cohort decision implies ETR for an unmet medical need.**
- **Requests for a cohort decision can be submitted by a pharmaceutical company, the Minister of Health or the College of Medical Directors at the RIZIV / INAMI.**
- **The cohort decision is taken by the College of Medical Directors, following the advice of the 'commission for advice on temporary compensation for the cost of a pharmaceutical product' (CAIT / CATT).**
- **The CAIT / CATT assesses whether the product targets an unmet medical need, based on the list of unmet medical needs defined by the General Council of the RIZIV / INAMI.**

1.3 Reflections on the legal framework for unmet medical needs

There is an anomaly in the law in that the budget estimate is based on a list of unmet medical needs, while the treatments for those needs have, in theory, not been defined yet. Moreover, the treatments seem to be limited to pharmaceutical treatments. It is therefore assumed that the law intends to define the list of pharmaceutical treatments meeting an unmet need, rather than a list of needs as such. This is in contrast with the spirit of a demand driven, rather than a supply driven system. If the healthcare system is to become more demand driven, i.e. driven by the real needs rather than by the new medicines launched on the market, it is necessary to define a list of unmet medical needs independent of the pharmaceuticals that are in the pipeline. The needs of patients, for that matter, often cannot be met by new pharmaceutical treatments and hence the focus should not solely be on new pharmaceuticals. For instance, the conclusion of the citizen labs organised by the King Baudouin Foundation was that citizens value psychosocial well-being and quality of life of the patient's environment very highly. These needs might be solved better by more efficient support systems than by a new drug.

Hence, the list should be a list of conditions. Indications of treatments in the pipeline could figure on the list, but it should not be an inclusion criterion.

Besides the stakeholders currently allowed to submit proposals, also patients, healthcare providers and other professionals should be allowed to propose the inclusion of conditions in the list of unmet medical needs, as they are often well-placed to identify real needs.

The focus on conditions rather than on interventions under development implies, however, the impossibility to estimate beforehand the required budget for next year's cohort decisions. This does not mean that it is not possible to define a budget for innovations coming to the market, independent of what is already in the pipeline. This could be pharmaceutical innovations but equally so surgical innovations, organisational innovations, (supportive) care innovations, etc. There is no reason to limit this budget to pharmaceutical interventions only.

Key points

- **The list of unmet medical needs is used to allow the General Council of the RIZIV / INAMI to estimate the budget required for financing the unmet medical needs in the following year.**
- **This is confusing, as it suggests that the list contains products in the industry's pipeline rather than diseases (quite likely so, as the industry can submit proposals for the unmet medical needs list). In contrast, one would expect the list to contain names of seriously debilitating or life-threatening health conditions for which no alternative treatment exists and for which a treatment is urgently needed, independent of what is being developed currently.**
- **To move from a supply-driven health care system to a demand-driven health care system, as was one of the purposes of the UMN law, it is necessary to start from diseases instead of products.**

- **The list of stakeholders who can submit proposals for inclusion on the unmet medical needs list should be expanded to include patients, healthcare providers and other professionals.**
- **Budgets for innovative treatments for high healthcare needs should not be limited to pharmaceutical innovations but be expanded to other types of innovations in the healthcare sector.**

1.4 Unmet medical need versus therapeutic and societal need

The law of 7 February 2014 and the Royal Decree of 12 May 2014 consistently use the term “unmet medical needs”, which is defined as a pharmaceutical product for the treatment of a severe or life threatening condition for which **no reimbursed alternative** treatment is available. In the regulation of the European Commission (EC 726/2004), used by the FAMHP for judging the eligibility of a product for the compassionate use or medical needs programme, the definition of an unmet need is, however, more restrictive: it only refers to chronically or seriously debilitating diseases or diseases considered to be life threatening and that cannot be treated satisfactorily **by an authorised medicinal product**.

This implies that the indication for which a product that has received approval from the FAMHP for a medical needs programme does not necessarily have to figure high on the list of unmet medical needs. If the indication under consideration can be managed satisfactory by a non-medical treatment, that indication can be judged as a low medical need. This is legitimate and should not be regarded as problematic.

We would like to remark that the term “unmet medical need” might be misleading and restrictive. We prefer using the term “therapeutic need” instead of unmet medical need for two main reasons.(1) A need is only really “unmet” if there is no alternative at all for helping the patient. However, rarely patients are not helped at all, especially in the case of severe conditions (e.g. symptomatic treatment, supportive care). As soon as something is done for the patient, that activity should be considered as the alternative with which the new treatment or clinical management strategy should be compared.(2) Related to this, we would argue that need is not a categorical feature which is “present” or “absent”, but that there are gradations of the extent to which needs are met. Some needs will be met more than others, but will still be important needs and therefore deserve attention.

The term “therapeutic need” refers to the need for a better treatment than the treatment currently reimbursed (or the best supportive care currently reimbursed). This concept allows to define gradations of need: the more effective the current treatment, the lower will be the therapeutic need, even if it concerns a very severe disease like diabetes. By defining need in this way, it will be possible to identify those diseases which should get priority in terms of investment decisions. Also conditions which are partially met can get into the list of needs, which seems more in line with the spirit of the law than including only conditions which are unmet. The objective of the legislator was presumably to target the high medical needs (i.e. severe diseases), for which no alternative treatment exists yet or for which the alternative treatment is insufficiently effective.

An additional remark with respect to the concept of unmet medical need is that it misses out on the concept of societal need. For example, a product can serve a high societal need if it can substantially reduce disease-related public expenditures, without jeopardizing the outcomes for the patient. This happens, for instance, when an effective authorized treatment is available but at a much higher cost than an equally effective much cheaper alternative. If the list of unmet medical needs is meant to guide public money to conditions for which there is a high need, it might be worthwhile to also consider creating a list of societal needs. It is better to separate the list of conditions with a high societal need from the list of conditions with high therapeutic need, because the objective of new interventions tackling these conditions will also be different: for high therapeutic needs the objective is to improve patient outcomes, while for high societal needs the main objective is to reduce public expenditures. It is clear that it would not be possible within the current legal framework to take cohort decisions for products targeting a high societal need, because for cohort decisions also the requirements for the medical needs programme as defined by the FAMHP need to be met.

In this study, we did not bind ourselves to the current legal framework for the unmet medical needs programme, but followed the underlying philosophy of the programme, being to identify the health(care)needs of patients and society in order to set the priorities right. The definition of unmet medical need used in this study is therefore larger than the definition used in the unmet medical needs programme and includes societal need besides therapeutic need.

Key points

- **Differences between the definition of "medical needs" used by the FAMHP and the definition of "unmet medical needs" used by the RIZIV / INAMI imply that indications accepted for a medical needs programme do not necessarily have to figure high on the list of unmet medical needs.**
- **Therapeutic need is a better concept than unmet medical need, because it includes conditions whose needs are partially met but remain severe despite the currently available treatment or care.**
- **It makes more sense to make a rank ordered list of therapeutic needs than a list of exclusively unmet medical needs as for most severe conditions some treatment or care is already provided.**
- **Creating a list of societal needs should be considered. For conditions figuring high on this list, the objective is to reduce public expenditures.**

2 OBJECTIVES OF THIS STUDY

The general objective of this study is to develop and assess the applicability of a multi-criteria decision analysis (MCDA) approach for the appraisal and ranking therapeutic and societal needs in healthcare, taking into account the public preferences with respect to the relative importance of appraisal criteria.

The starting point will be the multi-criteria decision analysis (MCDA) approach presented in a previous KCE study.[1] MCDA is used for complex decisions that involve multiple criteria. It makes the criteria explicit and takes the relative importance of these criteria into account in a transparent manner.

The main focus of the current study will be on the development of tools to facilitate the application of the MCDA approach; i.e. tools to summarize the evidence regarding a health condition, to score the conditions on a pre-determined scale, to aggregate the weighted scores and to construct the list of needs. Also the optimal process of using the MCDA approach for making the appraisals will be examined.

This study will not assess or analyse the pros and cons of the different forms and approaches of MCDA. MCDA is a set of approaches, rather than a specific one-solution methodology. Rather, it develops and assesses the applicability and acceptance of *one* MCDA approach for Belgium. Also, the study will not produce a list of therapeutic and societal needs that can be used in real life. The lists resulting from the pilot study will be produced for research purposes only.

The study will also not provide guidance on how to score particular evidence on the scales. For example, no guidance will be given on how to score a 2 year life expectancy loss due to illness on a 4-point Likert scale. Nevertheless, we will give guidance on how to deal with factors such as uncertainty in the MCDA.

Key points

- **The objective of this study is to develop and test a multi-criteria decision analysis approach for ranking therapeutic and societal needs in healthcare. Multi-criteria decision analysis makes the criteria for judging needs explicit and uses explicit weights that reflect the relative importance of these criteria for this judgment.**
- **This study does not assess or analyse the pros and cons of different types of MCDA.**
- **This study does not produce a list of unmet medical needs. It is a methodological study, developing a possible approach to creating such list.**

3 METHODS

3.1 Building on previous research

In 2010, KCE published a framework to improve accountability for reasonableness in reimbursement decisions, thereby focussing on feasibility, transparency and consistency (Table 2).[2]

Table 2 – Key questions and possible criteria for a drug reimbursement appraisal process (MCDA framework)

Decision	Question	Possible criteria
Medical, therapeutic and societal need	Does the product target a medical, therapeutic and societal need?	**Medical need:** Life-threatening / non-life threatening condition; Severe / mild symptoms; Poor initial health state; Baseline health level **Therapeutic need:** Effective alternative treatments available / not available **Societal need:** High / Low prevalence; Health inequality; disease-related public expenditures
Preparedness to pay out of public resources for a treatment	Are we, as a society, in principle, prepared to pay for a treatment that will improve this indication out of public resources?	Own responsibility Life-style related condition
Preparedness to pay out of public resources for the treatment under consideration	Are we, as a society, prepared to pay for this particular treatment, given that we in general would be prepared to pay for a treatment for this indication?	Safety and efficacy of the treatment compared to the alternative treatment(s); Curative, symptomatic, preventive; Therapeutic value; Significance of health gains
Preparedness to pay more	Given that we, as a society, are prepared to pay for this treatment out of public resources, are we prepared to pay more for this treatment than for the best alternative treatment?	Added therapeutic value; Potentially induced savings elsewhere in the health care sector; Quality and uncertainty of the evidence ; Acceptability of co-payments and/or supplements; Rarity of disease
Willingness to pay (price and reimbursement basis)	How much more are we willing to pay out of public resources for this particular treatment?	Added therapeutic value; Budget impact / ability to pay; Cost-effectiveness ratio; Medical, therapeutic and societal need; Quality and uncertainty of evidence; Limits to cost sharing

Source: le Polain et al. 2010[2]

In 2012; we published the report of a qualitative study about the acceptability and feasibility of different models for citizen- and patient involvement in health policy according to the different stakeholders in Belgian health policy.[3] The study showed that all stakeholders considered involving citizens and patients to be important. They would see them in a consultative role. Citizens would be consulted for more strategic decisions, e.g. about healthcare priorities, whereas patients would be consulted more in the context of operational decisions, e.g. drug reimbursement decisions. Patients would have a role as "experts by experience", whereas citizens would play a role as "payers for healthcare and potential patients". Because it would be impossible to consult citizens about important decisions in health policy, we searched for alternative ways to set up a one-time consultation that could deliver results that would be useful for many decisions, at least for a certain period of time. This was the topic of a next report.

In 2013-2014, KCE examined how such a framework could be put into practice, keeping in mind the purpose of increasing transparency and consistency of decision making processes and adding the important requirement of relevance of decision criteria. We focused on the aspect of assessing and appraising need (both therapeutic and societal need) and added value of new interventions. For these three domains, we identified relevant appraisal criteria and we measured the relative importance of these criteria for the appraisal according to the general public. A large population survey was performed, using discrete choice experiments for each domain (see KCE-report 234 on incorporating societal preferences in health care reimbursement decisions).[1] The criteria identified as being relevant for each domain and their weights according to the general public are presented in Table 3.

Table 3 – Criteria for defining therapeutic need, societal need and added value with their weights

Therapeutic need	Weight
• Impact of the condition on **quality of life**, given current treatments available	0.43
• Impact of the condition on **life expectancy**, given current treatments available	0.14
• **Inconvenience** of current treatment	0.43
Societal need	**Weight**
• Condition-related **public expenditures per patient**	0.65
• **Frequency** of the condition	0.35
Added value	**Weight**
• Change in quality of life	0.37
• Change in prevalence	0.36
• Change in life expectancy	0.14
• Impact on public expenditures	0.07
• Impact on treatment discomfort	0.06

The criteria were defined as follows:

- Impact of the condition on quality of life, given current treatments available refers to the extent to which the disease has an impact on the five dimensions of the EQ-5D, which is a frequently used quality of life instrument in research: mobility, self-care, usual activities, pain/discomfort, anxiety/depression. Mobility refers to the ability to walk about, self-care to the extent to which patients are able to wash and dress self, usual activities to the extent to which they are able to participate in social activities and go to work or school.
- Impact of the condition on life expectancy refers to the extent to the number of years people lose due to their disease, as compared with patients of the same age without the disease.
- Inconvenience of current treatment refers to the inconvenience caused by for instance the frequency of use (e.g. taking a drug once or more times a day), the administration route (e.g. syringes, oral drugs, via perfusion, by the patient him- or herself or by someone else), the place of administration (in hospital, at home, in a doctor's cabinet) and the side-effects of/morbidity caused by the current treatment. It should be noted that inconvenience has a broad meaning. It does not relate only to medication. For example, dialysis treatment is usually considered very inconvenient, be it at home or at the hospital. This type of inconvenience is clearly of a different order than the inconvenience of having to take a pill twice a day instead of once a day. This may explain the high weight of 'inconvenience' in our population survey.
- The frequency of disease refers to the prevalence or, for acute diseases, the incidence of the disease.
- The disease-related public expenditures refer to the total public expenditures per patient with the disease, including health care expenditures and invalidity benefits.

In that study, we also presented the principles of a multi-criteria decision analysis (MCDA) approach to assess the therapeutic and societal need for a better treatment than the currently available treatment and to assess the added value of new treatments. The current study focuses on the determination of therapeutic and societal need, as defining unmet medical needs is independent of interventions under development. The MCDA approach is explained in more detail in paragraph 3.2. In short, MCDA takes the different criteria that matter for a decision into account and weights these according to their relative importance. Two components are essential: the performance of diseases and interventions on the relevant criteria and the weights of the criteria. Weights reflect preferences *between criteria*, scores reflect priorities or preferences *within a criterion*.

Besides previous KCE research, research that is relevant for the current study has been performed by the King Baudouin Foundation (KBF). In 2015, the KBF performed citizen consultations about values and criteria that should guide decisions about the reimbursement of health interventions.

A selected group of 32 citizens (16 Dutch speaking, 16 French speaking, 16 women and 16 men of different socio-economic backgrounds and ages), discussed during the course of three week-ends which criteria they considered important for taking reimbursement decisions and why.

Additional criteria, compared to the ones identified by KCE, emerged from these consultations. Participants pleaded for a holistic patient-oriented health care approach (integrated care), that defines improvement of quality of life as the main priority. Respect for the sick person's autonomy and attention for the interests of his or her environment are considered very important. They emphasize the importance of a coherent healthcare policy that puts prevention high on the agenda. Public health insurance, based on the principles of solidarity and the right to health care are considered cornerstones of the health care system.

In addition to the conclusion that improving quality of life should be a more important policy objective than prolonging life expectancy, the participants expressed the wish to have a more demand driven health care system, in contrast to the current supply-driven system. Inefficient use of health care resources should be reduced. Effectiveness of interventions should be proven. Decisions should be more transparent and flexible (revisable in case of new evidence). Prevention and financial sustainability are key factors.

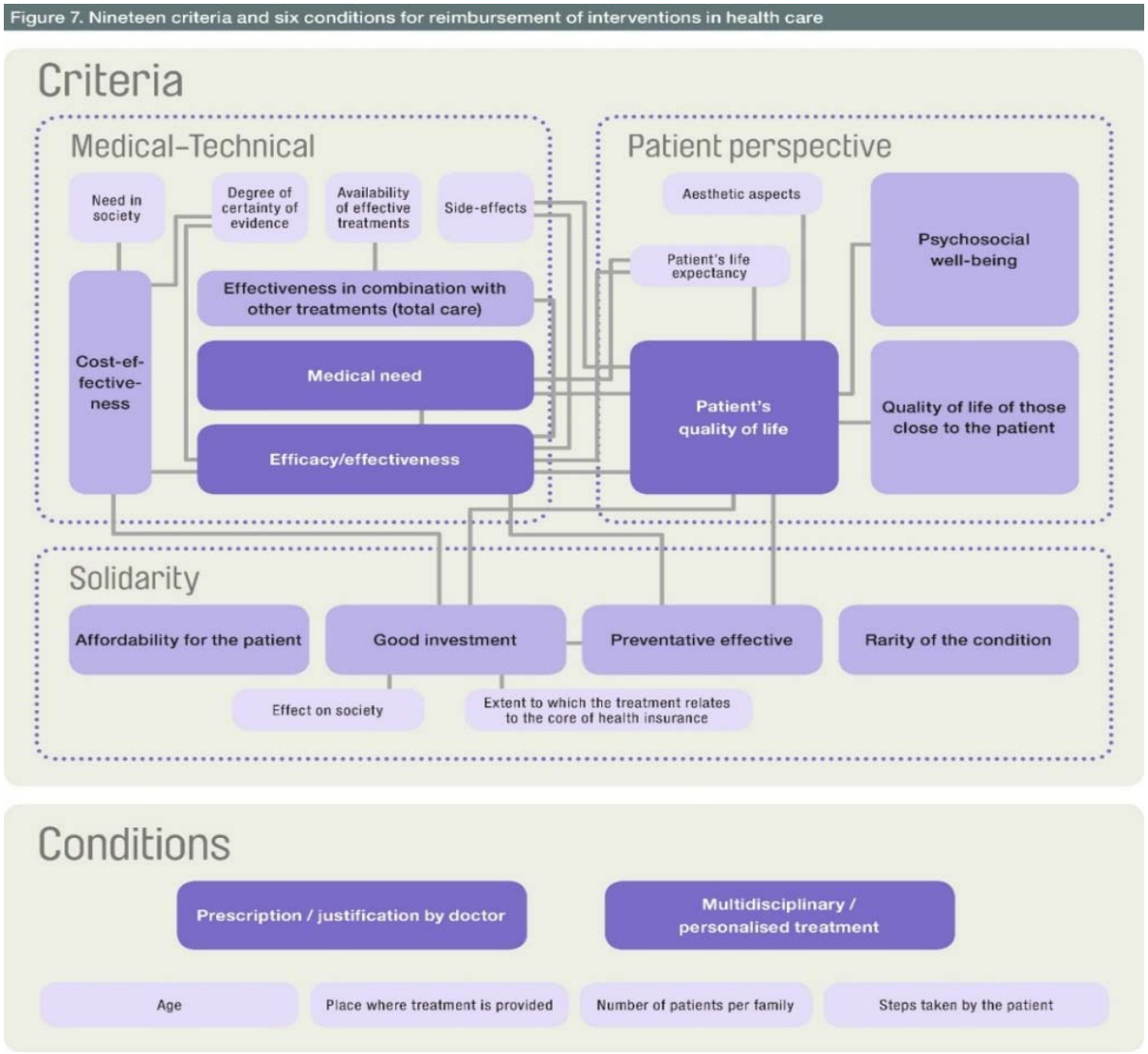
Personal responsibility for own health is linked to social responsibility to maintain solidarity and equity by the participants. Individual choices for a healthy or unhealthy lifestyle have an impact on the society but it is undesirable, according to the participants, to retain interventions from specific population groups based on the argument of personal responsibility. Nevertheless individual responsibility could be used as a trigger for prevention.

Age should not be a criterion for reimbursement, according to the participants.

After three week-ends, the participants listed 19 criteria and 6 conditions for reimbursement in health care (see Figure 3). Criteria relate to reimbursement decisions for products and services in general, whereas conditions refer to decisions regarding the reimbursement for particular patients in particular situations. The intensity of the colour Figure 3 relates to the relative importance attributed to the criteria and conditions by the citizens' panel. The darker the colour, the more important the criterion or condition was considered.

Convenience of treatment was not identified as an important criterion by the panel. However, it is a criterion for reimbursement according to the Belgian law.[4]

Figure 3 – Criteria and conditions for reimbursement identified by the citizen panel



Source: Raeymaekers, 2015 [5]

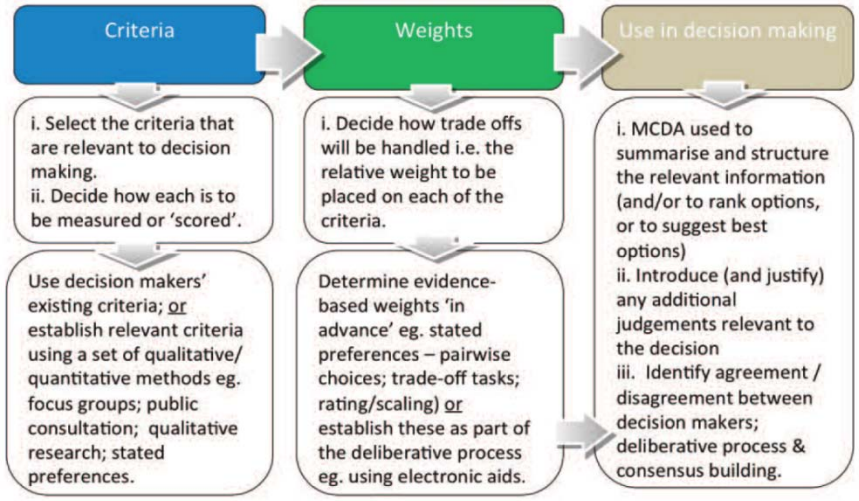
3.2 Multi-criteria decision analysis

MCDA can be defined as *"a set of methods and approaches to aid decision-making, where decisions are based on more than one criterion, which make explicit the impact on the decision of all the criteria applied and the relative importance attached to them."*[6].

The first two blocks of Figure 4 (criteria and weights) present key steps for *developing* an MCDA framework, the third block explains how to *use* MCDA in decision making.

MCDA always starts with the definition of the decision for which the MCDA is performed. The decisions can vary from assessing the benefit-risk of a new drug, to defining unmet needs, defining research priorities, or deciding on the reimbursement of a particular product or service.

Figure 4 – Steps to be taken in the development of an MCDA framework



Source: Devlin et al., 2011[6]

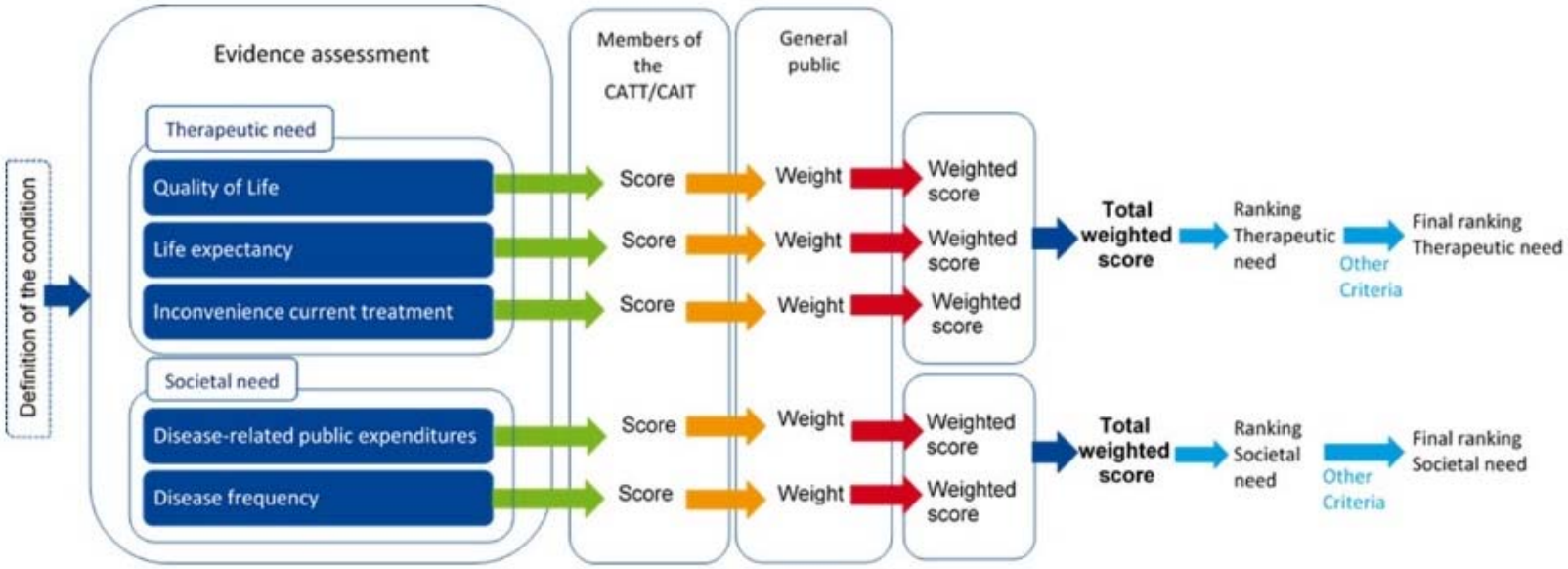
The actual use of MCDA in the decision making process regarding the appraisal of needs in various diseases is presented in Figure 5.

- First the disease or condition as well as the patient population under consideration needs to be well-defined
- Then an evidence or HTA report needs to be prepared, critically assessing and summarizing the existing evidence for the disease with respect to the various relevant criteria for the appraisal. These criteria must be defined a priori and correspond, in our case, to the criteria included in the population survey performed in 2014: impact of the disease on quality of life given current treatment, impact of the disease on life expectancy given current treatment, inconvenience of current treatment, impact of the disease on public expenditures and prevalence of the disease.
- This evidence is subsequently used by the members of the CATT / CAIT to assign scores to the criteria on a predefined scale (e.g. from 0 to 3), representing the "performance" of the disease on each of the criteria. For example, a score of 0 could be assigned to impact of the disease on life expectancy because the evidence did not demonstrate an impact of the disease on mortality. The individual scores have to be discussed amongst the commission members. Uncertainty or quality of the evidence upon which the scores are based can be included in the discussion and have repercussions on the scores.
- The scores are then weighted, using the criteria weights obtained from the general public (Figure 5).
- The weighted scores of the criteria relevant for assessing therapeutic need (quality of life, life expectancy and inconvenience of current treatment) are summed to obtain a total weighted score for the therapeutic in the disease under consideration; the weighted scores of the criteria relevant for assessing societal need (public expenditures and prevalence) are summed to obtain a total weighted score for the societal need related to that same disease.

  When this appraisal process is repeated for several diseases, the total weighted scores of these diseases can be ranked, giving a first idea of the relative level of need in each these diseases.

- After this process, the commission sneeds to consider the relevant criteria that have not yet been assessed as part of the MCDA (e.g. social vulnerability of the patients, impact of the disease on the quality of life of patients' acquaintances.

Figure 5 – MCDA process in decision making



MCDA should be considered as a *support* to decision-making, helping to structure complex decisions and deliberative processes that involve multiple criteria. MCDA could facilitate transparent and consistent decisions. It should not be seen as a prescriptive tool but rather as a tool to exercise the judgments made during multi-criteria decisions. It is a misconception that MCDA can only be purely quantitative. Inputs into the MCDA might be qualitative. For example, the introduction and justification of additional judgments relevant to the decision (point ii under "decision making" in Figure 4) is also part of the MCDA. The crucial feature of MCDA is that it makes qualitative assessments explicit and transparent. Deliberation can be used to define the inputs for MCDA and is, as such not in contradiction with MCDA principles. Hence, MCDA is not a mechanistic decision tool. As Klein stated in 1993: "*what really matters is how that debate is structured: how far it promotes reasoned, informed, and open argument, drawing on a variety of perspectives and involving a plurality of interests.*"[7]

The International Society of Pharmacoeconomics and Outcomes Research (ISPOR) has established an MCDA Task Force that is working on 'Emerging Good Practice Recommendations' for MCDA. The Task Force developed a checklist for MCDA, consisting of 7 items with 21 sub-items.[8]. The items and sub-items represent the 7 steps to be taken when developing and using MCDA in decision making (Table 4). They give a good guidance on how to start the MCDA process. We will follow this guidance to define the scope and process of the current study (see paragraph 3.4).

Table 4 – ISPOR MCDA Task Force recommendations for good MCDA practice

MCDA step		Recommendation
1. Decision problem	Definition	Develop a clear description of decision problem including the objectives, the stakeholders, and alternatives, and whether the objective is to value or rank alternatives.
	Validation	Validate the decision problem with decision makers and clinical experts
2. Selecting criteria	Develop	Report the sources used to identify criteria, the long list of criteria identified, and the rationale for excluding criteria.
	Report	Report the final criteria list, including definitions and measurement scales.
	Validation	Validate the criteria with stakeholders and against MCDA requirements (complete, non-redundant, non-overlapping, preferentially independent).
3. Measuring performance	Methods	Report and justify the sources used to measure performance
	Reporting	Report the performance matrix, showing performance of alternatives against criteria
4. Scoring and weighting	MCDA model	Report the MCDA model (value measurement, outranking, reference), and justify in reference to the decision problem and stakeholder preferences.
	Scoring/weighting methods	Justify the methods used for scoring and weighting with reference to the objectives of the analysis and the stakeholders involved in the analysis.
	Reporting	Report the values of scores and weights
	Validation	Validate the meaning of scores and weights with stakeholders
5. Aggregation	Reporting	Report the aggregation function used in the analysis
	Validation	Validate with stakeholders that the aggregate scores reflect how they expected their scores and weights to be used
6. Dealing with uncertainty	Report	Report and justify variation (ranges, distribution) in parameter inputs, including sources, and list all assumptions made (including structural assumption).
	Characterizing uncertainty	Consider the effects on the results of parameter uncertainty and assumptions.
	Uncertainty methods	At a minimum deterministic sensitivity analysis should be performed. Justify the method adopted to explore uncertainty with reference to the differences between techniques
	Characterizing heterogeneity	Consider the effects on results of variation in scores and weights and performance measures between sub-groups.
7. Interpretation	Robustness	The findings should be interpreted in light of the results of the validation undertaken and the results of the analysis of the impacts of uncertainty and heterogeneity.
	Meaning	The finding should be interpreted in light of the meaning of the particular scores and weights employed.

Source: Marsh, K., et al., 2015[8]

3.3 MCDA tools in literature

In this paragraph, we describe two multi-criteria decision analysis tools that were developed for other purposes than ours, but there are overlaps and both tools have inspiring elements for our own tool. The EVIDEM and the Evidence to Decision tool are both the result of collaborative efforts of research groups around the world.[9]

3.3.1 EVIDEM

The EVIDEM framework is a MCDA framework designed to evaluate *interventions* and facilitate their prioritization. The framework defines a comprehensive set of criteria organized in pragmatic tools. It is the result of the work of the EVIDEM Collaboration, an independent non-profit organization run by an international Board of Directors.

The EVIDEM framework consists of a "MCDA Core Model" and a "Contextual tool". The MCDA Core Model consists of operationalisable criteria for which the high and low ends of the scale can be defined a priori; i.e. there is agreement on which end of the scale should get priority. For example, severe diseases should get priority over mild diseases, diseases with many unmet needs should get priority over diseases with few unmet needs. Most of the criteria included in the Core Model include sub-criteria.[9]

The Contextual tool includes seven generic criteria/themes, with a number sub-criteria. The contextual criteria can be used as qualitative criteria for qualitative consideration during the decision making process or can be moved to the Core Model if they are operationalized. The EVIDEM framework is flexible in this respect and allows users to define the Core Model and Tool in a way that is most relevant to their setting. The Contextual tool encompasses normative contextual criteria such as mandate and scope of healthcare system, prioritary populations and access to healthcare plans, environmental impact and feasibility contextual criteria, such as affordability and system capacity.

The complete list of criteria with an explanation is available on the web-site of EVIDEM (https://www.evidem.org/components-decision.php).

Besides a framework, EVIDEM has also developed instruments for synthesizing evidence, both for the executive level (highly synthesized evidence) and for by-criterion evidence reports (detailed evidence), an instrument for assessing quality of evidence, tools for deriving criteria weights from decision makers using different techniques, and a software tool for analyzing and presenting the data. Recently, also a package for prioritizing research questions and software for analyzing and presenting the data for research question prioritization have been developed.

3.3.2 EtD: Evidence to Decision Framework

The Evidence to Decision framework (EtD) is a product of the DECIDE Collaboration. DECIDE is a 5-year project from January 2011 to December 2015 co-funded by the European Commission under the Seventh Framework Programme. The project's objective was "to improve the dissemination of evidence-based recommendations by building on the work of the GRADE Working Group to develop and evaluate methods that address the targeted dissemination of guidelines. GRADE is a systematic approach towards assessing and communicating the quality of evidence and the strength of recommendations. It has been developed to address the weaknesses of other grading systems and is now widely used internationally". (http://www.decide-collaboration.eu/tools-decision-making-and-dissemination-under-development)

Within the project, the EtD framework has been developed to help decision makers to move from evidence to decisions. The types of decisions can vary from making clinical practice recommendations to making reimbursement decisions. It is clear, however, that the origins of the framework lie within the domain of guideline development.

The similarities and differences between EVIDEM and EtD are interesting. Both frameworks apply a multi-criteria decision approach, and promote the use of scientific evidence as the basis of the decision making process. However, the emphasis on qualitative versus quantitative approaches in the actual implementation of the MCDA differs between the frameworks. The EVIDEM framework starts from the quantitative assessment and scoring of interventions, followed by quantitative weighting of the criteria scores and finally (qualitative) discussion. In the EtD framework, the judgement on the performance of an intervention on a particular criteria results from a discussion amongst panel members/decision makers. The EtD framework

uses a more qualitative approach than the EVIDEM framework, in that it does not calculate weighted total scores but presumes discussion on every individual judgement for each criterion. The EtD framework is described as meant for helping to structure discussions and identify reasons for disagreements, rather than for prioritizing interventions.

Another interesting difference between the EVIDEM framework and the EtD is that the EVIDEM framework considers "quality of the evidence" as a separate criterion in the decision making process, whereas the EtD takes the quality of the evidence into account in the judgments related to the criteria. The latter seems more appropriate, as the quality of the evidence might vary between criteria.

Like the EVIDEM framework, the EtD provides tools for summarizing the best available research evidence for each relevant decision criterion to inform judgments about relevant decision criteria. EtD offers a full interactive tool that can be accessed through the internet. The interactive summary of findings tables present the key messages from a systematic literature review in concise format: the most important outcomes of the interventions, the size of these effects and the certainty of evidence.

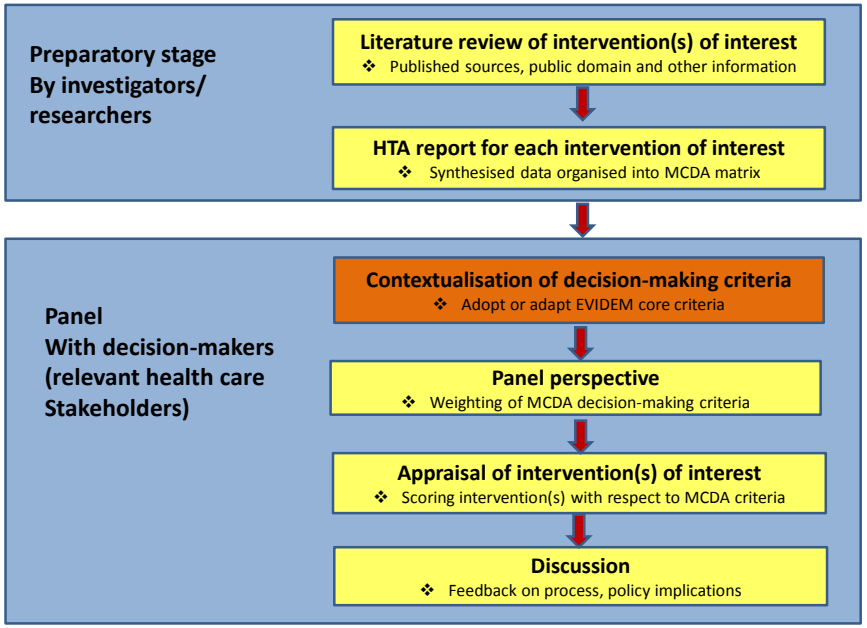
3.3.3 *Examples of applications of MCDA in coverage decision making*

While a number of countries are using MCDA in their reimbursement decision making processes, few have been described in literature. The application of MCDA in reimbursement procedures in Lombardy (Italy) and in Colombia will be described in a book that is currently in press. Canada pilot-tested a MCDA approach for the appraisal of medicines for coverage decisions.[10] Besides these three countries, the EVIDEM collaboration mentions on its website that also Brazil, Chile, China, Germany, Indonesia, Nepal, Romania, South-Africa and the United States of America are using or exploring the opportunity of using MCDA for coverage decision making.

3.3.3.1 *Colombia*

Colombia has piloted the use of a MCDA approach in reimbursement decision procedures in 2012-2013. The system tested works as shown in Figure 6.

Figure 6 – MCDA process applied in Colombia



Source: adapted from Goetghebeur et al (2012)

First, an HTA report on the interventions of interest is produced. Then, the criteria used in the EVIDEM framework are critically assessed and adapted to the local situation of Colombia. Fifteen criteria are included in the Colombian MCDA:

- Disease severity
- Size of population affected by the disease
- Improvement of efficacy/effectiveness
- Current clinical guidelines applicable in Colombia
- Type of medical service (clinical benefit)
- Budget impact on health plan

- Improvement of safety and tolerability
- Public health interest
- Improvement of patient-reported outcomes
- Current intervention limitations
- Attention to vulnerable groups of population
- Cost-effectiveness of intervention
- Completeness and consistency of reporting evidence
- Relevance and validity of evidence
- Attention to differential needs for health/health care

The criteria are then weighted by a panel of decision makers (academics, researchers and civil servants) (n=11) and citizens (n=201) irrespective of any healthcare intervention of interest.

For the pilot test of the MCDA framework for making reimbursement decisions, a mock reimbursement decision committee (n=7), consisting of individuals with potential decision-making responsibilities or academic interest in this field, appraised the value of four interventions: primary prophylaxis for severe haemophilia A, zinc supply for diarrhoea prevention, anastrozole as first line therapy for hormone-receptor-positive postmenopausal women with metastatic breast cancer, and ticagrelor + acetylsalicylic acid for patients with acute coronary syndrome without ST elevation and moderate to high cardiovascular risk. For each of these health technologies, a by-criterion MCDA matrix was created to assemble the HTA information. Based on the information included in this matrix, the mock committee could appraise the technologies of interest.

Respondents were asked to score each criterion individually on a 4-point scale (0-3), where 3 represents the highest level of fulfilment of each decision criterion and 0 the lowest. The calculation of the MCDA value estimates was done by summing the weighted scores of all criteria.

Scores were standardised to a 0 to 1 scale by dividing them by the maximum possible score of 3. The MCDA values are obtained by taking the sum of combined weights and scores for all decision criteria and thus have a value between 0 and 1, indicating lowest and highest priority respectively. The MCDA values allow for cross comparison of healthcare interventions. At the end of the scoring session, results are presented to participants to promote discussion.

The pilot test showed that scoring took on average 11.15 minutes (range 7-18 minutes) per-healthcare technology per participant. The weighted total scores ranked the interventions considered in the pilot test as follows (ranges of MCDA scores, reflecting variation between individual's valuation, are mentioned between brackets):

- zinc supply for diarrhoea prevention: average weighted score of 0.904 (range 0.595 to 0.977)
- anastrozole as first line therapy for hormone-receptor-positive postmenopausal women with metastatic breast cancer: average weighted score of 0.822 (0.698-0.934)
- primary prophylaxis for severe haemophilia A: average weighted score of 0.794 (0.782-0.986)
- ticagrelor + acetylsalicylic acid for patients with acute coronary syndrome without ST elevation and moderate to high cardiovascular risk ticagrelor: average weighted score of 0.708 (0.449-0.945)

Participants considered the EVIDEM framework useful as a means for incorporating HTA into decision-making. Comments related to the adequacy of information presented in the evidence summary and challenges for interpretation and valuation of some criteria.

3.3.3.2 Lombardy

Lombardy is a region in the North of Italy with a population of 9.8 million. It implemented a MCDA approach for decision making in 2012. There are two committees with a responsibility in decision making that can be supported by MCDA: the prioritization committee for emerging technologies and an appropriateness committee for diffused technologies. The prioritization committee decides on the priority of the appraisal of emerging technologies. Their decision is based on the scoring and weighting of **8 domains**: (1) technical relevance, (2) safety, (3) efficacy in clinical research, (4) effectiveness in clinical practice, (5) financial impact, (6) equity of access, (7) social impact and (8) organizational impact. The appropriateness committee makes an appraisal of the appropriateness of reimbursement of health technologies based on the scoring and weighting of **15 criteria**:

- Completeness and consistency of reporting evidence.
- Relevance and validity of evidence
- Disease severity
- Size of population affected
- Public health interest
- Type of medical service
- Improvement of safety and tolerability
- Improvement of efficacy/effectiveness
- Improvement of patient reported outcomes
- Clinical guidelines
- Comparative intervention limitations
- Budget impact on health plan
- Cost-effectiveness of intervention
- Impact on other spending
- Adherence to requirements of decision-making body

Each committee member assigns personal weights to either domains or criteria through an online form. The scale used for the weighting is a direct and constrained rating scale. For the 8 domains, for instance, each committee member is asked to assign a score of 8 to the domain considered as most important and 1 to the domain considered as least important, and then to distribute unassigned weights to the other domains. This is done for each emerging technology. The appropriateness committee members are asked a similar task, be it that they weight the criteria and not the domains.

Once the members have assigned their weights, the personal weights are discussed online and during meetings. Based on these discussions, each member can modify his weightings. After the weights are assigned and approved, members are given access to the full evidence available: manufacturer's requests, independent reports like the horizon scanning reports by AGENAS, other supporting evidence.

The scoring is done on a 0 to 4 scale. Committee members individually score the performance of the proposed technology with respect to available alternatives (or any actual standard of care) using an online form. The scores are labelled as follows: 0 = absence of relevant information; 1 = comparative lesser value; 2 = comparative similar value; 3 or 4 = comparative (slightly or highly) better value. The members also have to write a comment for each score.

Based on the scores and the weights, priority (or appropriateness) indices are calculated for each technology. This is done by summing the weighted criteria scores. A qualitative conclusion is assigned to the values as follows:

- if the index lies between 0 and 0.25 (i.e. one or more of the scores is 0 and the estimate averages are less than 1), the intervention cannot be evaluated in a robust way;
- if the index lies between 0.25 and 0.50 (i.e. the estimate's averages are among 1 and 2), the relative value of the intervention is less or equal to the value of alternatives;
- if the index lies between 0.50 and 1 (i.e. the estimate's averages are between 2 and 4), the proposed intervention has a better overall comparative value than the alternatives.

Qualitative analysis of comments written for each domain/criterion is performed. Comments are categorized by two reviewers. Disagreement is resolved by consensus. The comments are analyzed descriptively and reported. The most frequent and robust arguments are proposed as possible motivations for the decision.

Both the priority and appropriateness index and categorized comments are discussed and revised to verify the coherence between scores and comments, eliminate ambiguities and identify further areas of assessment.

The Lombardy region tested the reproducibility of the indices by letting two different subgroups repeating the judgment for some technologies independently of each other. They found high reproducibility.

3.4 Proposed application for needs appraisal in Belgium by means of MCDA

3.4.1 Scope and overview of methods

The scope of the MCDA approach and a general overview of the methods used in this MCDA are summarized in Table 5.

The methods are detailed in the following paragraphs.

Table 5 – Scope and process of the Belgian pilot study

MCDA step	
1. **Decision problem**	Ranking of diseases according to their therapeutic and societal need. Therapeutic need is defined from the perspective of the patient; societal need is defined from the perspective of the society. The decision problem is directly related to the requirement to establish a list of needs in the context of the ETR procedure for innovative medicines. Products eligible for ETR should target a condition that figures high on the list of needs.
2. Selecting criteria	The selection procedure of the criteria included in the MCDA has been described in KCE Report 234 on incorporating public values and preferences in health care decision making. In the pilot, the MCDA criteria are complemented with criteria emerging from the citizen laboratories executed by the KBF.[5, 11] Criteria included in the MCDA include: For therapeutic need: • Impact of disease on quality of life, given current treatment • Impact of disease on life expectancy, given current treatment • Inconvenience of current treatment For societal need: • Impact of disease on public expenditures • prevalence of the disease Additional criteria, to be considered qualitatively after the MCDA results have been produced are: • psychosocial well-being / patient frailty • impact of the patients' disease on carers' quality of life • autonomy

	• integration of care
3. Collecting and presenting evidence	Evidence with respect to each MCDA criterion is collected and summarized in evidence tables, according to a specific template (see paragraph 3.4.2). It might be necessary to collect primary information directly from patients or other experts if no evidence is available from literature. For aspects related to the impact of a disease on quality of life and the inconvenience of current treatment, it might be appropriate to consult patient representatives (see paragraph 3.4.3). The quality of the evidence is assessed. If no evidence is available, it is stated as such. Evidence on different dimensions of a criterion is presented as detailed as possible.
4. Scoring	A 4-point categorical scale is used for scoring the performance of a disease on each of the criteria (see paragraph 3.4.4). Scores are discussed in the commission when there is a large variation across individual scores. The objective of the discussion is to clarify elements in the evidence tables or allow people to justify their score. A second round of scoring is performed after the discussion. Criteria for which there is little disagreement amongst scorers, are not discussed. The individual scores given after the discussion round(s) are used in the next step.
5. Weighting the scores	The weights derived from the general public (Table 3) are multiplied by the individual scores. The weighted scores are then summed to obtain an individual total weighted score for each disease per respondent. See paragraph 1.1.1.
6. Aggregation	The mean and median of the individual total weighted scores are calculated for each disease.
7. Dealing with uncertainty	Uncertainty in the scores is dealt with by discussing the criteria for which there is much variation across individual scorers. This is the first step to reduce variation in scores due to misunderstanding of the evidence in the evidence tables or different appreciations with respect to the quality of the evidence. MCDA Model uncertainty relates to the use of the median weighted score as the value that determines the rank of a health condition. Variation in rankings following the use of mean weighted scores instead of median weighted scores are assessed. See paragraph 3.4.6.
8. Interpretation	When applying the MCDA to therapeutic need on the one hand and societal need on the other hand, two separate rankings are obtained. Commission members will have to agree on the final ranking. Additional criteria, not yet included in the MCDA (cfr. KBF research), might need to be discussed amongst decision makers after the rankings are produced based on the MCDA. If rankings have to change based on these discussions, they need to be justified, using the arguments related to the criteria of the KBF research. See paragraph 3.4.7. The Belgian MCDA model does not aggregate the median total weighted scores of therapeutic and societal need to obtain one overall score or ranking for need (see paragraph 3.4.8).

The **practical application** of the MCDA hence requires:

- an overview of the scientific evidence with respect to several health conditions on all criteria considered relevant for defining needs (i.e. the criteria mentioned in Table 3), including evidence collected from patients regarding impact of disease on quality of life and inconvenience of current treatment;
- the scoring of this evidence on the scoring scale chosen for the MCDA in step 2 of the MCDA development process;
- weigthing the scores assigned to each criterion according to their relative importance for the decision on therapeutic and societal need;
- the determination of the sum of the weighted scores for therapeutic – respectively societal– need;
- the ranking of all health conditions according to their total weighted score;
- the identification of additional elements not yet included in the quantified MCDA but nevertheless relevant for the ranking;
- judging and justifying how these additional elements should alter the ranking resulting from the quantitative part of the MCDA;
- agree on the final ranking of diseases in terms of therapeutic, respectively societal, need.

Therefore, the following tools and methods have to be developed:

- A template for summarizing by-criterion-evidence for each disease (see paragraph 3.4.2) and a method for collecting additional information from external experts or, in particular for impact of the disease on quality of life and inconvenience of current treatment, from patients (see paragraph 3.4.3).
- A scale for scoring diseases on each of the criteria (paragraph 3.4.4).
- A practical procedure for weighting the scores and aggregating the weighted scores of individual commission members (paragraph 1.1.1)
- Ways to deal with uncertainty in the (quality of the) evidence, especially with respect to the score to assign if the evidence is (highly) uncertain or the quality is low (paragraph 3.4.6)
- Ways to deal with criteria or values that are not covered by the criteria included in the MCDA. These could be ethical values that are hard to capture or quantified as operational criteria (see paragraph 3.4.7).
- Ways to deal with the two rankings for need (see paragraph 3.4.8)

3.4.2 Templates for summarizing by-criterion-evidence

Evidence on the impact of a health condition on the criteria used for defining therapeutic and societal need must be summarized in separate “summary of evidence”-tables for each criterion, hereafter called “evidence tables”. Results included in the evidence tables ideally come from well-conducted systematic reviews. Findings should be summarized in tables following a standard template.

The evidence table for each criterion contains the following elements:

- A list of all important dimensions of the criterion affected by the disease (dimensions of the criterion ‘impact on quality of life’ encompass, for instance, impact on self-care, on mobility, etc.; dimensions for the criterion ‘public expenditures’ encompass different types of direct healthcare expenditures (drugs, physcian visits), indirect healthcare expenditures (transportation to hospital), and transfer payments (such as disability benefits, increased child allowance etc.));
- The (quantified) magnitude of the impact of the disease on each of these dimensions, compared to a situation without the disease;
- Sources used to describe the magnitude of the impact of the disease on each of the dimensions, including study designs and numbers of participants where appropriate;
- Comments: may encompass comments about the uncertainty of the estimates, concerns about the quality of evidence (may vary by dimension) or any other comments. **Assessment of the quality of evidence** submitted should be performed by an independent reviewer. If multiple studies have been performed on a particular criterion, a quality assessment should ideally be performed for each study as well as for the collection of evidence. We present a tool for assessing the quality of individual studies and of the collection of evidence in Appendix 1. This tool is based on work of EVIDEM, but tailored to the criteria included in the Belgian MCDA. The tool can be used by assessors who

review the submitted evidence as a checklist and as a reporting tool. Based on the assessment, an overall grade can be given to the quality of the evidence provided for a specific criterion. This grade and its justification can be copied by the independent reviewer in the comments box of the evidence tables.

Given the five criteria considered for the assessment of the therapeutic and societal need in a particular disease, five evidence tables were created for the MCDA tool:

- The first table describes the impact of the disease on **health-related quality of life**, given the treatment currently already available to patients. E.g. it describes whether patients still have problems with usual activities, despite the treatment that is currently available and provided to them. Also impact on mobility, self-care, pain/discomfort and anxiety/depression is described if evidence is available.
- The second table describes the **inconvenience or discomfort of current treatment**. Inconvenience relates to the route of administration (e.g. injection or oral), place of administration (e.g. at the hospital or at home), frequency, duration, and overall burden of treatment in terms of treatment-related morbidity.
- The third table describes the impact of the disease on **life expectancy**, given current treatment.
- The fourth table describes the **frequency** of the disease in the population.
- The fifth table describes the **public expenditures** associated with the disease. Public expenditures encompass direct health care expenditures, sickness and invalidity insurance benefits, publicly financed home care services, etc. The score should be based on the public expenditure per patient.

The templates created for the Belgian MCDA pilot was inspired by the work of EVIDEM (December 2014: Instrument for synthesizing evidence) (Table 6). It is important that evidence on relevant dimensions of the criteria is presented separately if possible. In the EVIDEM framework, these are called sub-criteria. The better and the more complete the information on the respective dimensions or sub-criteria, the more evidence-based the scoring of the commission members will be.

Table 6 – Templates for the summary of evidence tables

Criterion: Impact of the disease on quality of life (QoL)		
Dimensions	**Magnitude of the impact of the disease**	**Sources of evidence**
Mobility Self-care Usual activities Pain/discomfort Anxiety/depression …		
Comments		

Criterion: Inconvenience of current treatment		
Aspects of current treatment causing inconvenience to patients	**Magnitude of the inconvenience induced by each of these aspects**	**Sources of evidence**
E.g. place of treatment, duration of treatment, timing		
Comments		

Criterion: Impact of the disease on life expectancy		
Dimensions	**Magnitude of the impact of the disease on each dimension**	**Sources of evidence**
Life years lost due to the disease		
Comments		

Criterion: Frequency of disease		
Dimensions	**Magnitude of the impact of the disease on each dimension**	**Sources of evidence**
Prevalence		
Comments		

Criterion: Disease-related public expenditures		
Dimensions	**Magnitude of each type of expenditures**	**Sources of evidence**
E.g. Direct public healthcare expenditures Indirect public expenditures (disability benefits, increased child allowances …)		
Comments		

3.4.3 Collecting patient input

When the evidence from the literature regarding relevant impacts of the disease or the treatment is insufficient or lacking, input might have to be collected directly from patients. Patient involvement in HTA (and HSR) research is described as a strong added value by the European Platform for Patients' Organisations, Science and Industry (EPPOSI). It is clear that patients, who lived the experience of a health condition and its management, or informal carers who are able to speak on behalf of patients, have special knowledge, which might shed a different light on the impact of a disease.[12] An understanding of patients' needs and preferences is essential for an accurate appraisal of the impact of a disease on their quality of life and the inconvenience of the current management of their disease. It also avoids that discussions amongst the experts about the score for impact on quality of life or inconvenience of current management becomes more based on their personal (clinical) experience.

Qualitative research (such as interviews and focus groups) can complement quantitative research (surveys measuring patient reported outcomes or psychosocial, socioeconomic and functional outcomes) to add patient experience to the knowledge base.(HTAi PCIG 2015)

When patient involvement is considered in the process of the MCDA for assessing needs, it is recommended to have a dedicated staff member or subcontractor whose role is to support patients to contribute to the evidence. To maintain patient group's engagement in such processes, it is also important to give feedback to the organisations who have contributed to the assessment, to share what contributions were most helpful and provide suggestions to assist their future involvement.[13]

3.4.3.1 *HTAi Interest Group Patient and citizen involvement in HTA*

A lot of interesting work on why and how to involve patients in health technology assessment has been performed by the HTAi Interest Group "Patient and citizen involvement in HTA".

Regarding the question of whether individual patients or rather patient groups should be consulted, the documents state:

> *"An individual patient and a patient group may provide different but equally useful perspectives and contributions.*
>
> *A patient and/or his/her caregiver can bring rich knowledge formed from living with the health condition, its diagnosis and treatment. They may also be informed about the experience of other patients through their personal interactions.*
>
> *A patient group may be comprised of patients or representatives of a patient community. If it is the latter, their knowledge will be less rich in specific experience but may be informed by a wider group of patients. A representative of a patient group may be able to consult their members to respond to the HTA processes. […] General organisations representing broader patient and user interests may not share the same perspectives as disease-specific groups."*

The group also mentions that it is important to have transparency about potential conflicts of interest individuals or patients groups may have, in much the same way as for every other stakeholder consulted in the context of an assessment. Therefore, potential interests should be registered when collecting information from patients and patient groups.

The group has developed tools and other resources that are all freely available from the HTAi website.[14]

The group has developed templates for written consultation of patient groups (Appendix 2). The template encompasses prompts for patient groups to help them explain patients' and care-givers'/carers' perspectives in a manner that is most likely to inform experts that have to make an appraisal of the therapeutic needs of patients based on the evidence provided.

It is of utmost importance that the reporting of the evidence collected from patients using a qualitative approach is done by an impartial person. The reporting person should also declare his or her potential conflict of interests.

3.4.3.2 PROMIS

Besides qualitative research, quantitative data on the health state of patients and their perceptions remain important. Interesting initiatives are taken with this respect. Of worth mentioning is *PROMIS.* PROMIS stands for "patient reported outcomes measurement information system". It is a system of validated and reliable measures of patient-reported health status for physical, mental and social well-being.(http://www.nihpromis.org/about/abouthome) The system encompasses generic tools that measure what patients are able to do and how they feel. The measures can be used as endpoints in clinical studies.

PROMIS was an initiative of seven research institutions and the National Institute of Health (NIH) in the US, to provide clinicians and researchers access to efficient, precise, valid and responsive adult- and child-reported measures of health and well-being. It started in 2004 and in 2010, a second round of PROMIS funding was provided by the NIH, expanding the network to 13 researchers at 12 research sites. The development of PROMIS took several years of research, including extensive literature reviews for items to be included, item classification, qualitative item reviews, archival analysis and item bank testing (in two waves).

PROMIS changed the way patient-reported outcome tools are selected and employed in clinical research and practice evaluation. It also established a national resource for accurate and efficient measurement of patient-reported symptoms and other health outcomes in clinical practice in the US. As of yet, the system is being translated into Dutch and Flemish by a consortium of research groups in the Netherlands and Belgium.[15]

The idea of PROMIS is that patients themselves can fill out the tools through a computerized adaptive testing (CAT) system or on paper. CAT implies that the questions depend on the answers they give to previous questions. After completion, the respondents receive a feedback report that explains how their responses compare to those of other respondents of the same population. The outcome measures included in PROMIS are generic, meaning that a wide variety of patients could fill out the questionnaires, and that they are applicable to adults and children.[15]

PROMIS, once implemented in routine practice, would provide a wealth of information on patient-reported outcomes, and would be an excellent quantitative input for the evidence tables that are used as a basis for the MCDA tool for appraising therapeutic needs. As of yet, it is too early to use the PROMIS data systematically, as the database needs to be set up and data need to be collected from patients first, but in the long run (i.e. >5 years), this would be a source of evidence that could still be complemented with a qualitative part where necessary.

3.4.3.3 Vlaams Patiëntenplatform (Flemish Patient Platform)

Another approach for including the patients' voice in the MCDA has been presented by the Vlaams Patiëntenplatform (VPP), the Flemish umbrella organisation of patient associations, in the context of a study about patient-driven innovation.[16] The approach aimed at defining needs of patients in function of the characteristics of their disease. Because the project aimed at identifying needs of a wide range of patients, they developed a framework in which several diseases could be fit.

A first step was to identify groups of patients with a comparable pattern of needs. In step 2, patients from these groups were recruited for a participative process used for mapping needs. The needs are then visualised by means of the so-called "persona's", which are healthcare users or patients that do not really exist but are described as such. The study also generated ideas for products and services that could meet the needs of patients and their carers, which eventually led to a patient-driven innovation agenda.

What is particularly interesting in this approach and relevant for the current study, is the generic nature of the method to define needs that resulted from it. Patient groups were clustered into four clusters, using two main dimensions characterizing their disease with each time two options (Figure 7):

- Dominantly metal disorder versus dominantly physical disorder
- High level of impairment versus low level of impairment

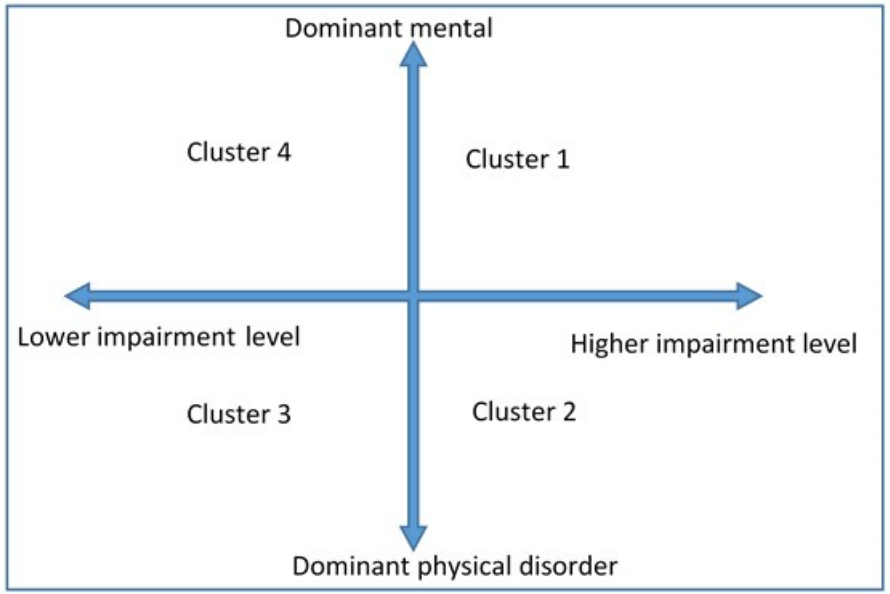
The methods used in the study to identify the needs in each cluster included diaries, pictures taken by the patients or their carers, interviews, surveys. The analyses of the individual data allowed the researchers to create cards that describe a fictitious person and his daily life. These were used to identify 12 domains of life affected by disease:

- Bodily functions
- Intellectual tasks and stress
- Activities of daily living
- Communication
- Mobility
- Physical capacity and movement
- Social contact
- Leisure activities
- Technology
- Self-care, informal care, care environment
- General attitude
- Finances

Needs in each of these domains were discussed for each cluster of patients amongst people of patient organisations, collaborators at VPP, representatives from (health) care centres, industry and research institutes. The brainstorming sessions gave rise to a list of ideas for innovation for each cluster of patients.

As the clusters are defined in generic terms (based on the two dimensions mentioned before), the outcomes of the needs assessment could also be used for patients that were not directly included in the study or described as persona's with a particular disease. Instead, each disease could be located in the figure and presumed to have similar needs as those defined for the persona's located in the same cluster.

Figure 7 – Patient clusters



3.4.3.4 Chosen approach for Belgian MCDA pilot study

We decided to test the feasibility of gathering additional patient input for the evidence tables for one case included in the pilot study. The chosen case was that of Alzheimer's disease. The purpose of involving patients directly was considered mainly important for obtaining evidence on the impact of the condition on the health related quality of life of patients –and in the case of Alzheimer also on their informal caregivers- and the inconvenience of current care or management.

A collaboration with the "Vlaams Patiëntenplatform" (VPP), the "Ligue des Usagers des Services de Santé" (LUSS) and "Patienten Rat & Treff" (PRT) was set up to develop a procedure for patient involvement in the unmet medical needs appraisal procedure of the CATT – CATT that would be feasible for future implementation. VPP, LUSS and PRT are all umbrella organisations of patient associations in Belgium, VPP for the Flemish-

speaking patients, LUSS for the French-speaking patients and PRT for the German-speaking patients.

We explored various options. The basic decision to take is “who to involve” and “how to involve them”. For the “who”-question, basically two options exist: individual patients or representatives of patients (i.e. patient groups). In the case of Alzheimer, other relevant groups to involve are the patient’s family and their informal caregivers. For the “how”-question, we also distinguished four options: written questionnaires, individual interviews, group interviews and focus groups.

The written approach requires the least effort from the researcher who has to collect the information and is in that sense the most practical approach. However, the lack of face-to-face contact risks to create a distance between the patient group and the decision maker that reduces the feeling of real engagement of the patient group in the endeavour. We therefore considered it preferable to ensure a face to face contact.

The question was then who to interview: individual patients or patient representatives. It should be noted that for the collection of information directly from individual patients legal requirements need to be fulfilled. An approval of an ethical committee is required for the collection and analysis of information derived directly from patients (e.g. by questionnaires or interview). The privacy commission needs to be notified of the processing of personal data. An insurance for the interviewees has to be taken. We opted for a more practical –and more feasible approach in real life when data need to be collected for many health conditions- being to interview representatives of patients and their family or caregivers during two separate face-to-face meetings. An interview guide was established, based on the work of the VPP[16] and on the HTAi PCIG template[14].

Because our pilot test not only involves patients but also patients’ caregivers and their family, information on criteria that are not directly included in the MCDA, but are nevertheless deemed important by the citizens of the citizens laboratories of the KBF, could be collected. The information thus collected was included in additional summary of evidence tables.

Input was collected from patients with severe burn injuries and representatives/informal caregivers of patients with Alzheimer’s disease by means of semi-structured face-to-face interviews.

First, the accounts of 29 patients who were involved in a previous KCE study on aftercare for patients with severe burn injuries, were re-used to learn about unmet medical needs. The burned patients were recruited with the help of the care coordinators of the six Belgian burn care centres. They agreed to send invitations to a target list of patients. Patients were asked to contact the research team if they were interested to participate. These semi-structured face-to-face interviews were conducted in the period between January and April 2013.

The respondents shared their experiences with the care they received after a hospitalisation for severe burn injuries. They were asked especially about what went right and what went wrong during their care trajectory, about how severe burns impacted their daily lives, participation in social activities, financial situation and reintegration. The interviews lasted between 2 and 2.5 hours and the location was chosen by the interviewee. All interviews were audio-recorded and transcribed verbatim. They have been analysed by means of Nvivo specialised computer software.

Second, the VPP was asked to contact their member that could represent the Alzheimer patients and their informal caregivers. The *Alzheimer Liga Vlaanderen* (Flemish Alzheimer Ligue) was identified. Two representatives/informal care givers of patients with Alzheimer disease have been interviewed. It was the Alzheimer Liga itself who invited two of its members to participate. A KCE researcher interviewed the two informal caregivers at the same time by means of a topic list in January 2016.

The respondents talked about the most important problems Alzheimer patients experience, the treatments, the impact of Alzheimer’s disease on communication, mobility, social contacts, suffering, physical functioning, family life, finances, quality of life and their unmet needs. The interview lasted about two hours and took place in the offices of the Alzheimer Liga. The interview was audiotaped but not transcribed verbatim. A summary was made based on field notes.

Key points

- **Patient involvement in MCDA is important for completing the summary of evidence tables when evidence on impact on patients' quality of life or inconvenience of current treatment is lacking.**
- **Involvement of patients can happen in a quantitative and qualitative way. Both approaches are complementary. Standardised quantifiable patient-reported outcomes measures are currently lacking, which poses a problem for the consistency in scoring the impact on these outcomes.**
- **While solutions are being explored (e.g. PROMIS), qualitative approaches for filling the evidence gaps could consist of group interviews with patient representatives. These should be performed by a neutral person, who derives the key messages from the interviews. The key messages could then be included in the summary of evidence tables.**

3.4.4 Scoring scale

3.4.4.1 Scales used in KCE's discrete choice experiments in the general public

In the survey performed in the general public in 2014[1] discrete choice experiments were used to derive the weights of the criteria used to define therapeutic and societal need. Each criterion had a pre-defined number of levels (Table 7).

Table 7 – Criteria for therapeutic and societal need with their respective levels

Type of need	Criteria	Levels
Therapeutic need	Inconvenience of current treatment	Patients experience ... • much inconvenience from current treatment • little inconvenience from current treatment
	Quality of life with current treatment	Patients currently have a quality of life of ... • 8 out of 10 • 5 out of 10 • 2 out of 10
	Life expectancy with current treatment	Patients ... • no longer die from the condition • die 5 years earlier than people without the condition • die almost immediately from the condition, despite current care
Societal need	Prevalence of the condition	The condition is... • rare: less than 2000 people in Belgium have the condition • not so frequent: between 2000 and 10 000 people in Belgium have the condition • rather frequent: between 10 000 and 100 000 people in Belgium have the condition • very frequent: more than 100 000 people in Belgium have the condition
	Condition-related public expenditures	• Little public expenditures per patient • Much public expenditures per patient

The simplest way to apply the results of this survey directly in the decision-making context would be to let the committee members score each health condition on these criteria, using the levels that were used in the survey. This would allow an easy application of the estimated multi-nomial model, resulting in a “utility” sore for each disease. The higher the score, the higher the need. This would obviate the need for using the level-independent criteria weights as presented in Table 3.

The advantage of this approach is that one remains closer to the immediate results of the discrete choice experiments presented to the general public in the survey.

However, there are two main problems with this approach:

- For MCDA, it is important to score each criterion on a **common scale** (i.e. identical in terms of the number of response categories and visual presentation), because a lower number of categories would imply a lower maximum score. For example, if public expenditures has only two levels (corresponding to a score of 1 or 2) and prevelance has four levels (corresponding to a score of 1, 2, 3 or 4), the maximum score differs. Hence, a maximum score on prevalence gets a higher weight than a maximum score on public expenditures, independent of the relative importance of each of these criteria according to Table 3. The challenge is thus to develop a common scale for scoring health conditions on each of the criteria. Reimbursement committee members then need to develop experience with positioning scientific evidence with respect to a condition on that scale.
- The number of levels might be considered insufficient by decision makers. When judging the public expenditures related to a particular disease, decision makers might wish to give an intermediate score, rather than "little" or "much". It is important to develop a scale that allows **sufficient differentiation**, but at the same time is intuitive and easy to use.

The scale format referred in the latter point can have an important effect on the responses given as well as on the associated psychometric properties.[17] Weijters et al. (2010) state that the choice of a particular rating scale format can be broken down into two major components:

(1) the number of response categories to be offered, including the choice for an odd of even number of categories and

(2) the labelling of response categories.[18]

The number of response categories affects scale reliability and can have profound effects on both the cognitive burden and the sensitivity of the scoring design.[17] More response options allow to better discriminate between assessments. On the other hand, too many response options may reduce the clarity of the meaning and the scoring task becomes more complex.[17] Weijters et al. (2010) developed a decision framework for selecting a response scale format. For the selection of the number of response categories, a trade-off has to be made between maximizing the potential information transmission versus minimizing respondent demands.

For populations with high cognitive ability, they recommend the use of scales with more response categories.[18] Grondin et al. (2010) conclude from a literature review that a response scale using from five to nine categories should produce relatively good results.[17] They also found that scales with odd numbers of response categories may be more reliable than scales with even numbers of response alternatives because they represent reality in a better way. However, empirical studies examining the relation between reliability and number of response categories have produced conflicting results.

As for the labelling of response alternatives, Weijters et al. (2010) recommend to fully label the scale when the objective of the scale measurement is to derive mean or median scores. This means that every response category gets a label, not just the extreme values on the scale.

Key points

- **For the purpose of MCDA, a common scale needs to be used for scoring diseases on the different criteria and the scale needs to be intuitive and applicable.**
- **Based on findings from published literature reviews, a fully labelled four-level ordinal scale was constructed for each criterion included in the MCDA.**

3.4.4.2 Scoring scale in EVIDEM

None of the MCDA frameworks examined recommends a particular scale. EVIDEM presented Likert scales with precise labels for the response categories, but agencies using the EVIDEM framework have also used other scales (e.g. Colombia used an ordinal 0-3 scale). Likert scales are scales with discrete response categories. Some examples given by EVIDEM are presented in Table 8

Table 8 – Examples of scales for scoring decision criteria and scoring instructions (EVIDEM)

Criteria	Scoring scale Low → High	Scoring instructions
Disease severity	Not severe → Very severe O 5 Very severe O 4 O 3 O 2 O 1 O 0 Not severe	Score the severity of disease by considering its impact on the patient; consider the disease as a whole, not only the aspect of the disease that the intervention is targeting. (Economic aspects are not considered here). *Score from an absolute point of view (not relative to comparative interventions)*
Size of population affected	Very rare disease → Common disease O 5 X > 500/10 000 O 4 X < 500/10 000 O 3 X < 100/10 000 O 2 X < 10/10 000 O 1 X < 5/10 000 (rare) O 0 X < 2/100 000 (ultra rare)	Score the size of the population affected by the condition keeping in mind that this would be the size of the population potentially benefitting from the intervention. (Economic aspects are not considered here - see economic cluster). *Score from an absolute point of view (not relative to comparative interventions)*
Unmet needs (i.e. comparative intervention's shortcomings)	**No unmet needs → Many & serious unmet needs** O 5 Many and serious unmet needs O 4 O 3 O 2 O 1 O 0 No unmet needs	Score limitations of comparative interventions with respect to treatment or prevention of the targeted condition. (Limitations due to cost of are not considered here).

Source: EVIDEM Collaboration, December 2013[19]
Note: this is an extraction from the full table presented in the EVIDEM document.
The scoring instructions are not necessarily in line with the definitions used in the population survey described in KCE-report 234.

The EVIDEM framework is continuously in evolution. The previous version still included labels for each response category. Labels have the advantage of giving an idea of what could mean a particular level on the scale. It might lead to more consistency in the scoring across diseases or interventions. The disadvantage of giving examples, however, is that they might sometimes create confusion. In the version of June 2015, the labelling has been dropped.

The EVIDEM framework gives a concrete definition of each criterion. This helps raters interpreting the content of the criterion. A standard definition also allows consistency in what different raters take into account when scoring a criterion. For example, "disease severity" is defined as "*Severity of the health condition of patients treated with the proposed intervention (or severity of the health condition that is to be prevented) with respect to mortality, disability, impact on quality of life, clinical course (i.e., acuteness, clinical stages)*"; "size of population affected" is defined as "*Number of people affected by the condition (treated or prevented by the proposed intervention) among a specified population at a specified time; can be expressed as annual number of new cases (annual incidence) and/or proportion of the population affected at a certain point of time (prevalence)*".

Because we used a slightly different MCDA framework than the one presented by EVIDEM, we cannot copy-paste their definitions and criteria. For example, while EVIDEM combines the mortality and quality of life in one criterion, defined as "severity of disease", we chose to consider these as two separate criteria in the MCDA framework, in order to allow taking the public preference for each separately into account in the decision making process. By including both criteria in a more general 'disease severity' criterion, the implicit weights of the decision makers will be used. As shown in KCE report 234, decision makers and the general public have different preferences regarding the relative importance of the impact of a disease on life expectancy and the impact on quality of life, so the separation actually does matter.

As for the definition of prevalence, we think it is not good to use "and/or" in the definition, as this may give rise to confusion or inconsistency amongst raters: if some base their score on the incidence, others on the prevalence and still others weight both and make an overall judgement, scores may vary to a large extent. It is important that the scoring instructions are explicit about how to deal with prevalence on the one hand and incidence on the other hand.

Key points

- **It is important to provide a clear and unambiguous definition of the criteria to be scored.**
- **The definition of response categories should also be clear.**

3.4.4.3 Chosen scoring scale for Belgian MCDA application

It is important to use a common scoring scale for the evidence for all criteria. For each of the five criteria considered, a scale needed to be defined.

The scoring of the criteria needs to take into account the assessment of the magnitude of the impact on all dimensions, the meaningfulness of the demonstrated magnitude of the impact and the quality of the evidence upon which the magnitude estimates are based.

No literature was found on the most appropriate type of scale or scale characteristics for this kind of purpose. The general conclusion in most of the literature on preference measurement scales is that the appropriateness of a scale is context-dependent.

We tested a **four-category scale** for scoring the magnitude of the impact of a disease on criteria, in line with the scale used by the EtD framework and the previous version of the EVIDEM framework. In contrast with the scales used by the EtD and EVIDEM framework, we opted for fully labelled response categories (see Table 9).

Table 9 – Scoring scale and response category labels for each criterion

Criterion	Response categories and labels
Impact of the disease on life expectancy, given current treatment	0 : The disease has no impact on life expectancy 1 : The disease has some impact on life expectancy (patient loses a small proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment) 2 : The disease has a high impact on life expectancy (patient loses a large proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment) 3 : The disease has a very high impact on life expectancy (patient dies almost immediately, despite the best available current treatment or care)
Impact of disease on quality of life given current treatment	0 : The disease has no negative impact on the quality of life of patients, compared to people without the disease 1 : The disease has some negative impact on the quality of life of patients, compared to people without the disease 2 : The disease has a high negative impact on the quality of life of patients, compared to people without the disease 3 : The disease has a very high negative impact on the quality of life of patients, compared to people without the disease
Inconvenience of current treatment	0 : Current treatment is not or only slightly inconvenient to patients 1 : Current treatment is somewhat inconvenient to patients 2 : Current treatment is highly inconvenient to patients 3 : Current treatment is very highly inconvenient to patients
Frequency of disease	0 : less than 2000 people in Belgium have the condition (less than 1 per 5500) 1 : between 2000 and 10 000 people in Belgium have the condition (between 1 per 5500 and 1 per 1100) 2 : between 10 000 and 100 000 people in Belgium have the condition (between 1 per 1100 and 1 per 110) 3 : more than 100 000 people in Belgium have the condition (more than 1 per 110)
Current disease-related public expenditures per patient	0 : The disease currently has a very small impact on public expenditures per patient 1 : The disease currently has some impact on public expenditures per patient 2 : The disease currently has a high impact on public expenditures per patient 3 : The disease currently has a very high impact on public expenditures per patient

3.4.5 *Weighting and aggregating*

In the Belgian MCDA approach, the weights derived from the general public (Table 3) are multiplied by the individual scores. The weighted scores are then summed to obtain a total weighted score per individual.

This is different from the MCDA examples applied elsewhere. Often, the weights are derived from the decision makers and each time a decision has to be taken. The disadvantage of such an approach is that the weights will vary depending on the disease and hence there risks to be a lack of consistency between appraisals. For the appraisal of needs, it is important to give criteria the same weight in each health condition, to avoid manipulation of the results.

An example of scores and the procedure for weighting for one health condition and one scorer is provided in Table 10.

Table 10 – Example of a scoring table and the weighting procedure for one disease and one scorer

Criteria therapeutic need	Score	Weight	Weighted score
Impact of disease on quality of life	1	0.43	0.43
Impact of disease on life expectancy	2	0.14	0.28
Inconvenience of current treatment	2	0.43	0.86
TOTAL WEIGHTED SCORE THERAPEUTIC NEED			**1.57**

Criteria societal need	Score	Weight	Weighted score
Prevalence of disease	3	0.35	1.05
Disease-related public expenditures	1	0.65	0.65
TOTAL WEIGHTED SCORE SOCIETAL NEED			**1.7**

The total weighted score for this disease should be compared with the total weighted score for other diseases to give it a meaning. Therefore, the appraisal exercise is repeated for many interventions. A ranking can be made based on the total weighted scores.

However, the ranking is not based on the total weighted scores of one individual only. Once the individuals have given their scores, the individual weighted scores need to be aggregated and a measure of central tendency (median or mean) needs to be chosen. We examined the differences in ranking of health conditions appraised with the MCDA tool if the mean total

weighted scores were used versus the median total weighted scores during the pre-test study.

3.4.6 Dealing with uncertainty

Evidence on health or health care is almost never uncertain. It is important to take uncertainty into account in the appraisal process.

The lower the quality of the evidence, the higher the uncertainty. Aspects to be considered in the assessment of the quality of the evidence are the risk of bias, indirectness of the evidence (e.g. combining results of different studies or assuming relationships between parameters that are not directly measured), imprecision (confidence intervals), inconsistency, and publication bias. Results are considered imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. These elements should be described in the fourth column of the summary of evidence tables.

If the available evidence on a particular criterion is highly uncertain, there are several possibilities: the score could be downgraded, making the impact of the disease on that criterion seem less important; uncertainty could be taken into account after the ranking process as an additional criterion that could be used to change the ranking of a disease; people could be allowed to give several scores assuming different levels of uncertainty and then a sensitivity analysis is performed; the scoring could be put on hold until more data are collected or published.

For the first approach, an example could be that there is evidence of very low quality that a disease has a moderate impact on life expectancy. In that case, the score for the criterion 'impact of disease on life expectancy' should be 0 or 1 at most. If the 95%CI upper or lower CI crosses the minimally important impact, the impact should get a score of 0.

In the second approach, first the magnitude of the impact is scored as if there is no uncertainty, and a remark is made that the score is based on low quality evidence. When it comes to deliberating about the rank order of diseases based on the weighted total scores, the uncertainty of the evidence could come in as a reason for lowering the ranking of a disease. We decided to apply this approach in the pilot study for the following reasons:

- If uncertainty would be a reason for giving a lower score, there are actually two possible reasons why a disease gets a low score: because the evidence really points towards a low impact of the disease *or* because the evidence that the impact of the disease is moderate, is highly uncertain. For transparency, it is not good to allow two different rationales for one single score.
- If it is made explicit that the ranking of a disease has changed because of the low quality of evidence, it is immediately clear that the ranking could increase if more evidence becomes available.

The impact on the ranking will depend on how risk averse the commission members are with respect to the different criteria. The more risk averse for important criteria, the higher the impact on the ranking will be.

Key point

Uncertainty of evidence is treated as a separate criterion that is not covered in the MCDA and could give rise to changes in the ranking of diseases *after* the ranking based on the total weighted scores has been created. The impact will depend on how risk averse the commission members are with respect to the different criteria.

3.4.7 Dealing with criteria not covered in the MCDA

A method for dealing with criteria and values that are not included in the MCDA had to be developed because the citizens panel of the KBF showed that criteria such as the impact of the disease on the patient's environment (caregivers, family, others), are also very important for resource allocation decisions.

The chosen approach for the pilot study was to present the results of the citizen's panels to the members of the CAIT / CATT to inform them about the values placed on specific criteria by the citizens, beyond those included in the MCDA. The presentation would be given after the ranking based on the MCDA criteria and the aggregated total weighted scores. The commission could then discuss whether the ranking should be changed based on these additional criteria.

It was decided not to apply the same scoring techniques to these additional criteria and to weight them against the other criteria already included in the quantitative part of the MCDA because this would diminish the validity of the approach. The weights measured in the general public, as described in KCE

report 234, were obtained by means of a discrete choice experiment that included only the criteria included in the MCDA as described. Hence they are only valid for decisions that have to weight these criteria. Therefore, it is considered better, from a validity point of view, not to change the weights when additional criteria are considered in the decision, because there would be not valid or justifiable way to change them. Moreover, from a transparency and consistency point of view, it also seems better to keep a core approach that will be applied to all diseases in a standardised manner.

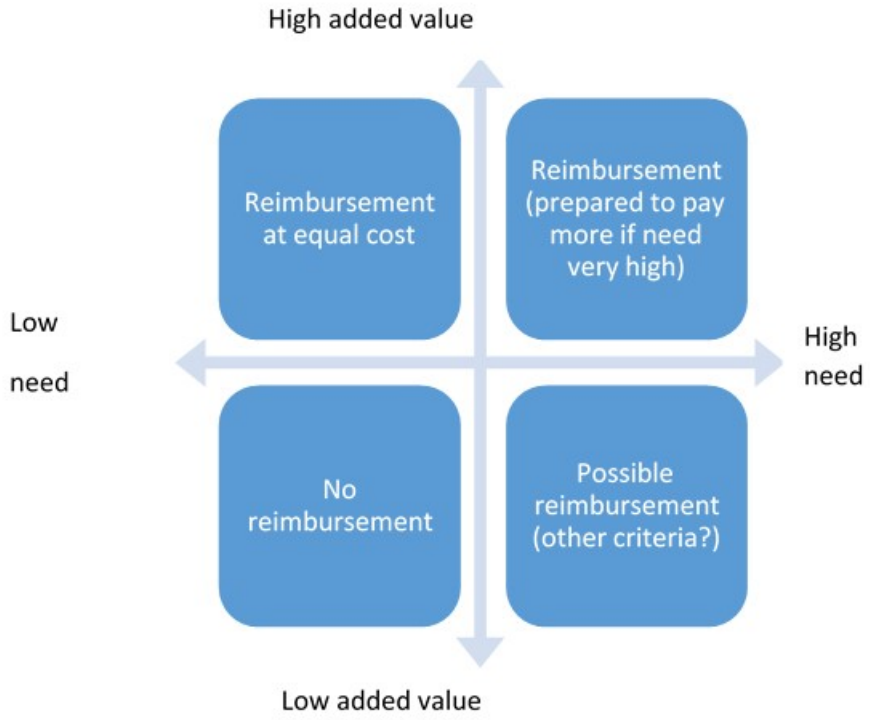
Evidence for some of these additional criteria can be collected from patients and their caregivers during the interviews or focus groups organised for collecting information on the impact of the disease on quality of life. This may require additional questions in the interview guide. The additional information must be summarized and included in separate evidence tables. These evidence tables can have the same structure as those for the other criteria.

3.4.8 Dealing with two lists of "need"

The Belgian MCDA model does not aggregate the median total weighted scores of therapeutic and societal need to obtain one overall score or ranking for need. This has been a deliberate choice. Merging the scores would require a judgment about what is most important: the patient perspective or the societal perspective. While this could in principle be done, we do not have the data to do this in a scientifically valid way, nor in an evidence-based way. Keeping the two perspectives separate and work with the two lists as they are, has the advantage that it is immediately clear on which aspects any new intervention for the diseases which are high on the needs list should show an added value: if a new intervention is developed for a need that is high on the societal needs list, it should either reduce the frequency of the disease (i.e. be curative) or reduce the public expenditures per patient. If both perspectives are merged by aggregating the weighted scores, for instance, the risk that new interventions are accepted for cohort compensation for the wrong reasons increases (e.g. the target indication is high on the list because the prevalence is high, but the intervention only offers added benefit in terms of quality of life).

It is clear, for instance, that the overall need will be judged higher if both therapeutic and societal need have a high total weighted score, and the overall need will be judged lower if the total weighted score for one of the types of need is low. In the diagram representing the decision framework (Figure 8) this will translate into a higher or lower value on the X-axis, but if either therapeutic or societal need is high, always at the right side of the origin.

Figure 8 – Decision framework for cohort decisions



3.5 Testing and evaluating the proposed Belgian approach for needs appraisal

The application of the complete MCDA approach, including the creation of the evidence tables, scoring, weighting, ranking, discussion and deliberation on the resulting priority list, was tested in two phases: a pre-test phase and a pilot phase. The pre-test phase was performed with KCE experts, while the pilot phase was performed with members of the CAIT / CATT.

The pre-test and pilot test were performed on real cases, i.e. evidence tables relating to real diseases, using real evidence. The objective of the pre-test was to assess the feasibility of using the tool, to test the validity of the chosen scale, and to examine the validity of the outcomes (i.e. the perceived validity of the ranking of diseases resulting from the application of the tool). The pilot test aimed at testing the feasibility and acceptability of the amended tool in the CAIT / CATT, i.e. the commission that would become the main user of these tools in the future.

3.5.1 Pre-test evaluation of the MCDA procedure

3.5.1.1 Sample selection

The pre-test was performed with experts at KCE. Eleven experts with a biomedical background were invited to participate. We attempted to mimic as much as possible the actual implementation of the tool by carefully selecting the group of people for the pre-test study. The selected group of experts was as large as the group of experts that will be responsible for proposing a list of needs to the General Council of the RIZIV / INAMI, i.e. the CAIT / CATT. As the members of the CAIT / CATT, all of the experts invited to participate in the pre-test had a medical background and experience with interpreting and assessing evidence. They were either involved in the development of guidelines for good clinical practice or health technology assessments. Eleven KCE experts participated in the pre-test.

This number and profile was chosen because the actual commission (CAIT / CATT) also has eleven members with a biomedical background. Only experts that were able to participate in the two meetings that would be organised for the pre-test, were included, as it was considered important that each participant would be able to complete the entire pre-test process. Two experts pointed out that they would not be able to participate in one of the two scheduled meetings. Their invitations were withdrawn and two other experts were invited.

3.5.1.2 Selected cases

In selecting the cases to be included in the pre-test, a balance was sought between types of diseases: mental versus physical diseases, rare versus frequent diseases, and diseases causing sudden death or not causing sudden death. The cases selected for the pre-test, with their general features are presented in Table 11.

Table 11 – Sample of cases selected for pre-test and pilot test, with their features

Case	Lethal	Not lethal	Mental	Physical	Not so frequent	Frequent
Invasive meningococcal disease	x			x	x	
Major depression		x	x			x
Severe heart failure	x			x	x	
Refractive errors		x		x		x
Amyotrophic Lateral Sclerosis	x			x	x	
Alzheimer		x	x			x
Mesothelioma	x			x	x	
Deep mixed partial thickness burns of the skin in children		x		x		x

3.5.1.3 By-criterion summary of evidence tables

The by-criterion summary of evidence tables were created for each disease based on KCE reports and submissions from the industry to the CAIT / CATT in 2015. A few of the evidence tables used in the pre-test study are presented in the Appendix 3.

3.5.1.4 Scoring process

A LimeSurvey was set up, including the instructions for completing the survey, the full evidence tables and the scoring scales. The survey and related material was also sent by e-mail to all participants, because the evidence tables were sometimes quite large and the lay-out options in LimeSurvey were not adequate for creating nice evidence tables. The e-mail also included a personal link to LimeSurvey, which allowed participants to complete part of the survey and return to it later.

The experts were asked to score the magnitude of effect of a disease on each of the criteria included in therapeutic and societal need. They were asked to put comments in comment boxes whenever they felt the need (e.g. for highlighting uncertainty, issues with wording, issues with comprehension).

3.5.1.5 Discussion on tools and processes

Participants were invited to two meetings. A **first meeting** was organised to discuss the pre-test participants' experiences with the scoring process. Participants were invited to comment on the proposed methodology and the scoring rules. The results of this meeting could give rise to adaptations in the scoring rules presented, as well as the clarity of the evidence tables.

A **second meeting** was organised to discuss the results of the MCDA exercise. First, a discussion was held on the criteria for which the scoring differed substantially between raters. A selection of cases was discussed to stay within the predefined timeframe of 2 hours. All raters were given the opportunity to justify their score and explain why they gave that score. A second scoring round during the meeting (individual scoring), allowed to

assess whether the scores became more consistent and whether changes in the ranking of diseases occurred.

The results of the citizen laboratories performed by the King Baudouin foundations were then presented to the pre-testers, in order to allow a discussion on whether the criteria highlighted by the citizen labs but not yet included in the MCDA evidence tables, should alter the ranking of the diseases. After the meeting, the possible changes in the ranking of the diseases, due to the re-scoring of some items and consideration of the results of the citizen labs was assessed.

3.5.2 Pilot study

The pilot study was performed with all members of the CAIT / CATT. The objective of this pilot phase was threefold:

- to examine whether the guidance was clear and applicable,
- to examine whether the proposed procedure of deliberation on the outcomes of the ranking exercise was feasible and acceptable,
- to test an approach for involving patient groups in the completion of the summary of evidence tables.

The research process for the pilot testing was similar to the pre-testing process, although a few differences are noticeable:

- a first meeting was organised to explain the assignment in the context of the pilot study and to present the evidence. CAIT / CATT members were given the opportunity to ask questions
- during the second meeting both the methodology and the results were discussed:
- patient input for the evidence tables relating to “impact of disease on quality of life” and “inconvenience of current treatment” was collected for the case of Alzheimer’s disease. A qualitative approach was used (see paragraph 3.4.3.4).

The cases included in the pilot test were the same as those of the pre-test. However, the evidence tables were modified along the lines suggested during the pre-test study before they were distributed to the members of the CAIT / CATT.

The actions and outputs produced during the MCDA process are presented in Figure 9. The independent expert in the pilot case was a researcher from KCE, but could in practice be someone from the RIZIV / INAMI who is not a member of the CAIT / CATT.

Figure 9 – Flowchart of MCDA process

3.5.3 Statistical analysis

The MCDA process foresees discussions on the criteria with large variability between raters. Because the number of response categories was limited to 4 for each criterion, the variability in scores across raters was presented bulb charts.

After the first adaptation of individual scores following the discussions, the individual criteria scores were multiplied by the respective criteria weights. Individual total weighted scores were calculated for therapeutic need and for societal need for each disease. They were obtained by summing the weighted criterion scores for the criteria that fell under the therapeutic need domain and the societal need domain respectively.

The aggregates of the total weighted scores considered were the median and the mean. Two rank orders of diseases were created for each type of need (societal and therapeutic): one based on the median total weighted scores and one based on the mean total weighted scores.

No comparisons were made between the results of the pre-test and the results of the pilot test because the versions of the evidence tables were not the same. Changes in content as well as in form were made following the suggestions made by the participants in the pre-test study.

4 RESULTS

4.1 Pre-test study

4.1.1 Methodological issues

Methodological issues were discussed during the first meeting.

4.1.1.1 Scoring: minimal important impact

In the assessment of therapeutic need, each criterion is compared to a situation without the health condition. For example, quality of life with the condition and given current treatment is compared with quality of life without the condition; life expectancy with the condition and with current treatment is compared to life expectancy without the condition. The distinction between a score of 0 (no impact) and a score of 1, 2 or 3 (some, high, very high impact) depends on the clinical meaningfulness of a particular observed impact. As long as the observed clinical impact, even if positive, is below the threshold of meaningfulness, the score should be 0.

A challenge is to define the minimal important impact for several criteria, such as quality of life, inconvenience of current treatment and life expectancy. This definition can differ across contexts: the same impact can be important in one context and not important in another. For example, the results of the population survey showed that a negative impact of disease on quality of life is less important for the definition of need if patients are currently already in bad health than when patients are currently in relatively good health. Thus, depending on the co-morbidity of patients, the minimal important impact of a condition can differ and requires a different score.

There is no valid method to define a meaningful impact. Expert opinion combined with critical reflection will often be necessary. Experts could in this respect also be patients (e.g. for impact on quality of life and for inconvenience of current treatment). Especially in dubious cases, for which experience with the condition is needed to be able to judge the impact, the input from patients will be needed.

A practical solution to this problem could be to have a plenary discussion within the committee about the importance of objectively measured impact (e.g. by means of the functional scales) whenever large variations in the

scores assigned to a particular criterion are observed between the experts of the committee.

4.1.1.2 Evidence tables: content

The experts participating in the pre-test study made several remarks with respect to the content of the evidence tables:

- Often, the need lies not in the treatment of the disease in general but rather in a specific aspect of the disease. For example, while current insulin therapy is very effective for treating diabetes, the disease might still provoke complications that are less well met by current treatments, for instance painful neuropathy or blindness. When the focus is on one particular aspect of a disease (e.g. complication), a very precise definition of that aspect is necessary in the introduction, and all evidence should relate to that aspect, not to the disease in general. If it is on the disease in general, including all its complications, a clear description of the disease and its complications is necessary.
- The evidence tables should be completed or reviewed by a neutral person before they are presented to the committee members for scoring the diseases. Validity, reliability and objective presentation of the evidence should be reviewed. A similar approach as the one applied for the drug reimbursement committee, where independent experts of the RIZIV / INAMI review and assess the evidence submitted by the pharmaceutical companies, could be envisaged.
- The evidence should be presented in a comprehensible and harmonized way as much as possible. Better use of absolute figures instead of ratios, no abbreviations, etc. This comment has led the researchers to look for better templates for summarizing the evidence for the MCDA exercise. EVIDEM has published, in 2014, an instrument for synthesizing evidence, a possible template for highly synthesized by-criterion reports, and an instrument for assessing the quality of evidence. These instruments and templates were re-worked to fit the Belgian MCDA model. They could not be applied in the pilot test yet, but were developed for new submissions next year (see Appendix 4). Note that harmonization, as asked for by the pre-testers might not always be possible. In the ideal world, comparable data, measured by means of the same instruments, are available for all diseases, but this is not reality and still a decision has to be taken. The templates should give guidance on what kind of evidence to include (with quality standards). However, if this evidence is not available, other evidence must be allowed, as it it impossible to make a judgement based on nothing at all. Evidence that does not fit the guidance, could, however, be given less weight in the appraisal, in order to stimulate researchers to include the measures that comply with the requirements for the MCDA evidence tables in their primary research.

A revised version of the evidence tables was created, taking most of the comments into account (a few examples in Appendix 5). This version was subsequently used in the pilot study with the CAIT / CATT.

4.1.1.3 Applicability

The experts participating in the pre-test formulated a couple of caveats regarding the applicability of the appraisal of the evidence in real life:

- It should not be taken for granted that all committee members have an equal understanding of the jargon used in the evidence tables. The committee members might have different domains of expertise. Hence, utilities, particular quality of life measures, clinical outcome measures might not necessarily be familiar to all committee members.
- Similary, not all committee members might be equally used to interpreting scientific evidence, confidence or credibility intervals, hazard ratio's etc.

To solve these issues of applicability, it was suggested to create a vade mecum of specialised/difficult terms (see Appendix 6) and/or give an oral presentation of all evidence before starting the scoring, giving the committee members the opportunity to ask questions for clarification. This is the approach used by the Guideline Development Groups (GDGs). GDGs are panels of experts judging and discussing about the existing evidence related to interventions in order to decide whether or not to include the interventions in a clinical practice guideline. The discussions are preceded by a presentation of the evidence by a neutral person. It has been mentioned also that providing a vade mecum and giving an oral presentation may still be insufficient and more formal training sessions might be needed. This depends on the level of expertise in interpreting scientific evidence of the commission members.

One expert expressed his concern about the quantification of the appraisal process. He referred to the GDG approach, where an appraisal of the evidence is also made by a panel of experts, but in a more qualitative way. The EtD framework, for instance, contains categorical scales, which are used to structure the discussion within the guideline development panel (see paragraph 3.3.2). Panel members have to justify their appraisal for one criterion. This approach could also be envisaged in the Belgian MCDA model for assessing needs, i.e. committee members could be asked to justify their score. This will especially be important for criteria for which there is a large variation in scores amongst the committee members.

Other concerns raised related to the order in which the diseases are presented. The participants wondered whether the order could influence the scoring due to a benchmark or reference case effect. A possible solution to this problem is to first present all diseases and start scoring afterwards.

Participants were questioning the relevance of scoring the prevalence of a disease, as this is a purely quantitative criterion, for which there cannot be a discussion. However, this is only true if epidemiological data are available and reliable. If not, the prevalence has to be estimated based on other data and might be uncertain. It is important, as explained in paragraph 3.4.6 on dealing with uncertainty, to allow raters to make a comment on the uncertainty of certain data, so that this uncertainty can be taken into account after the ranking exercise. Leaving out the criterion takes away this opportunity and then the uncertainty regarding prevalence risks to be overlooked.

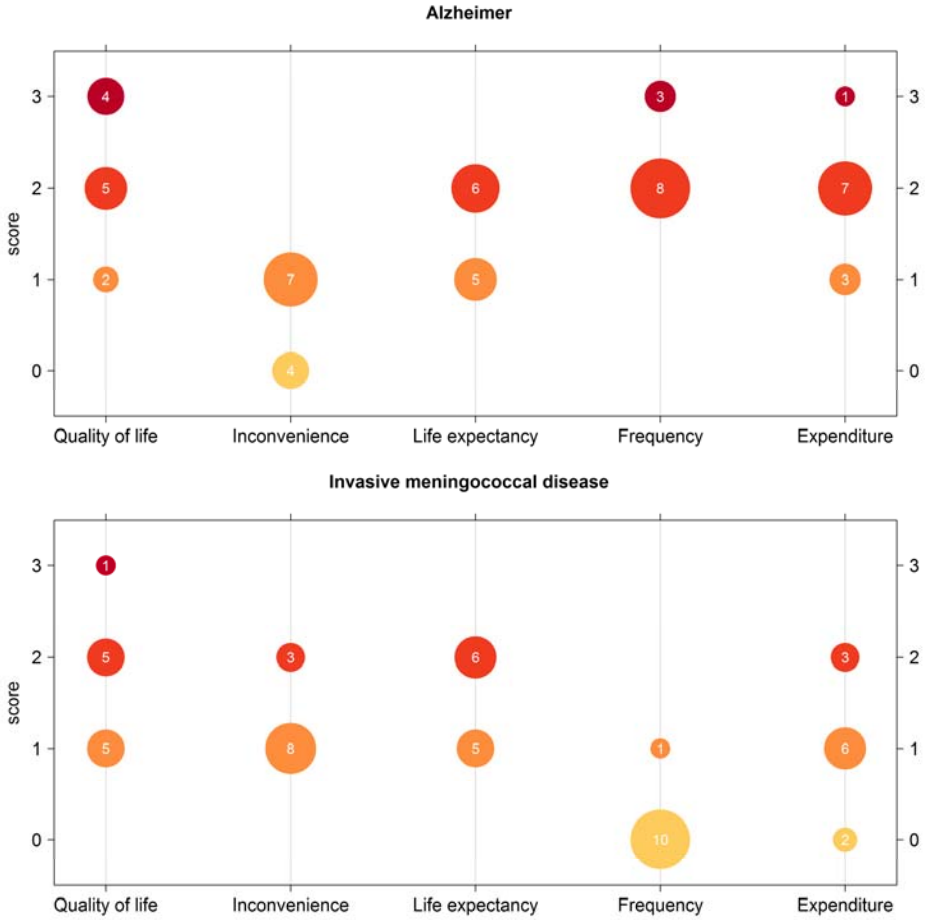
Participants recommended to change the response categories for the criterion “quality of life” to “no impact”, “some impact”, “moderate impact” and “high impact”. The distinction between high and very high impact was considered hard to make, while the distinction between “moderate” and “high” is more clear.

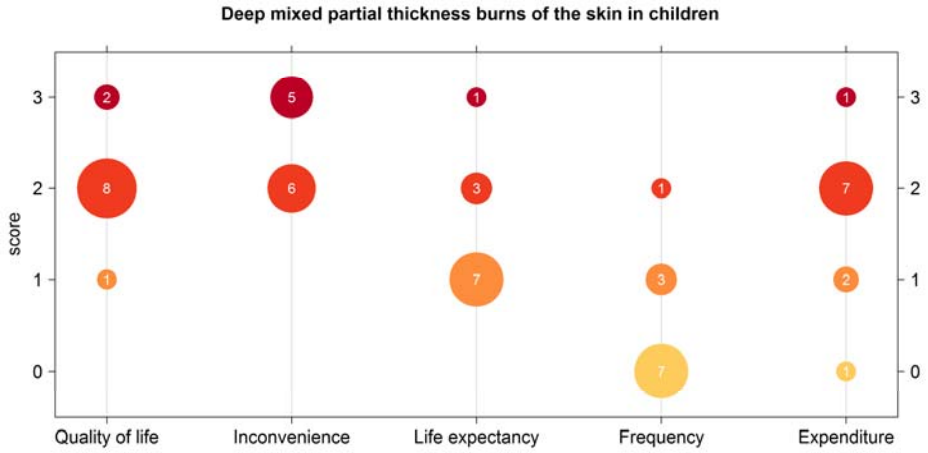
Finally, suggestions for clarifications were made regarding the content of some evidence tables. For example, it should be made clear that if the impact of meningococcal disease is to be judged, it should be judged for the entire population affected by meningococcal disease and averaged out, and not only for these with severe sequelae.

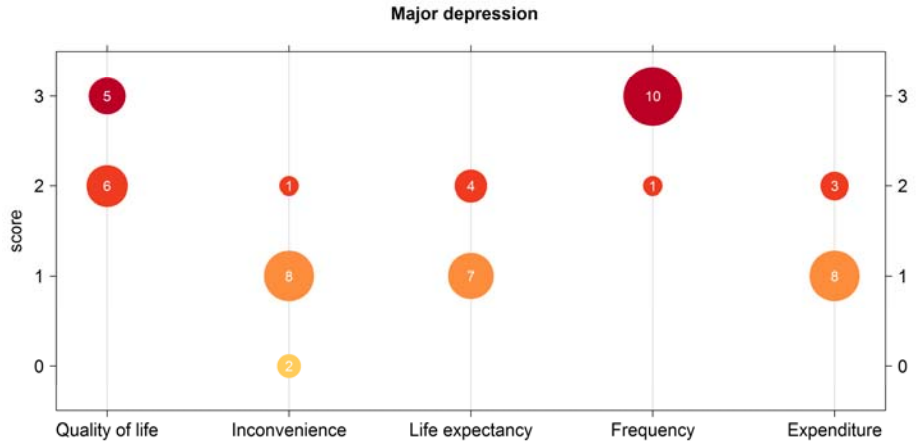
4.1.2 Variation in scores

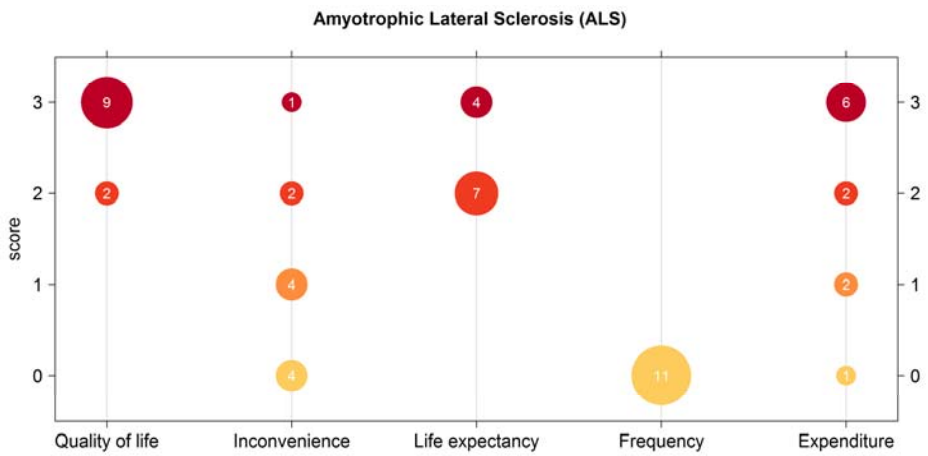
During the second meeting, an overview of the scores for all criteria in all diseases and their variation between raters was presented (Figure 10).

Figure 10 – Distribution of scores per criterion and disease in pre-test study – first round

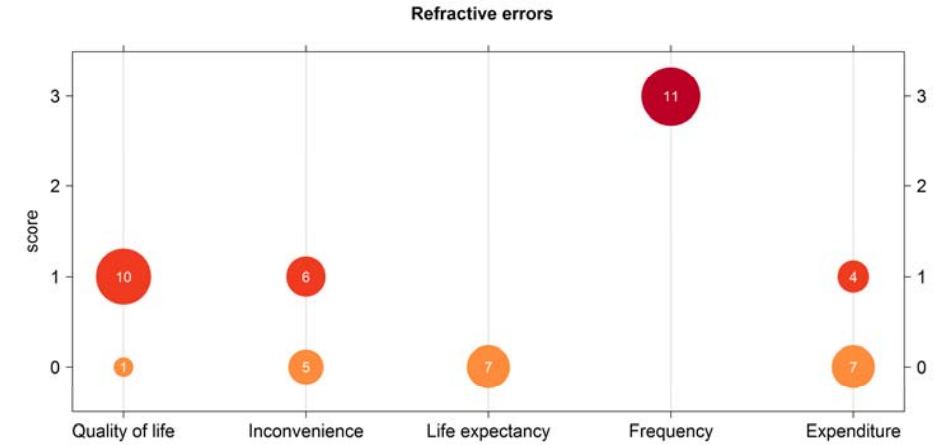


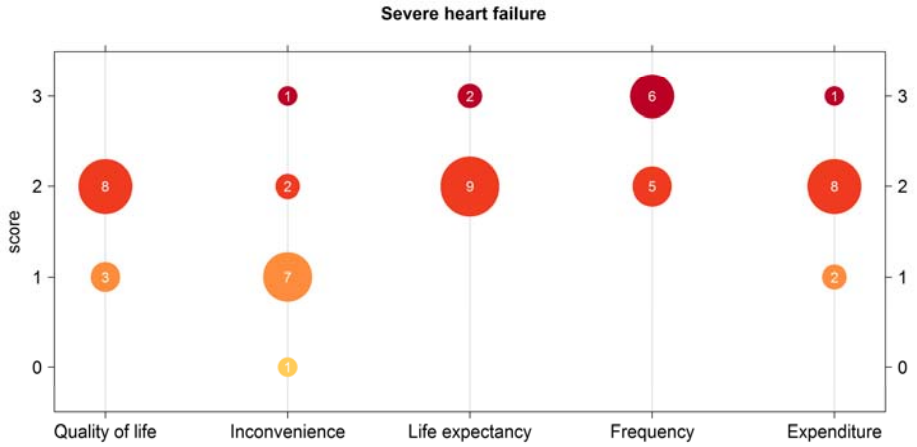

Deep mixed partial thickness burns of the skin in children
score
Quality of life
Inconvenience
Life expectancy
Frequency
Expenditure


Major depression
score
Quality of life
Inconvenience
Life expectancy
Frequency
Expenditure


Amyotrophic Lateral Sclerosis (ALS)
score
Quality of life
Inconvenience
Life expectancy
Frequency
Expenditure


Mesothelioma
score
Quality of life
Inconvenience
Life expectancy
Frequency
Expenditure





The colour of the bulbs indicate the level of severity of a disease on a particular criterion: the darker colours represent higher severity. The number of bulbs per criterion indicate the variation in scores between respondents: three bulbs means that at least three different scores were given by the respondents. The size of the bulbs represent the number of respondents that gave a particular score: the larger the bulb, the higher the number of respondents that gave that score. The number of respondents is each time shown in the bulbs.

Some criteria were given the same score by all respondents. This was mainly the case for the frequency criterion, for which the levels are defined in numbers and hence little discussion about the score is possible if valid prevalence figures are available.

Discussion was held about the following criteria:

Severe heart failure:

- Impact of severe heart failure on quality of life
- Inconvenience of current treatment for severe heart failure
- Impact of severe heart failure on life expectancy
- Frequency of severe heart failure
- Severe heart failure related public expenditures per patient

ALS:

- Impact of ALS on quality of life
- Inconvenience of current treatment for ALS
- Impact of ALS on life expectancy
- Frequency of ALS
- ALS-related public expenditures per patient

Mesothelioma

- Impact of mesothelioma on quality of life

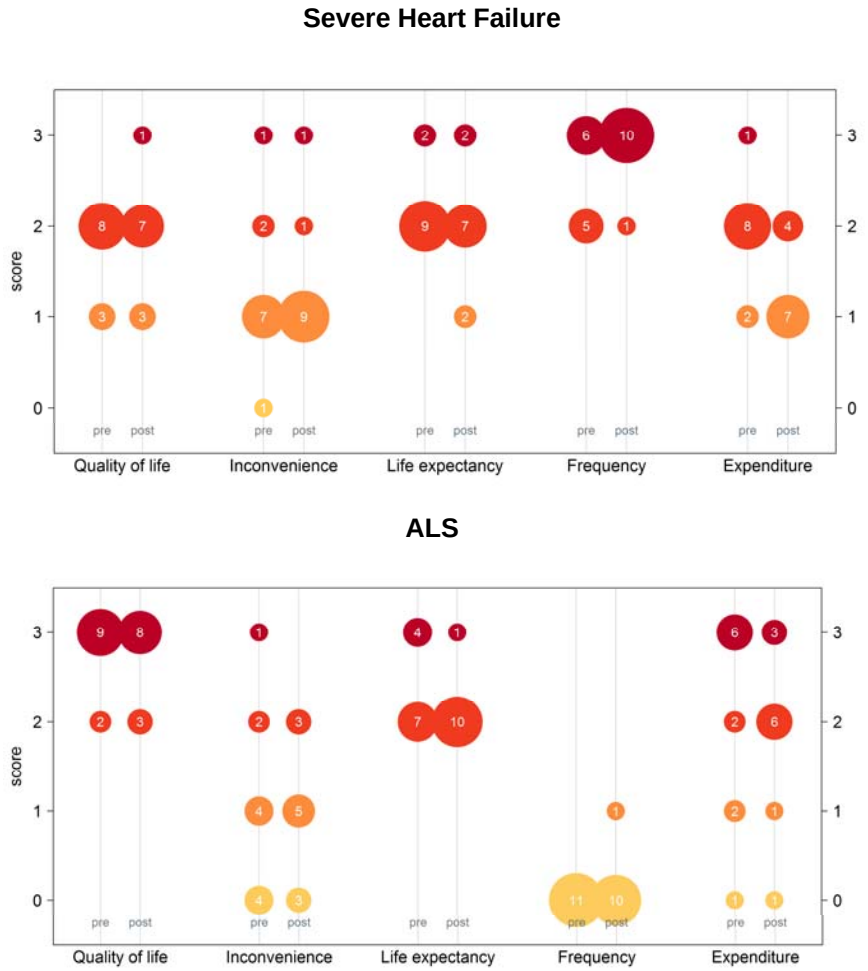
Deep mixed partial thickness burns of the skin in children

- Impact of deep mixed partial thickness burns of the skin in children on quality of life
- Inconvenience of current treatment for deep mixed partial thickness burns of the skin in children

- Impact of deep mixed partial thickness burns of the skin in children on life expectancy
- Frequency of deep mixed partial thickness burns of the skin in children
- deep mixed partial thickness burns of the skin in children related public expenditures per patient

After the discussions, a second round of scoring was performed. The comparison between the distribution of scores for the selected criteria before and after the discussion are presented in Figure 11.

Figure 11 – Distribution of scores in pre-test study – comparison pre- and post-discussion



Mesothelioma*

** Only discussion about impact of the disease on quality of life*

Deep mixed partial thickness burns of the skin in children

All the participants changed at least one score following the discussions. For 10 criteria, the scores became more homogeneous between participants after discussion. In six criteria, the scores became more heterogeneous after discussion. These were impact of heart failure on quality of life and life expectancy, frequency of ALS and ALS-related public expenditures and inconvenience of treatment for deep partial thickness burns of the skin in children and impact of these burns on life expectancy.

4.1.3 Ranking of therapeutic need

The rankings of diseases in terms of therapeutic need, based on the mean and median total weighted scores before discussion, are presented in Table 12. Two observations can be made.

First, the ranking based on the median weighted score and the ranking based on the mean weighted score are identical.

Second, more ties are observed with the use of median total weighted scores (e.g. depression and severe heart failure, meningococcal infection and Alzheimer's disease) than with the use of mean total weighted scores.

Table 12 – Ranking therapeutic needs based on median total weighted score and based on mean total weighted score respectively –before discussion

Rank	Disease	Median score	Disease	Mean score
1	Mesothelioma	2,57	Mesothelioma	2,45
2	Burns in children	1,86	Burns in children	2,16
3	ALS	1,71	ALS	1,97
4	Major depression	1,57	Major depression	1,64
5	Heart failure (shared 4th place)	1,57	Hart failure	1,60
6	Meningococcal infection	1,43	Meningococcal infection	1,47
7	Alzheimer (shared 6th place)	1,43	Alzheimer	1,43
8	Refractive errors	0,86	Refractive errors	0,63

It is dangerous to draw conclusions from pre-post discussion comparisons, as not all diseases were discussed during the second meeting. Hence, we cannot be sure whether the weighted scores of the other diseases would also not have changed following the discussion. However, we present it for completeness and transparency about the process and its outcomes.

After the discussion and second scoring round, the ranking changed only if the ranking was based on the median weighted score. This was due to the change in the median weighted score for ALS.

Table 13 – Ranking therapeutic needs based on median total weighted score and based on mean total weighted score respectively – after discussion

Rank	Disease	Median score	Disease	Mean score
1	Mesothelioma	2,57	Mesothelioma	2,61
2	ALS	2,00	Burns in children	2,00
3	Burns in children	1,86	ALS	1,90
4	Major depression	1,57	Major depression	1,64
5	Heart Failure (shared 4th rank)	1,57	Heart Failure	1,61
6	Meningococcal infection	1,43	Meningococcal infection	1,47
7	Alzheimer (shared 6th rank)	1,43	Alzheimer	1,43
8	Refractive errors	0,86	Refractive errors	0,63

After the discussion, the median score for ALS changed, as a consequence of which the ranking based on the median weighted score and the mean weighted score was no longer identical.

Compared to the ranking before the discussion, no differences can be observed in the ranking based on the mean weighted score.

4.1.4 Ranking of societal need

The rankings of diseases in terms of societal need, based on the mean and median total weighted scores before discussion, are presented in Table 14. There were some differences between the ranking based on median total weighted scores and the ranking based on the mean total weighted scores. ALS and major depression switched in ranking, as were refractive errors and burns in children. As in therapeutic need, we also observe a tie when median scores and not when mean scores are used.

Table 14 – Ranking societal needs based on median total weighted score and based on mean total weighted score respectively - before discussion

Rank	Disease	Median score	Disease	Mean score
1	Heart failure	2,35	Heart failure	2,13
2	Alzheimer	2,00	Alzheimer	1,98
3	ALS	1,95	Major Depression	1,85
4	Major Depression	1,70	ALS	1,42
5	Burns in children	1,30	Refractive errors	1,29
6	Refractive errors	1,05	Burns in children	1,28
7	Meningococcal infection	0,65	Meningococcal infection	0,74
8	Mesothelioma (shared 7th rank)	0,65	Mesothelioma	0,71

The impact of the discussion and second scoring round on the ranking was higher for societal need than for therapeutic need. Alzheimer and heart failure switched places after the discussion and heart failure became a lower societal need than Alzheimer. Note that Alzheimer was not discussed during the second meeting, so we can only conclude that the discussion reduced the weighted score of heart failure. ALS, which was also discussed during the second meeting, also received a lower weighted score after discussion, thereby lowering its place in the ranking. Finally, deep mixed partial

thickness burns in children obtained a higher weighted score after discussion and hence increased in the ranking.

Table 15 – Ranking societal needs based on median total weighted score and based on mean total weighted score respectively – after discussion

Rank	Disease	Median score	Disease	Mean score
1	Alzheimer	2,00	Alzheimer	1,98
2	Heart failure	1,70	Heart failure	1,90
3	Major Depression (shared 2nd rank)	1,70	Major Depression	1,85
4	Burns in children	1,65	Burns in children	1,46
5	ALS	1,30	ALS	1,33
6	Refractive errors	1,05	Refractive errors	1,29
7	Meningococcal infection	0,65	Meningococcal infection	0,74
8	Mesothelioma (shared 7th rank)	0,65	Mesothelioma	0,71

4.2 Pilot study

4.2.1 Input from representatives of patients with Alzheimer's disease

From the pilot interviews with representatives/informal caregivers of patients with Alzheimer's disease the main findings are about how patients' needs change in function of the stage of the disease and the non-medical character of Alzheimer's patients' needs.

In an early stage patients realise things go wrong, sometimes they suspect Alzheimer's disease without being diagnosed, but they try to hide it from the outside world as long as possible. By consequence they become more on their own, turn inside and get more isolated. Also at this stage the disease causes all kinds of misunderstandings and conflicts with spouses, children, colleagues etc. In a more advanced stage other challenges come to the fore, especially in the interaction between the patient and the informal caregiver. For example, Alzheimer patients get locked-up in their past. It is of no use to discuss about things that went wrong in the recent past, as they simply cannot remember it happened. It is far better to accompany them in their world and confirm their experiences of the world instead of constantly trying to correct them or confront them with the real world.

Alzheimer patients' needs are not so much medical, but rather social, emotional and practical. They need relatives and friends being patient and compassionate. They need proximity of beloved ones. They need to feel safe and secure. They like quietness and dislike noisy places. They need to be treated with respect, also in how they are talked to. In addition Alzheimer patients' need practical help. In Flanders a lot of practical help is available, but procedures take a lot of energy of informal caregivers and take too much time. The progression of the disease is sometimes faster than procedures, making the demanded support or devices superfluous once they arrive. Many people do not know where to find support and practical information. There is a need for one coordinator, a social worker, who guides people through the system and informs them about their rights (e.g. financial allowances). Also it would avoid people needing to tell their story over and over again to several carers.

4.2.2 Input from patients with severe burns

In contrast to Alzheimer patients, patients with severe burns have a lot of medical needs, but they also need financial, psycological and social support.

From discharge onwards and mostly for the rest of their lives, treatment consists mainly of physiotherapy, wound care and wearing pressure garments. These treatments bring along several inconveniences. All three treatments take a lot of time, which makes it more difficult for burn patients to participate in outdoor activities such as going to school or work. Also the care appointments (e.g. long physiotherapy sessions) limit their participation in normal social activities.

In addition, treatments are expensive, which may lead to financial problems, especially if patients are unable to return to work.

In addition, aftercare for severely burned patients is characterized by several physical complaints often resulting in substantial functional loss. The functional disabilities are associated with boredom (they cannot do anything, but do not feel sick), loneliness, shame (especially in case of visible scars),

difficulties accepting their physical limitation, being dependent on others, and appearance.

Functional disabilities give way to the need for psychological and social support. Patients fear the future. They report uncertainties regarding how their body functions will evolve, in a positive way or further down, especially in interaction with normal aging processes. Burn patients experience difficulties in coping with their changed body because of problems with the self-image, identity and the reactions of others. Some patients report stigmatizing behaviour or remarks from others.

4.2.3 Methodological issues

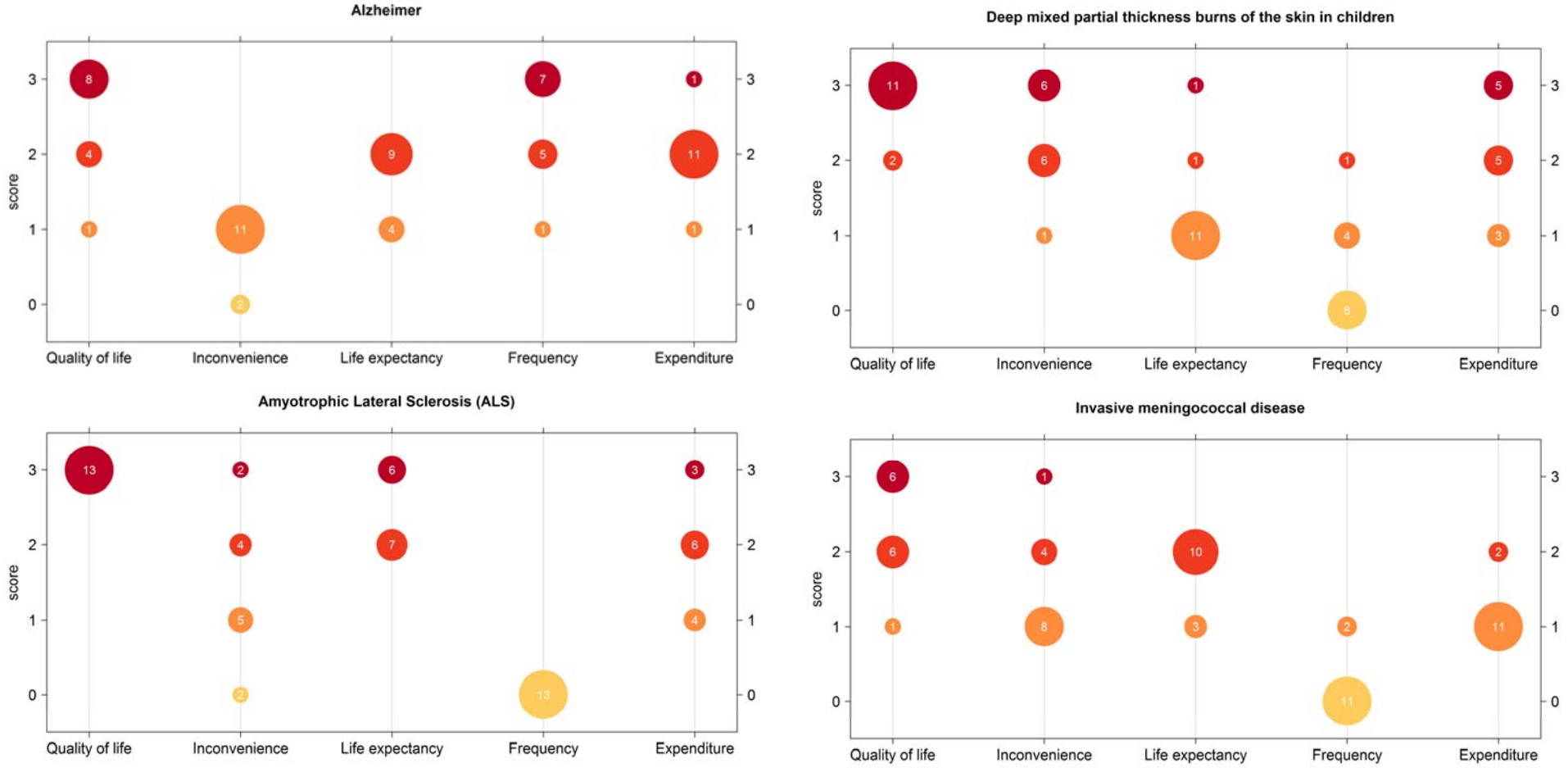
Seven CAIT / CATT members, of which 6 participated in the pilot test of the MCDA tool, were present during the second meeting. The first part of the meeting was devoted to the discussion of the methodology. The following methodological issues were raised:

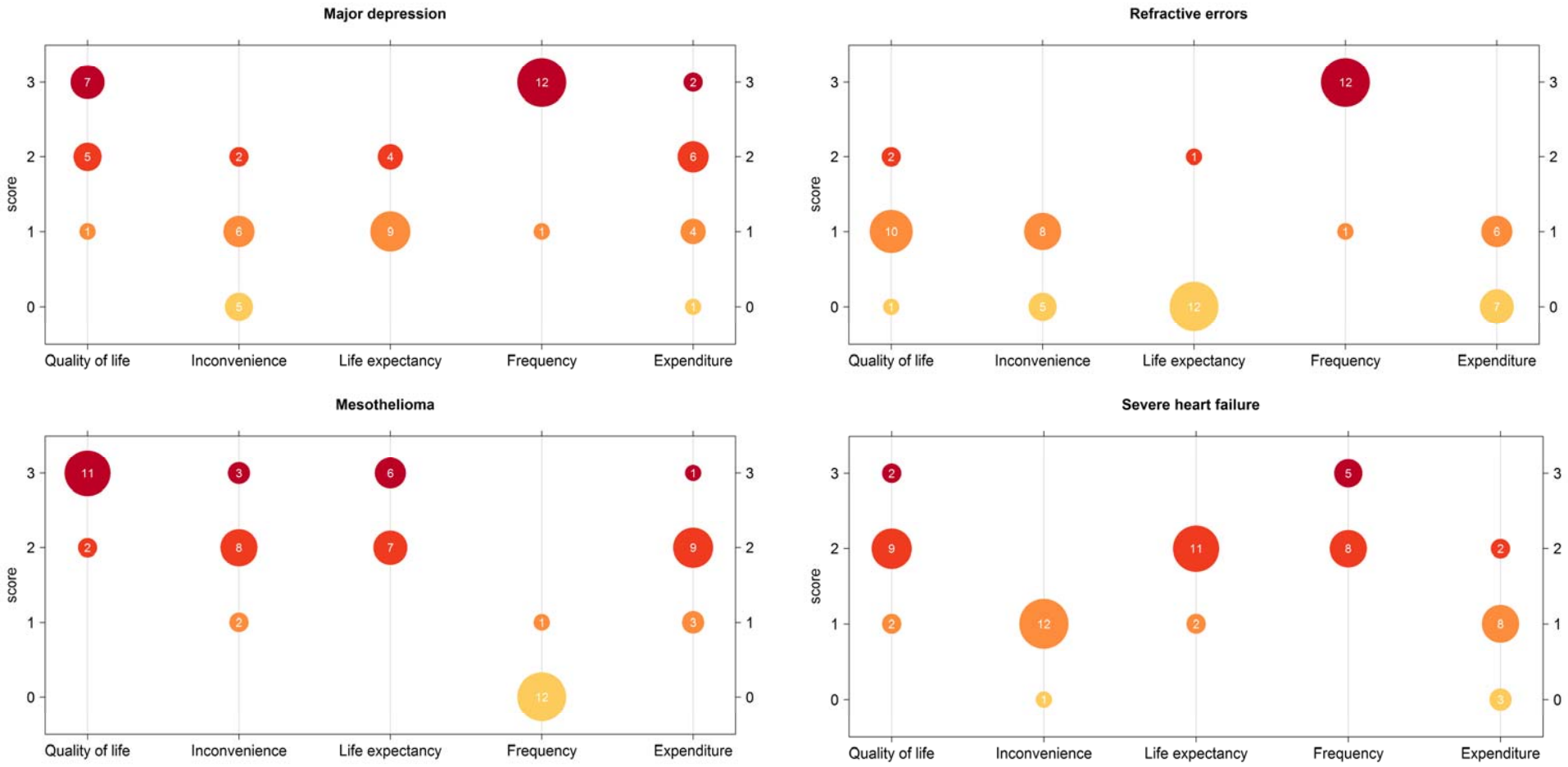
- For the case of invasive meningococcal disease, two participants found it difficult to give one overall score for each criterion, because the disease has an acute phase which can be life-threatening, and a chronic phase with possible sequalae. The evidence provided encompasses both phases.
- The level of evidence is not clear from the evidence tables. Also, study results are mostly presented as ranges rather than as exact figures. This is not surprising, but renders the scoring more difficult. The suggestion of KCE to involve an independent expert to make a critical assessment of the submitted evidence was considered useful.
- The role of patients and patient organisations should be worked out. One committee member highlighted the fact that the appraisal by the committee members is probably biased by personal experiences (or absence thereof) with the conditions under consideration. He considered it useful to involve patients in the process for all diseases. The precise modalities should be worked out (e.g. whether the patients should be able to vote as well, or just share their experiences with their disease without scoring). Qualitative research to capture the perspective of the patients should occur according to standards for good qualitative research.
- One participant missed the ability to give scores between the 4 pre-defined levels and would have appreciated a scale with more levels.
- The participants highlighted the importance of providing unambiguous definitions of the criteria. For example, for one respondent it was at a certain moment in time not clear that disease-related public expenditures referred to all types of public expenditures and not to costs for the patient. Also the definition of “inconvenience of current treatment” needs to be made very clear, as it is expected that the evaluation of “impact of the disease on quality of life” will influence the score for inconvenience and vice versa. If a treatment is very inconvenient, it is likely to impact upon the quality of life of patients. However, with clear and unambiguous definitions this should be avoidable. Although definitions were provided at the front page of the pilot package, we can conclude from this feedback that it is important to get used to the definitions. In the beginning, this may require repeating the definitions on the pages relating to the criteria. It can be expected, though, that after some time, after several scoring exercises, the committee will get used to the definition and will no longer need the definitions. Nevertheless, for newcomers, it may still be useful to repeat them regularly.
- The evidence tables should avoid suggestive language at all times. For example, “patients support the treatment generally well” should be avoided.
- The CAIT / CATT members present considered the MCDA exercise useful. Earlier in 2016 they had performed as similar exercise, with mainly the same criteria but without applying the weights from the general public and the standardised evidence tables. They also found little changes in the ranking following the discussions. The rankings were also considered to make sense. The approach was very much appreciated.

4.2.4 Variation in scores

An overview of the scores for all criteria in all diseases and their variation between raters is presented in Figure 12.

Figure 12 – Distribution of scores per criterion and disease in pilot study – first round




Major depression
Refractive errors
Mesothelioma
Severe heart failure
score
Quality of life
Inconvenience
Life expectancy
Frequency
Expenditure

The presence of only 6 members of the 13 who participated in the pilot test on the second meeting reduced our ability to measure the impact of the discussion on the variation in scores and the ranking.

Our proposal to have discussions about the criteria for which at least three different scores were assigned by the group was rejected by the committee members present. Variability in scores was not considered problematic, especially when the variability could be attributed to one or two respondents having another score than the majority of the respondents.

Nevertheless, a discussion was held about criteria for which there was obvious disagreement in scores and the whole range of possible scores was used. These were:

- Burns-related public expenditures
- Inconvenience of current treatment for ALS
- Major depression-related public expenditures per patient

For burns-related public expenditures, one respondent explained that he had considered patient costs only and had, therefore, given a low score (1=minor impact on public expenditures per patient). Others had considered the long-term costs over a lifetime from the societal point of view, and therefore gave a higher score.

Regarding inconvenience of current treatment for ALS, the variability in scores seemed to be explained by the implicit weight given by the respondents to the different aspects of treatment that may occur. For example, respondents who thought about the non-invasive mechanical ventilation that some patients may need when their disease advances gave a higher score than respondents who considered the treatment for the group of patients who do not need ventilation or considered that not all ALS patients will need ventilation. A similar conclusion was drawn for the variability in scores for the impact of invasive meningococcal disease on quality of life and the inconvenience of treatment for invasive meningococcal disease. Depending on whether respondents implicitly put the highest weight on the acute phase relative to the chronic phase, they gave higher, respectively lower, scores; or, if respondents mainly weighted the severe sequelae in the chronic phase (e.g. amputations), they are also inclined to give a higher score. None of the participants changed their score for inconvenience of current treatment for ALS following the discussion.

For major-depression-related public expenditures, the issue was that the cost figures were not clear-cut. Committee members suggested that more information should be collected from the RIZIV / INAMI databases when this criteria needs to be scored.

After the discussions, the respondents were given the opportunity to change their score for these two criteria. One person changes his score for major-depression-related public expenditures from 0 to 2, another person changed his score for burns-related public expenditures from 1 to 2.

4.2.5 Ranking of therapeutic need

The ranking of therapeutic needs before the discussion on the criteria is presented in Table 16.

Table 16 – Ranking of therapeutic needs in pilot study before discussion

Rank	Disease	Median weighted score	Disease	Mean weighted score
1	Mesothelioma	2,57	Mesothelioma	2,46
2	Burns in children	2,29	Burns in children	2,42
3	ALS	2,14	ALS	2,26
4	Meningococcal infection	2,00	Meningococcal infection	1,90
5	Alzheimer's disease	1,86	Alzheimer's disease	1,69
6	Heart Failure	1,57	Depression	1,57
7	Major Depression	1,43	Heart Failure	1,52
8	Refractive errors	0,86	Refractive errors	0,75

This ranking was considered to be intuitively sound. Referring to the results of the citizens labs of the KBF, one participant noted that for the chronic conditions the impact of the disease on the quality of life of patients' environment (family, caregivers, friends) can be higher because of the longer duration of the disease. This would be an argument to put Alzheimer higher on the list. Ironically, this impact cannot be an argument for severe life-threatening diseases (e.g. mesothelioma), because the impact is high but over relatively shorter period of time.

As nobody changed their scores for therapeutic-need related criteria after discussion, the ranking remained as it was before the discussion.

4.2.6 Ranking of societal need

The ranking of societal need before the discussion on the criteria is presented in Table 17.

Table 17 – Ranking of societal needs in pilot study before discussion

Rank	Disease	Median weighted score	Disease	Mean weighted score
1	Alzheimer's disease	2,35	Alzheimer's disease	2,16
2	Major Depression (shared 1nd rank)	2,35	Major Depression	2,10
3	Heart failure	1,35	Burns in children	1,56
4	Burns in children	1,30	Heart failure	1,43
5	ALS (shared 4th rank)	1,30	Refractive errors	1,30
6	Mesothelioma (shared 4th rank)	1,30	ALS	1,25
7	Refractive errors	1,05	Mesothelioma	1,23
8	Meningococcal infection	0,65	Meningococcal infection	0,80

An observation from the pre-test study that is confirmed here is that the use of median weighted scores leads to several ties.

The participants agreed that this ranking of societal needs was not very surprising. A short discussion was held about the place of mesothelioma. This ranks relatively low. The explanation is that the cost is now relatively low because there is no good treatment. However, the therapeutic need is very high in this disease. The discussion continued by highlighting that for other diseases it could be the other way around. For example, in Hepatitis C, a very effective but very expensive treatment is now available. Therapeutic need then lowers but the need for a less expensive treatment increases (societal need).

Because of these different perspectives and possibilities, the distinction between therapeutic and societal need was considered very relevant and important. Other examples were mentioned beyond those presented in the pilot study where the therapeutic need may be very high, but the societal need may be low and vice versa.

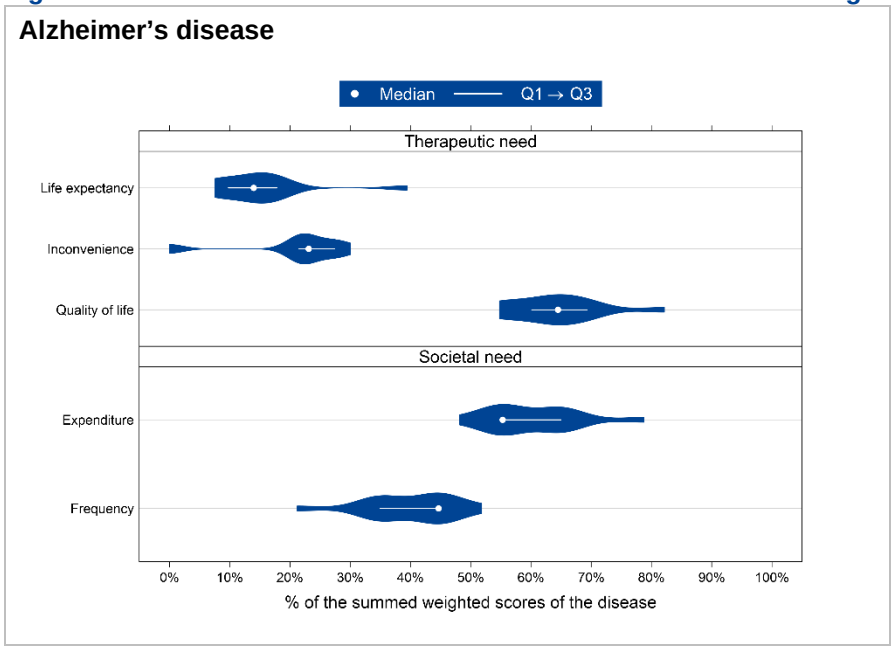
After discussion and the changes in scores of two participants, the ranking for societal need did not change. The median weighted scores for major depression and burns in children remained the same. The mean total weighted scores changed from 2.10 to 2.20 for major depression and from 1.56 to 1.61 for burns in children. None of these changes impacted upon the ranking.

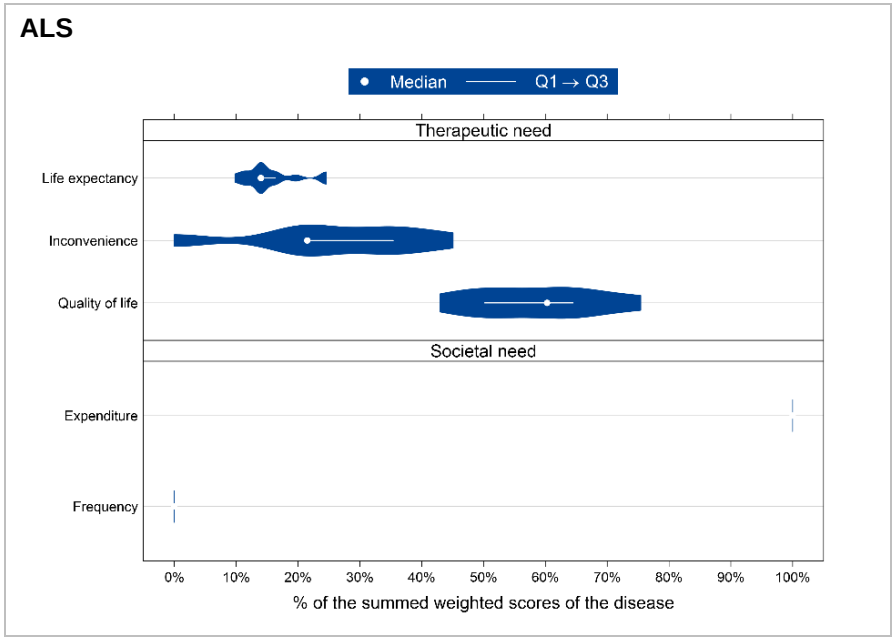
4.2.7 Contribution of each criterion to the total score

The contribution (in %) of each criterion to the total weighted score of each disease that is used for ranking the diseases is presented in Figure 13. The figure shows the distribution of the contributions amongst the participant. For each participant the contribution of his weighted score for each criterion in his individual total weighted score for the disease was calculated and plotted. This gives the distribution.

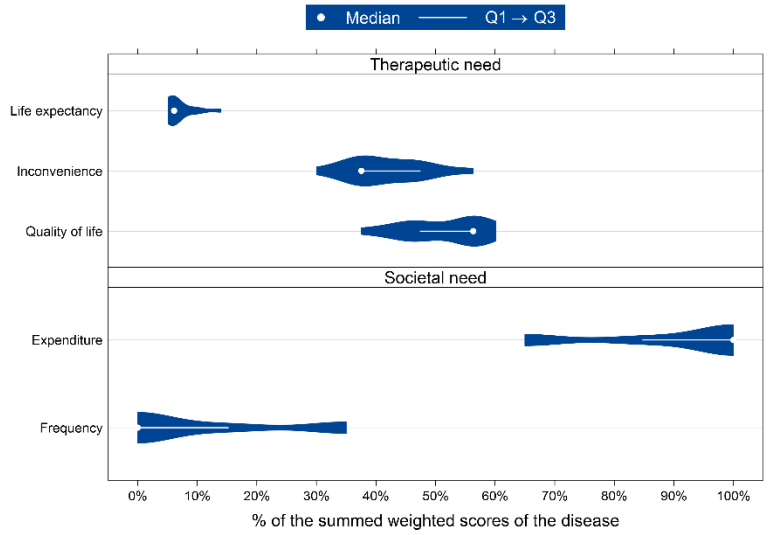
This information can be useful to assess the face validity of the results. This useful suggestion was given by one of the validators of this report, after the pilot study was finalised. Therefore, we could not assess the perceived usefulness of this information and presentation for the members of the CAIT / CATT. It seems worthwhile to test its usefulness in real life.

Figure 13 – Contribution of criteria to overall score used for ranking

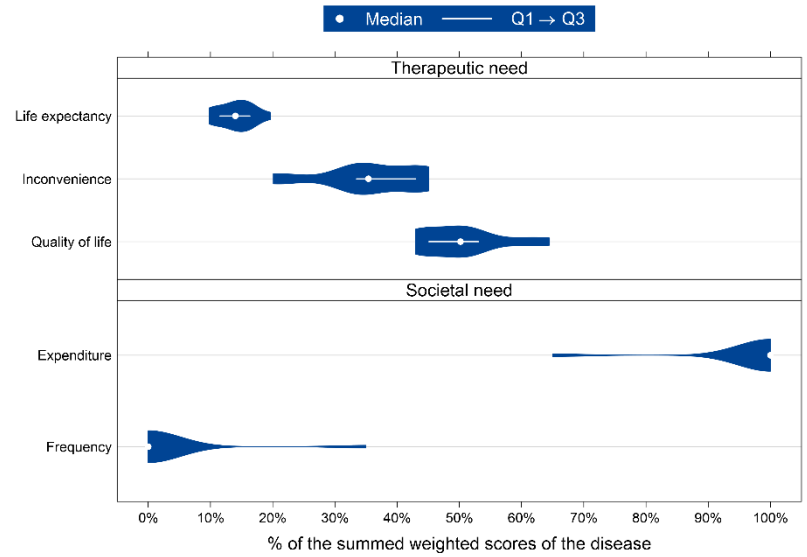




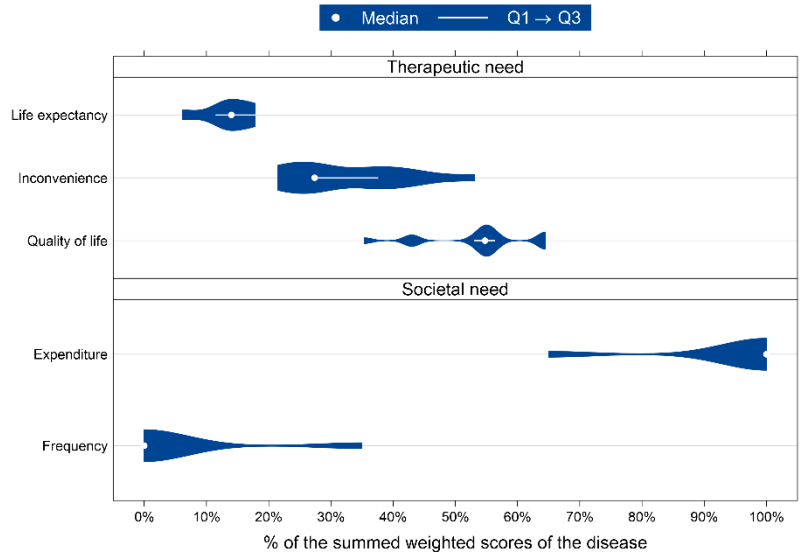
Burns in children



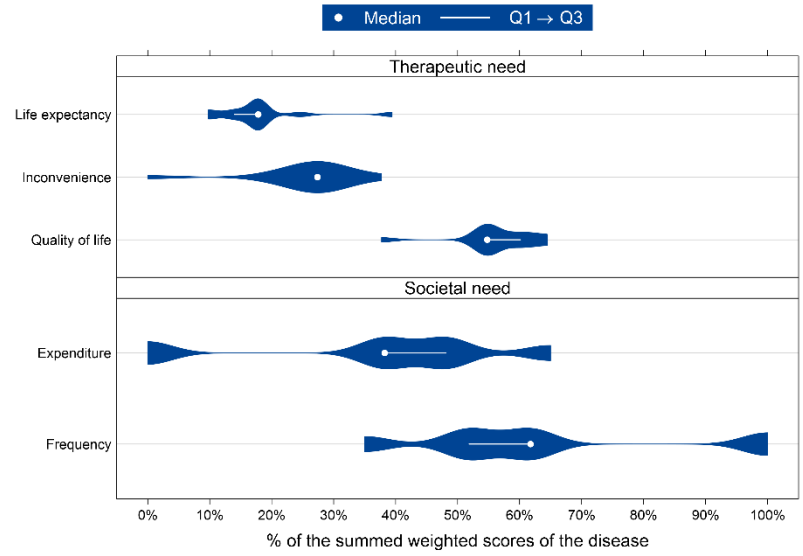
Mesothelioma



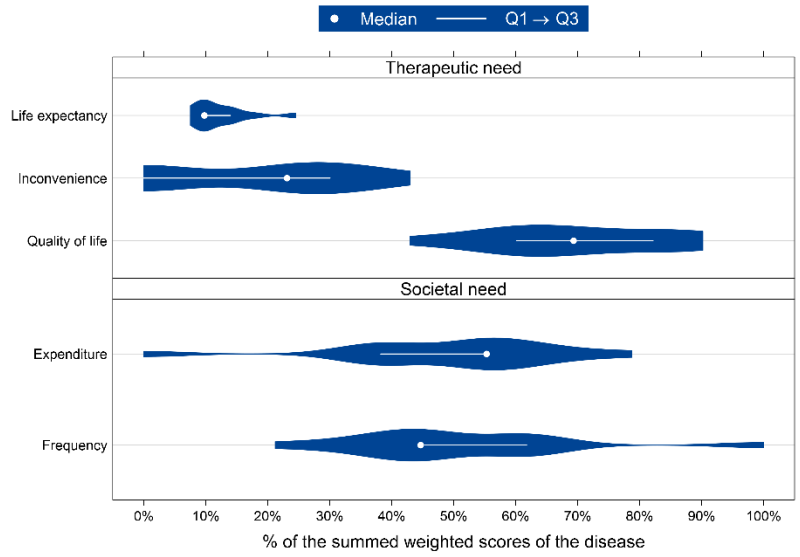
Meningococcal infection



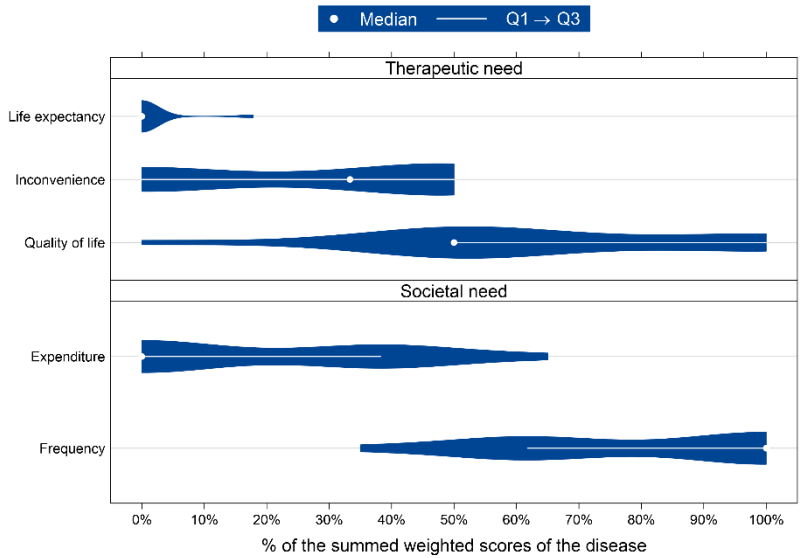
Heart Failure



Major Depression



Refractive errors



5 DISCUSSION

The appraisal of evidence on the impact of a disease of several criteria is a subjective process. Variations in scores between raters may be legitimate. MCDA is not going to resolve disagreements between decision makers in their appraisals of a problem. The merit of the MCDA system is that it provides an approach that encourages transparency and an explicit accounting of the judgments involved. The endeavour of this report should be considered in the light of this objective: to increase the accountability for reasonableness of decisions about the ranking of needs.

5.1 Methodological issues

The current study describes the development and testing of *a* MCDA approach to ranking therapeutic and societal needs. A number of methodological choices are made, which could be challenged.

First, the criteria included in our MCDA are generic, i.e. applicable to several diseases. This is necessary for the purpose of ranking needs, which relate to a variation of diseases and conditions. For decisions regarding the reimbursement of health interventions for a particular disease (the part of the MCDA framework not tested in this study but presented in report 147[2]), criteria could be made disease-specific. Examples have been described in literature (e.g. Angelis and Kanavos 2016[20]). In these approaches, relevant criteria are defined case by case, i.e. *de novo* for each disease under consideration. The advantage of this is that the appraisal is more specific for the diseases under consideration. The disadvantage is that the repeated selection of the relevant decision criteria for every new indication may be cumbersome and time-consuming. Moreover, it precludes the use of generic weights and hence needs to be combined with an approach deriving the criteria weights each time another disease is considered, thus adding to the intensity and time-consumption of the decision makers or advisory committees. For group wise revisions within a therapeutic class, this may be worthwhile. For case-by-case reimbursement decisions, this may not be feasible within the time constraints committees have for making a decision or formulating an advice.

The current model does not allow for the inclusion of additional criteria on an ad hoc basis, as the weights are determined from the general public only for the criteria in the model. When new criteria are introduced, their relative importance compared to the other criteria needs to be established. It is not feasible to do this in the general public each time a decision must be taken. Thus, when this approach is followed, it will include the decision makers' preferences, which are not necessarily identical to those of the general public.

Second, the derivation of the weights is done separate from the derivation of scores. The weights were obtained from previous research in the general public, the scores were obtained directly from the committee members. The rationale for using weights from the general public was to improve the accountability for reasonableness of decision makers towards the public. However, it could be argued that the decision makers are representatives of the general public and hence their preferences reflect those of the general public. We would counter-argue, though, that these decision makers are hardly aware of the preferences of the general public and benefit from research being done on this topic in the general population.

More sophisticated approaches exist where the weighting and the scoring are both obtained from the committee members and value functions are derived from the committee members.[20] Value functions can be derived directly or indirectly. The use of an ordinal scale is a direct way to derive value scores. An indirect way would be to ask a series of questions to the committee members to uncover their preferences by considering differences in the criteria scale and their relation to the value scale.[20] Members first value differences between possible outcome values for each decision criterion. Then, they value the strength of their preference for each difference compared to each other difference in outcome for that criterion. The outcome values can be quantitative or qualitative and are based on the existing evidence with respect to the outcomes of the different options under consideration. During the entire process the committee members do not have information on the identity of the different options. They score and determine the relative importance of each of the criteria independent from the interventions considered. This avoids bias in the derivation of the value function. Once the value function is derived, the weights for the different criteria are obtained by asking to compare the relative importance of each criterion for appraising the options and finally the result of the appraisal rolls out of the system.

Third, in contrast to examples in literature, we proposed to appraise therapeutic and societal needs *given currently available treatments.* This

means that in the appraisal of the therapeutic needs, for instance, the severity of the disease and the shortcomings of the current diseases are merged in one criterion. It could be argued that this may lead to misconceptions in decision makers, because some 'current treatments' are not addressing patient needs very well. This is due to the way in which healthcare interventions are developed and evaluated. Some treatments may make patients even worse off. In these cases, the needs may arise from improper treatment rather than from the disease as such. This applies to both therapeutic and societal needs, where for therapeutic needs the quality of life and life expectancy may be impaired due to the inappropriate treatment and for societal needs public expenditures may be due to the inappropriate treatment. By combining both current treatment and disease in one criterion, the distinction between what is due to the disease and what is due to the (inappropriate) treatment may not be clear for the decision maker.

However, we would counter-argue that the fact that some treatments make patients worse off is an undesirable situation that should be avoided. It could be avoided if the split approach proposed in this and previous KCE research, i.e. to first assess needs and only afterwards appraise the added value of interventions based on the identified needs, is systematically applied in the future. In the long run, this will preclude such undesirable situations. Moreover, it can be expected or at least hoped that these cases are only a minority in our current healthcare system and that most treatments are highly effective and meeting patient needs. We would therefore argue that it is extremely important to take the effectiveness of current treatments into account when assessing the needs of patients, to avoid wrong conclusions. For example, diabetes is a very severe disease, but with current insulin treatments, an important need of diabetes patients is met, so these particular needs will not be appraised very high. However, some aspects of diabetes are much less treatable or preventable, hence patients still have particular needs. These remaining needs also need to be defined and appraised, in the light of the absence of an effective treatment.

In our approach the limitations of current treatments are included in the appraisal of the impact of disease on quality of life and life expectancy given current treatment and in the criterion inconvenience of current treatment. An alternative approach could be to make a separate criterion for disease severity (in absolute terms, i.e. when no treatment is given) and one for 'limitations of current treatments in addressing the needs as defined in the description of the disease severity'. If the evidence on the disease needs without treatment is available, this approach would allow to clarify the state of knowledge on the disease and the needs that are not met by the current treatment, and rank the unmet needs based on this. Of course, this assumes another approach to the weighting, e.g. such as the one presented in the second discussion point above. This will be more time consuming. Another issue with this approach is that the evidence on the disease severity 'without treatment' will often not be available because most patients do receive some treatment. Although it can be argued that absence of evidence should not be a reason to develop a better model, the question is how applicable this theoretically better model then will be in real life.

Fourth, the scale we used for scoring the evidence on the different criteria for each disease is simple. It could be argued that a scale with more response options would be better, as it would allow more differentiation between several diseases. On the other hand, too many response options might make the scoring task more difficult and not necessarily improves the validity of the scores. There is no gold standard for scoring the impact of a disease on different criteria. The feasibility of the simple scoring scale with 4 response options has been demonstrated in this study. The necessity of potential refinements will or will not emerge from the regular use of the scale for real-life decision making.

Key points

- **For groupwise revisions of reimbursed products within a therapeutic class, a methodology that defines the MCDA model from scratch, i.e. defining the relevant criteria, measuring the relative importance of these criteria (weights), may be worthwhile.**
- **For case-by-case decisions, where indications vary usually from product to product, a more generic approach such as the one presented in this report, may be more feasible.**
- **Whether the simple 4-point ordinal scale is sufficient for scoring diseases on the different criteria will have to be demonstrated through its use in real life.**

5.2 Further steps in the unmet medical needs programme

The current study focused on a methodology for ranking needs. The next step is to judge whether new interventions claiming to address a high need on this ranking is eligible for a financial compensation (early temporary reimbursement) in the context of the unmet medical needs programme. For this, a similar MCDA approach could be used, using the criteria and weights identified and measured in KCE report 234.[1] The criteria to be covered for this decision are the directly related to those covered for therapeutic and societal needs assessment. Indeed, the appraisal of an intervention will depend on its impact on the criteria that determine therapeutic and societal need; i.e. the impact of the intervention of the quality of life of patients, on life expectancy and on the inconvenience of the treatment. The first step would be to apply an MCDA similar to the one presented in this report to the following criteria:

- Impact of the intervention on quality of life
- Impact of the intervention on life expectancy
- Impact of the intervention on the inconvenience of the treatment for the patient
- Impact of the intervention on the public expenditures related to the disease
- Impact of the intervention on the frequency of the disease

The application of a MCDA approach to appraise the added value of an intervention, will lead to a value score (weighted total score). Rather than using it to estimate a "cost-per-MCDA value", we would recommend to keep the information on the total costs and the total MCDA value separate, in order to allow more legitimate decisions from the societal point of view, taking the needs into account. Whereas the incremental cost-effectiveness approach suggests that interventions with an added value are *d'office* allowed to have an added cost, this is not the case in our model. Whether or not society is willing to pay for an intervention and the amount it is willing to pay will depend on the level of need in the indication targeted by the treatment. An added value, on whatever criterion, should not necessarily imply a higher price.

When taking a cohort decision, account should be taken of the criteria that determined the need. If, for example, patients experience a need because their current treatment is highly inconvenient to them and their relatives, a treatment that improves self-care but does not change the inconvenience of the treatment should not get the same priority as a treatment that reduces the inconvenience of treatment and does not change self-care.

Of particular importance are the appraisals of preventive interventions. Prevention is not an explicit criterion in the proposed MCDA approach. However, it was a very important criterion in the citizens' panel of the KBF. Obviously, if the therapeutic need is considered very high, prevention of the disease should be valued high as well. A preventive intervention will get the highest score on all therapeutic need criteria, as well as on the frequency criteria of societal need (if the intervention is effective at least). Hence, although prevention is not explicitly included in the MCDA design as a criterion, it will play an important role in the decision process due to the split design of the MCDA approach, separating the appraisal of need from the appraisal of the added value of an intervention.

5.3 Future research

This study should be considered a first concrete step in a process towards more transparency and consistency in decision making. It proposes the use of an MCDA tool to rank needs and make cohort decisions. However, it should be clear that an MCDA tool is not and should not be static. The challenge is now to start using this tool in real life and learn from experience to improve it, i.e. not to leave it as it is but revise it on regular occasions to ensure it is still serving its purpose. This may require sessions of reflection with the CAIT / CATT and the General Council of the RIZIV / INAMI, where the extent to which the application of the tool leads to decisions that are in line with the objectives of the Belgian healthcare system in general (sustainability, equity, quality) and the unmet medical needs programme in particular (giving priority to those who are worst off in terms of disease severity and for whom no reimbursed treatment options are available) can be discussed. These discussions should lead to suggestions for improvement and concrete steps to implement these improvements.

In the current MCDA, the criteria are more or less fixed –at least in the quantitative part- because they are the ones for which population preference weights were collected in previous research. The application of the tool might learn that other criteria should be included or at least be considered for inclusion. That would require a new population survey, if the involvement

of the general public in decision making is still considered important and valuable. However, this is not problematic, as in any case the survey would need to be re-done after some time, even if the criteria would not change, because the preferences of the public might change.

Further research is needed on the appropriateness, validity and reliability of different scoring scales for the MCDA approach. For example, how to deal with chronic diseases that have several stages with a variable severity (e.g. ALS and Alzheimer)? Or how to deal with acute diseases with severe sequelae for a very small portion of the patients (e.g. meningococcal infection)?

Also more research on the robustness of the approach would be welcome. Ways to combine the total weighted scores for therapeutic need and societal need in a scientifically sound way should also be explored. More empirical data are needed for this.

MCDA is an interesting approach for several kinds of decisions. It would be worthwhile to consider the development and testing of MCDA approaches for other decisions as well, such as for prioritizing health technologies for HTA.

Key points

- **The major benefit of MCDA compared to classical cost-effectiveness analysis is that it makes it more transparent and explicit which and how other criteria that are relevant for the decision are taken into account.**
- **The MCDA framework allows to make a distinction between interventions with an added value that actually meets patient or society needs and interventions that have an added value but do not meet the needs of patients or society. It hence does not support the idea that interventions with an added value on whatever criterion deserve a price premium compared to the current intervention.**

APPENDICES

APPENDIX 1. TOOL FOR ASSESSING AND REPORTING THE QUALITY OF EVIDENCE

Epidemiological data

Epidemiological data could provide evidence on the impact of a disease on **life expectancy** and on the **frequency** of disease.

(Adapted from The EVIDEM Collaboration under a Creative Common License; http://creativecommons.org/licenses/by/2.5/ca/ , June 2015; www.evidem.org)

Evaluation of an individual study – Epidemiological data			
Instructions: - *Perform a critical analysis of study(ies) to comment on all dimensions below, provide a critical overview and a quality grade* - *If multiple studies are available, complete the instrument for each individual study*			
Disease: **Setting:**		**Study (reference):**	
	Questions	**Critical overview**	**Quality grade**
OVERALL ASSESSMENT OF EPIDEMIOLOGICAL DATA	Is the study question relevant (setting, condition of interest)? Is the design appropriate (sample type and size, data sources and measurements, analyses, statistics)? *See dimensions below*		☐ Low ☐ Moderate ☐ High ☐ Excellent
Dimension	**Questions**		**Comment**
1. Study population	Is the study population (sample) representative of the target population (usually the general population) for which prevalence or incidence estimates are sought? Were the sampling method (e.g. random sampling) and the sampling frame (e.g. registry, insurance claims database) to select the study population valid? Were valid methods used to assess and address sampling bias? Are the items above fully reported?		

2.	**Setting and period of data collection**	Are the setting, including locations, clinical setting and period of data collection, reported? Are they relevant and valid?
3.	**Condition of interest**	Is the condition, for which prevalence or incidence estimates are reported, clearly defined, including case definition, diagnostic criteria and codes, disease stage, age of onset, duration frequency of symptoms, comorbidities? Is the condition, for which prevalence or incidence estimates are reported, relevant and valid? Does it correspond to the condition, or aspect of the condition, for which the need has to be assessed?
4.	**Outcome measures**	Are the reported outcome measures (e.g. prevalence at birth, point prevalence, incidence…) clearly defined? Are they relevant with respect to the condition of interest?
5.	**Data sources / measurements**	Are the methods used to establish the presence/absence of the condition fully reported? Are these methods valid (e.g. validity of diagnostic test)? Are the sources used to establish the occurrence of the condition (e.g. outcomes measured, medical records, reimbursement databases, registries, self-report) reported? Are they valid? (All diagnosable cases in study sample identified? Any misclassification prevented?)
6.	**Study sample size**	Was the study sample size sufficient to arrive at reliable epidemiological estimates, including estimates for population subgroups?
7.	**Analyses**	Are the analyses performed clearly described? Are the analyses appropriate and comprehensive with sound and relevant statistics (e.g. missing data, control of confounding, population subgroup analyses)?
8.	**Conclusions**	Are the conclusions supported by the results?

Summary of quality of evidence – Epidemiological data

Evidence on frequency of the disease and impact on life expectancy

Instructions:

- *Provide an overall quality grade of the epidemiological evidence, based on the critical assessment of the individual studies*
- *Copy your text under 'critical overview' and the overall quality grade in the evidence table of the relevant criterion under "comments".*

Disease:

Setting:

	Questions	Critical overview	Quality grade
OVERALL ASSESSMENT OF EPIDEMIOLOGICAL EVIDENCE	How relevant is the existing research with regard to epidemiological data? Are conclusions valid across the range of studies (conclusions across studies consistent or conflicting)? Are individual studies relevant and valid? *See dimensions below*		☐ Low ☐ Moderate ☐ High ☐ Excellent

Patient-reported outcome data

Patient-reported outcome data could provide evidence on the impact of a disease on **quality of life** and on the **inconvenience of current treatment.**

(Adapted from FANLTC (www.facit.org), TSQM-II (Atkinson et al, 2005) and AHRQ National Quality Measures Clearinghouse database)

Evaluation of an individual study - Patient-reported outcomes			
Instructions: - *Perform a critical analysis of studies and reports to comment on all dimensions below, provide a critical overview and a quality grade* - *If multiple studies are available, complete the instrument for each individual study*			
Disease: **Setting:**		**Study:**	
	Questions	**Critical overview**	**Quality grade**
OVERALL ASSESSMENT OF CONDITION-RELATED PATIENT-REPORTED OUTCOMES	Is the study question relevant (perspective, disease management strategy, time horizon and patient population)? Is the design appropriate (qualitative/ quantitative)? *See dimensions below*		☐ Low ☐ Moderate ☐ High ☐ Excellent
Dimension	**Questions**		**Comment**
1. Study population	Is the study population (sample) representative of the target population (usually the general population) for which information is sought? Was the sampling method (e.g. random sampling) to select the study population valid? Were valid methods used to assess and address sampling bias? Are the items above fully reported?		
2. Setting and period of data collection	Are the setting, including locations, clinical setting and period of data collection, reported? Are they relevant and valid?		
3. Condition of interest	Is the condition, for which data are reported, clearly defined, including case definition, diagnostic criteria, disease stage, age of onset, duration and frequency of symptoms, comorbidities? Is the condition, for which data is reported, relevant and valid? Does it correspond to the condition, or aspect of the condition, for which the need has to be assessed?		

4. Outcome measures	Are the reported outcome measures (e.g. physical, emotional and functional well-being, …) clearly defined? Are they relevant with respect to the condition of interest?
5. Data sources / measurements	Are the methods used to establish the presence/absence of the condition fully reported? Are these methods valid (e.g. validity of the instrument used)? Are the sources used to establish the occurrence of the condition (e.g. outcomes measured, medical records, interviews, questionnaires, focus groups, registries, self-report) reported? Are they valid?
6. Study sample size	Was the study sample size sufficient to arrive at reliable results (e.g. data saturation, response rate, …)?
7. Analyses	Are the analyses performed (quantitative or qualitative) clearly described? Are the analyses appropriate and comprehensive with sound and relevant statistics?
8. Conclusions	Are the conclusions supported by the results?

Summary of quality of evidence – Patient-reported outcome data

Evidence on impact of the condition on quality of life and inconvenience of current treatment

Instructions:

- *Provide an overall quality grade of the evidence, based on the critical assessment of the individual studies*
- *Copy your text under 'critical overview' and the overall quality grade in the evidence table of the relevant criterion under "comments"*

Disease:

Setting:

	Questions	**Critical overview**	**Quality grade**
OVERALL ASSESSMENT OF PATIENT-RELEVANT OUTCOME DATA	How relevant is the existing research with regard to patient-reported outcome data? Are conclusions valid across the range of studies (conclusions across studies consistent or conflicting)? Are the individual studies relevant and valid?		☐ Low ☐ Moderate ☐ High ☐ Excellent

Public expenditure data

Data on public expenditures per patient could provide evidence on the criterion of **condition-related public expenditures per patient.**

Evaluation of an individual study - Condition-related public expenditures per patient			
Instructions: - *Perform a critical analysis of study(ies) to comment on all dimensions below, provide a critical overview and a quality grade* - *If multiple studies are available, complete the instrument for each individual study*			
Disease: **Setting:**		**Study:**	
	Questions	**Critical overview**	**Quality grade**
OVERALL ASSESSMENT OF CONDITION-RELATED PUBLIC EXPENDITURES PER PATIENT	Is the study question relevant (perspective, disease management strategy, time horizon and patient population)? Is the design appropriate (how close to real market evolution, costs included, quality of sources)? *See dimensions below*		☐ Low ☐ Moderate ☐ High ☐ Excellent
Dimension	**Questions**		**Comment**
1. Study population	Are population characteristics reported? Is the setting clearly described? Is the study population relevant (age, gender, disease stage, comorbidities, setting,…)?° Does it correspond to the actual population for which the needs are to be assessed?		
2. Current management of the disease	Is the disease management strategy reported in the study? Is the management strategy in the study in agreement with the current management strategy in real life in Belgium?		
3. Perspective and cost included	Is the perspective of the analysis the public payer's perspective? Are all costs relevant to the public payer's perspective included (costs of current disease management strategy, including costs healthcare interventions and costs of treating complications, or sequelae, invalidity benefits…)?		
4. Time horizon	Is the time horizon reported? Is the time horizon long enough to capture all meaningful public expenditures related to the disease?		

5. Clinical event pathways	Is the event pathway fully reported? In case of modelling for cost estimation, does the model reflect a realistic event pathway according to current knowledge?
6. Parameters estimates	Are all parameters used in the study (resource use, unit costs, event probabilities) reported? Are the sources and methods used to estimate the model parameters valid?
7. Input data	Are input data, their sources and methods for estimates valid (claims, patients, epidemiology, unit costs for intervention, clinical data)?
8. Analyses	Are the analyses performed described? Are the analyses appropriate and comprehensive? Has a sensitivity analysis been performed (in case of modelling)?
9. Conclusions	Are the conclusions supported by the results?

Summary of quality of evidence – condition-related public expenditures per patient

Instructions:

- *Provide an overall quality grade of the evidence, based on the critical assessment of the individual studies*
- *Copy your text under 'critical overview' and the overall quality grade in the evidence table of the relevant criterion under "comments".*

Disease:

Setting:

	Questions	**Critical overview**	**Quality grade**
OVERALL ASSESSMENT OF CONDITION-RELATED PUBLIC EXPENDITURES PER PATIENT	How relevant is the existing research with regard to condition-related public expenditures per patient? Are conclusions valid across the range of studies (conclusions across studies consistent or conflicting)? Are the individual studies relevant and valid?		☐ Low ☐ Moderate ☐ High ☐ Excellent

APPENDIX 2. TEMPLATE FOR WRITTEN CONSULTATION OF PATIENT GROUPS (ADAPTED FROM THE HTAI PATIENT GROUP SUBMISSION TEMPLATE)

Available from: http://www.htai.org/interest-groups/patient-and-citizen-involvement/resources/for-patients-and-patient-groups.html (accessed 22/06/2016)

<HTA committee name>

Health Technology Assessment (HTA) on <Name of Medicine in Condition X>

Purpose of this form

We recognise that patients have unique knowledge about what it's like to live with a specific disease or medical condition. They can describe advantages and disadvantages of therapies, which may not be reported in published literature. They can tell us what they would most value from a new treatment.

This submission form has been created to help patient groups provide information for the assessment of a particular medicine. It provides prompts to draw out the unique patient knowledge that has the greatest potential to influence the decisions made by HTA staff and appraisal committees.

Section 2 provides guidance to patient groups on how to complete this submission form.

Section 3 asks you for some background information about your patient group.

Sections 4-8 are the main part of the form for you to complete, describing the views and experiences of patients and their care-givers/carers.

We recognise that completing this form requires substantial resources and so we commit to making all patient submissions available to all involved in the appraisal process, particularly HTA reviewers and committee members. Furthermore, our assessment reports and/or HTA advice/recommendations will document how the information from patients was considered in developing our conclusions/recommendations.

If requested, we will provide you with further feedback about how the submission from your patient group was used and influenced decision-making.

How to complete this form

In the main sections of this form, you are asked to describe the challenges patients face in living with the condition being studied, experiences of using current therapies, expectations from the new medicine and, if you are aware, the potential benefit or drawbacks from the new medicine being assessed.

Each question has a series of prompts in a box that are intended to assist you in providing the information that will be helpful to HTA reviewers and committees in understanding the impact of the condition and its treatment. Please address any of the prompts that your group feels is important and describe any other relevant issues that are not captured in the list of prompts.

In all parts of this form the term "patient" refers to anyone living with, or who has lived with, the condition for which the new medicine is indicated.

Please provide clear facts, information and summaries of experiences to give a concise, accurate and balanced overview of a range of patients' and care-givers' (carers') perspectives/views. State the source of your information (e.g. web survey, helpline analysis, social networking, focus group, patients' records, one-to-one conversations with those in clinical trials, patient stories) and provide clear references where they are available.

There is no need to send us published scientific papers, as we already have access to those. However, if you have views about the interpretation of a paper about a particular clinical trial, we would be happy to hear them.

For any of the sections in the form, if there are groups that should have special consideration, please indicate the specific needs/issues of that group (e.g. children, women/men, ethnic groups, those living in a particular location, those with other disabilities, disease sub-types).

If you require help in understanding HTA related terms, please refer to the HTAi glossary for patients or visit the training resources on the HTAi website.

If you have any questions when completing this form, please contact

<NAME, PHONE, EMAIL – Contact person from HTA organisation>

Information about your group

Name of group:

Key contact name:

Role:

Email:

Phone:

Postal address:

Type of group (tick all that apply):

- Registered charity ☐
- Fellowship ☐
- Informal self-help group ☐
- Unincorporated organisation ☐
- Other
  Please state_____

Purposes of group (tick all that apply):

- Advocacy ☐
- Education ☐
- Campaigning ☐
- Service ☐
- Research ☐
- Other
  Please specify_____

Describe your membership (number and type of members, region that your group represents, demographics, etc)?

In line with how we treat other stakeholders, we ask you to complete our declaration of interests.

Did anyone help you prepare this submission? YES / NO

If yes – who helped you and in what way?

Are you willing for this submission to be shared on our website after removal of financial information and personal information that could identify patients? YES / NO

We may invite you to meetings where this HTA is to be discussed. Would a member of your group be willing to attend such meetings? YES / NO

Impact of condition

How does the condition or disease for which the medicine is being assessed, affect patients' quality of life?

Issues to consider in your response

- Aspects of the condition that are most challenging (e.g. symptoms, loss of ability to work, loss of confidence to go out, inability to drive, social exclusion).
- Activities that patients find difficult or are unable to do.
- Aspects of the condition that are the most important to control (e.g. symptoms that limit social interaction or ability to work such as difficulty breathing, pain, fatigue, incontinence, anxiety).
- Support required for daily living (physical or emotional).
- Types of patients that are most affected by the condition (e.g. men/women, children, ethnic groups).
- Challenges in managing this condition when patients also have other medical conditions.
- What patients would most like to see from a new treatment (e.g. halting of disease progression, improvement in a particular symptom).

How does the condition or disease affect carers/unpaid care-givers?

Issues to consider in your response

- Challenges faced by family and friends who support a patient to manage the condition.
- Pressures on carers/care-givers daily life (e.g. emotional/psychological effects, fatigue, stress, depression, physical challenges)

Experience with current therapies

How well are patients managing their condition with currently available therapies?

(Currently available therapies may include any form of medical intervention such as medicines, rehabilitation, counselling, hospital interventions etc. If no specific therapy is available that should be stated.)

Issues to consider in your response

- Main therapies currently used by patients for this condition and how they are given (tablet, injection, physiotherapy, hospital check-ups, etc, at home, in hospital; dose and frequency, ease of access)
- Extent to which current therapies control or reduce the most challenging aspects of the condition.
- The most important benefits of current therapies.
- The burden of therapy on daily life (e.g. impact at different stages of disease, interruption to work, stigma, clinic visits to receive infused medicines, need for weekly blood tests or describe a typical episode of therapy over a week or period of treatment).
- Side effects from the therapies that are difficult to tolerate.
- Concerns about long-term use of current therapy.
- If the current therapy is a medicine:
  - Challenges in taking it as prescribed (e.g. swallowing the pill, self-injecting, use of a device to deliver the medicine, taking after food, not being able to lie down for 30 minutes after taking medicine).
  - Ways in which the dosing is modified compared to what is prescribed (e.g. dividing the dose to avoid unwanted side effects, missing doses to fit into daily life).

Additional information

Please include any additional information you believe would be helpful to the HTA reviewers and committee (e.g. ethical or social issues).

Key messages

Section moved to top when presented to HTA Committee

In no more than five statements, please list the most important points in of your submission.

For example:

- The biggest challenges of living with this condition are…
- Current therapies are inadequate because…
- This new medicine will be important for patients because…

APPENDIX 3. EXTRACT FROM THE PRE-TEST STUDY PACKAGE

Instructions

In your package you find descriptions of 8 diseases, organised in five summary-of-evidence tables per disease. Each table describes the available evidence with respect to one criterion.

- The first table describes the impact of the disease on **health-related quality of life**, given the treatment currently already available to patients. E.g. it describes whether patients still have problems with usual activities, despite the treatment that is currently available and provided to them. Also impact on mobility, self-care, pain/discomfort and anxiety/depression is described if evidence is available.
- The second table describes the **inconvenience or discomfort of current treatment**. Inconvenience relates to the route of administration (e.g. injection or oral), place of administration (e.g. at the hospital or at home), frequency, or duration of treatment.
- The third table describes the impact of the disease on **life expectancy**, given current treatment.
- The fourth table describes the **frequency** of the disease in the population.
- The fifth table describes the **public expenditures** associated with the disease. Public expenditures encompass direct health care expenditures, sickness and invalidity insurance benefits, publicly financed home care services, etc. The score should be based on the public expenditure per patient.

The first column of each table mentions the dimensions of the criteria affected by the disease, e.g. the impact of the disease on self-care as a dimension of quality of life; the second column provides the evidence with respect to these dimensions, the third column specifies the sources of the information.

Please score each criterion on the scale presented below each table. Sumbit your responses via LimeSurvey (you have received a link by e-mail).

You should give **your own appreciation of the impact of the disease** on the selected criteria **for the entire** patient **population**, even if you would like to differentiate between sub-groups.

You should consider the evidence provided as being up-to-date and validated. We do not ask you to assess the correctness or validity of the information provided.

In the comments box below the scoring table you can add whatever you like (e.g. your overall confidence in the evidence provided) that could be useful for the discussions during the face-to-face meetings.

Invasive meningococcal disease

Invasive Meningococcal Disease (IMD) is a severe disease affecting mainly young children and adolescents. IMD clinically presents as meningitis, septicimia or both, sometimes leading to septic shock and in a minority of cases to arthritis or pericarditis.

Criterion: Impact of the disease on quality of life (QoL)		
Dimensions	Magnitude of the impact of the disease	Sources of evidence
Patients with sequelae might have problems with: • Mobility • Self-care • Usual activities • Pain/discomfort • Anxiety/depression	Invasive Meningococcal Disease (IMD) may lead to sequelae in 3-19% of surviving cases: • hearing loss: 1.9-7.2% • scar and/or necrosis: 1.2-6.4% • amputation: 0.3-8% • epilepsy or seizures: 1.6-3.3% • renal failure: 2% About 16% of all patients with IMD shows psychological disorders. For survivors with any sequelae, utility losses range from 0.07 to 0.4 on a 0-1 scale, as measured with the EQ-5D. Values reported for specific sequelae vary considerably. • Utilities for the health state 'mild hearing loss' ranged from 0.49 with VAS to 0.91 with EQ-5D, on a 0 (death) to 1 (perfect health) scale; while those for 'severe hearing loss' (i.e. bilateral hearing loss or deafness) ranged 0.28 with HUI-3 to 0.86 with SG. • Utilities for the health state 'mild cognitive impairment' ranged 0.24 with HUI-3 to 0.74 with SG while those for 'severe cognitive impairment ranged 0.22 with VAS to 0.39 with SG. Despite using different methodologies, two papers report similar values for bacterial meningitis with recovery (without any severe sequelae): using the HUI-2 Koomen reports a utility of 0.93, while Bennet reports a utility of 0.98 (compared to 1).	KCE-study 231 Sequelae: literature review; 7 studies published after 1995 and performed in similar settings (Europe and North America) QoL: systematic literature review, 17 studies
Comments	Robust proxies for the frequency of sequelae in the Belgian context are difficult to establish due to different settings, case recruitment, follow-up period, sequelae definitions and measures in published studies. Furthermore, phenotypes and clones widely differ across geographical areas, and they may be associated with different risk of death and sequelae. In addition, the frequencies of specific sequelae do not take into account the risk of sequelae that were caused by other diseases. A case-control study in the UK compared these frequencies to those in controls; for instance, hearing loss was found in 1% of the controls due other causes such as e.g. otitis media.	

Scoring "impact of disease on quality of life given current treatment" (please check one box)

The disease has

☐ 0: **no** negative impact on the quality of life of patients, compared to people without the disease

☐ 1: **some** negative impact on the quality of life of patients, compared to people without the disease

☐ 2: a **high** negative impact on the quality of life of patients, compared to people without the disease

☐ 3: a **very high** negative impact on the quality of life of patients, compared to people without the disease

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Inconvenience of current treatment		
Aspects of current treatment causing inconvenience to patients	Magnitude of the inconvenience induced by each of these aspects	Sources of evidence
Depends on the sequelae	Meningococcal disease requires hospitalisation. Following hospital discharge, survivors of meningococcal disease that do not develop sequelae have to have 2 to 3 control visits with a specialist internal medicine. For those with sequelae, the inconvenience of treatment depends on the type of sequelae: • hearing loss requires hearing aid • severe neurological sequelae might require institutional care • severe skin scars or necrosis may require a skin graft • treatment of epilepsy requires lifelong medication intake • renal failure may require dialysis • minor amputations (finger, toe) may require the use of a prosthesis • major amputation (limb) require a prosthesis	KCE report 231
Comments	The description of the inconvenience caused by the treatment of the sequelae of the IMD is based on a description of the costs associated with the sequelae, rather than the real inconvenience reported by patients.	

Scoring "inconvenience of current treatment" (please check one box)

☐ 0: Current treatment is **not or only slightly** inconvenient to patients

☐ 1: Current treatment is **somewhat** inconvenient to patients

☐ 2: Current treatment is **highly** inconvenient to patients

☐ 3: Current treatment is **very highly** inconvenient to patients

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Impact of the disease on life expectancy		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Case fatality ratio	Case fatality ratio: 5.4% (95% CI 4.1-7.3). Case fatality ratio is higher among infants (5-6%), lower among older children (4%) and very high in the elderly (14-16%)	NRC data (National Reference Centre) matched with MZG-RHM data (Minimale Ziekenhuisgegevens – Résumé Hospitalier Minimum)

Scoring "impact of disease on life expectancy given current treatment" (please check one box)

The disease has

☐ 0: **no** impact on life expectancy

☐ 1: **some** impact on life expectancy (patient loses a small proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 2: a **high** impact on life expectancy (patient loses a large proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 3: a **very high** impact on life expectancy (patient dies almost immediately, despite the best available current treatment or care)

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Frequency of disease		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Incidence	The incidence of Invasive Meningococcal Disease (IMD) is 1.3/100 000 persons per year (139 cases).	Annual numbers of IMD per year and age group: Hospital clinical records (MZG-RHM: Minimale Ziekenhuisgegevens – Résumé Hospitalier Minimum): ICD-9 code 036.

Scoring "prevalence of the disease" (please check one box)

The prevalence of the disease is

☐ 0: less than 2000 people in Belgium have the condition (less than 1 per 5500)

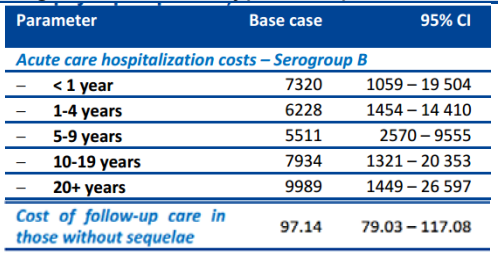
☐ 1: between 2000 and 10 000 people in Belgium have the condition (between 1 per 5500 and 1 per 1100)

☐ 2: between 10 000 and 100 000 people in Belgium have the condition (between 1 per 1100 and 1 per 110)

☐ 3: more than 100 000 people in Belgium have the condition (more than 1 per 110)

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Disease-related public expenditures		
Dimensions	Magnitude of each type of expenditures	Sources of evidence
Direct health care expenditures		Estimates originate from a combination of sources: - Coupled MZG-RHM/AZV-SHA data and NRC data, selection based on ICD-9-CM diagnostic code “36” as principal or any diagnosis indicating IMD - Incidence of sequelae - Nomenclature Calculations in KCE reports

Parameter	Base case	95% CI
Acute care hospitalization costs – Serogroup B		
– **< 1 year**	7320	1059 – 19 504
– **1-4 years**	6228	1454 – 14 410
– **5-9 years**	5511	2570 – 9555
– **10-19 years**	7934	1321 – 20 353
– **20+ years**	9989	1449 – 26 597
Cost of follow-up care in those without sequelae	97.14	79.03 – 117.08

Cost of follow-up care in those with sequelae varies by age:
Acute one-off costs:
- 0-4 year: 1434 [1044-1885]
- 5-17 year: 91 [66-119]
- 18-19 year: 45 [33-60]
- 20-65 year: 3445 [2508-4529]

Lifelong annual cost:
Between 3215 [2340-4226] and 3355 [2442-4411], depending of the age of the patient when getting the disease.

Scoring “current disease-related public expenditures per patient” (please check one box) [f]

The disease currently has

☐ 0: a **very small** impact on public expenditures per patient

☐ 1: **some** impact on public expenditures per patient

☐ 2: a **high** impact on public expenditures per patient

☐ 3: a **very high** impact on public expenditures per patient

Comments (here you can write down concerns regarding the evidence level or other)

[f] For your information: total direct public health care expenditures in Belgium are about € 31 billion a year; mean health care expenditure per citizen is about € 2,300/year.

Alzheimer

Clinical Alzheimer’s disease (AD) is preceded by a slowly progressing accumulation in the brain of amyloid plaques and neurofibrillary tangles with hyperphosphorylated tau protein. In many cases AD is present in combination with some degree of cerebrovascular damage (mixed dementia). With the exception of some genetically well-defined forms of AD, at present a definitive diagnosis of AD still requires histopathological confirmation of a probabilistic clinical diagnosis.

Alzheimer's disease is an incurable brain disease, wherein cells in some parts of the brain cease to function and in some parts of the brains to function and die. Alzheimer 's disease is irreversible and despite extensive research, the causes of the disease and treatment methods are still unknown. The symptoms of the disease are forgetfulness, personality changes, disorientation, and loss of speech. Alzheimer 's disease is the most common form of dementia. The risk of dementia increases with age. A high age is the most important known risk factor for Alzheimer's dementia and for the most other forms of dementia.

Memory impairment is usually one of the first symptoms of AD. As the disease progresses cognitive deficits start to interfere with activities of daily living (ADL) and behavioural problems may appear. These behavioural and psychological signs and symptoms in dementia (BPSD) commonly include depression, apathy, agitation, disinhibition, psychosis, wandering, aggression, incontinence and altered eating habits. They contribute significantly to caregiver burden, institutionalization (placement in an elderly home), and decreased quality of life for patients with dementia.

Criterion: Impact of the disease on quality of life		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Self-care Usual activities Anxiety/depression	Memory impairment is usually one of the first symptoms of AD. As the disease progresses cognitive deficits start to interfere with activities of daily living (ADL) and behavioural problems may appear. These behavioural and psychological signs and symptoms in dementia (BPSD) commonly include depression, apathy, agitation, disinhibition, psychosis, wandering, aggression, incontinence and altered eating habits. They contribute significantly to caregiver burden, institutionalization (placement in an elderly home), and decreased quality of life for patients with dementia. Most people with dementia, however, keep a sufficient quality of life even in the most advanced stages of their disease. This can be explained by –amongst others- coping strategies.	KCE report 111

Scoring "impact of disease on quality of life given current treatment" (please check one box)

The disease has

☐ 0: **no** negative impact on the quality of life of patients, compared to people without the disease

☐ 1: **some** negative impact on the quality of life of patients, compared to people without the disease

☐ 2: a **high** negative impact on the quality of life of patients, compared to people without the disease

☐ 3: a **very high** negative impact on the quality of life of patients, compared to people without the disease

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Inconvenience of current treatment		
Aspects of the current treatment causing inconvenience to patients	Magnitude of the inconvenience induced by each of these aspects	Sources of evidence
	Non-pharmacological interventions There is lack of standardization of non-pharmacological interventions in Alzheimer's disease. Many variations and combinations of interventions are evaluated in studies. The fact that few studies address the same issues in comparable ways is a major methodological problem, limiting the level of evidence which can be associated with such unique interventions. Non-pharmacological interventions include, for instance, cognitive stimulation or training, reminiscence therapy, "snoezelen" multisensory stimulation, music therapy, massage and touch, exercise or activity therapy, ergotherapy, etc. There is no evidence on the inconvenience of these interventions. Among these interventions, promising are cognitive stimulation (alone or add-on to therapy with inhibitors of acetylcholinesterase), ADL-rehabilitative care, music therapy, massage/touch, physical activity, and maybe interventions aiming at improving communication/interaction. For all the other interventions, insufficient high quality evidence is available to support or reject the therapy; including environmental changes like group living, intensified care programs or special dementia care units. Many patients with Alzheimer's disease are taken care off at home. Care at home for as long as possible is often preferred over institutionalization both by the patient and the family-caregiver. However, caring for a person with dementia at home is intensive and burdensome. Caregivers are at high risk of psychosocial morbidity and associated breakdown in care. Ample evidence is available on this impact. For interventions involving formal or informal caregiver(s), some moderate level of evidence was found for a positive effect of several forms of psychosocial interventions and psychoeducation on informal caregiver depression and stress. Education and training of staff were found to be promising interventions. Support measures preventing caregivers from becoming overburdened and depressed result in a delay of institutionalisation, as shown in a meta-analysis of 13 support programs. One example from this meta-analysis is the study by Mittelman et al., for which the longest follow-up period is available. This large randomized controlled trial (RCT) studied over a 9.5-year period 406 spouse caregivers of community dwelling AD patients in New York City. Enhanced counseling and support consisted of six sessions of individual and family counseling, support group participation, and continuous availability of ad-hoc telephone counseling. This intervention was associated with a delay in median time to placement of 557 days. In addition, self-rated health in intervention group caregivers was significantly better than control group caregivers.	KCE report

	Pharmacological treatment Drugs are not curative, but claim to slow down the deterioration of patients with dementia. Donepezil, Galantamine, Rivastigmine and Memantine are all drugs approved in Belgium for the treatment of dementia (apart from Memantine which is approved for moderate to severe dementia, all drugs are approved for mild to moderately severe dementia). Inhibitors of acetylcholinesterase (ChEIs) (donepezil, galantamine and rivastigmine) have dose-dependent adverse effects, including vomiting, nausea, diarrhoea and anorexia. These gastrointestinal adverse event rates limit dose increases of ChEIs. They can be partly avoided by means of a slower dose titration rate. Also cardiac side-effects (bradycardia, AV-block) have been reported but these are less frequent. It is well established that the use of both typical and atypical antipsychotics in patients with dementia is associated with an increased mortality rate. Their use should be restricted, e.g. to hostile, aggressive patients.

Scoring "inconvenience of current treatment" (please check one box)

☐ 0: Current treatment is **not or only slightly** inconvenient to patients

☐ 1: Current treatment is **somewhat** inconvenient to patients

☐ 2: Current treatment is **highly** inconvenient to patients

☐ 3: Current treatment is **very highly** inconvenient to patients

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Impact of the disease on life expectancy		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
	The average life expectancy after diagnosis of AD is 5 to 6 years, or 8 years after the first symptoms. Life expectancy depends on the age at the moment of the diagnosis: as high as 10.7 years for the youngest patients (65-69 years) to a low of 3.8 years for the oldest (90 or older at diagnosis). Roughly, this is about half of the life expectancy of a person with the same age but without AD.	KCE report

Scoring "impact of disease on life expectancy given current treatment" (please check one box)

The disease has

☐ 0: **no** impact on life expectancy

☐ 1: **some** impact on life expectancy (patient loses a small proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 2: a **high** impact on life expectancy (patient loses a large proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 3: a **very high** impact on life expectancy (patient dies almost immediately, despite the best available current treatment or care)

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Frequency of the disease		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Prevalence	In 2005, the proportion of people with dementia in the European Union was estimated at 1.14% to 1.27% of the population. For Belgium the estimate was 1.22% to 1.35%, corresponding to 127 174 and 140 639 subjects.(www.dementia-in-europe.eu, consulted May 14, 2008). The number of Alzheimer's Disease patients in Belgium is not well documented but can be estimated at about 75 000 patients in 2008 (presentation Prof Patrick Santens, Ghent, 2008). This corresponds to about 55% of all dementia patients. The published prevalence of dementia ranges from 6.3% to 9.3% for subjects 65 years of age and older. About one in three persons aged 90 years and older has dementia. A Belgian study based on GP consultations reported a prevalence of 11% in subjects 65 and older, living at home. The prevalence varies strongly with the study context: from a prevalence of dementia diagnosed by a GP of 2% in subjects 60 years and older, to 44.1% to 47% for institutionalized subjects over 60 years of age. About two thirds of the dementia patients in Belgium are women, mainly because the life expectancy is higher for women than men, but also because of a somewhat higher incidence in elderly women.	KCE report

Scoring "prevalence of the disease" (please check one box)

The prevalence of the disease is

☐ 0: less than 2000 people in Belgium have the condition (less than 1 per 5500)

☐ 1: between 2000 and 10 000 people in Belgium have the condition (between 1 per 5500 and 1 per 1100)

☐ 2: between 10 000 and 100 000 people in Belgium have the condition (between 1 per 1100 and 1 per 110)

☐ 3: more than 100 000 people in Belgium have the condition (more than 1 per 110)

Comments (here you can write down concerns regarding the evidence level or other)

APPENDIX 4. SUBMISSION TEMPLATE FOR THE UNMET MEDICAL NEEDS LIST

NAME OF THE DISEASE:

BRIEF DESCRIPTION (e.g. affected organ(s)/body part, cause, age at onset):

Please complete the tables below with the data and/or evidence currently available. If no data or evidence is available for specific items, you can leave those cells blank. Tables should preferably contain numbers. Textual remarks should be provided in the comment box below each table.

Impact of the disease on quality of life

Please describe here the impact of the disease on various aspects of quality of life, and in particular the impact on mobility (ability to walk around), self-care (ability to wash and dress self), usual activities (work, study, housework, family or leisure activities), pain/discomfort (related to the disease but not due to the treatment) and anxiety/depression, **given current treatment or care**. Provide EQ-5D evidence if available for patients under current treatment. Include potentially differentiating characteristics of sub-groups (e.g. special risks). Reference all statements.

Describe in the **first column** the dimensions of quality of life (QoL) still affected by the disease, despite current treatment. For the dimensions mentioned in the table below, a Y/N suffices.

Describe in the **second column** for each dimension mentioned in the first column the estimated magnitude of the impact of the disease on that dimension, with its range of uncertainty (e.g. 95% confidence interval). Multiple estimates can be presented if several studies are available. If there is a large variation between subgroups of patients, this should be highlighted here. Please describe in that case the characteristics that differentiate the subgroups.

Describe in the **third column** the sources used to estimate the magnitude; more specifically mention the study type(s) (e.g. register of patient data, RCT, observational study, interviews), the number of patients included and the full reference in case of published evidence.

Impact of the disease on quality of life		
Description of the dimensions of QoL affected by the disease	Magnitude of the impact of the disease on each dimension	Sources used to estimate the magnitude: number of observations and/or studies
Mobility		
Self-care		
Usual activities		
Pain/discomfort		
Anxiety/depression		
…		

Comments:

Inconvenience of the current treatment or care

Please describe in detail what the current treatment entails for patients.

Describe evidence to understand to what extent the treatment of the disease/health problem causes hindrance, trouble, or difficulty for the patient, related to:

- How easy or difficult is it to use the medication/undergo the therapy in its current form?
- How easy or difficult is it to plan when to use the medication/undergo the therapy each time?
- How convenient or inconvenient is it to take the medication/undergo the therapy as instructed?
- What are the adverse effects of the current treatment?

Inconvenience of treatment can be related to the frequency of use, the administration route, place of treatment, support required for the treatment and morbidity caused by the treatment (e.g. through side-effects).

It is acknowledged that inconvenience of a treatment can have an impact on the criterion "impact of the disease on quality of life given current treatment". However, this is definitely not always the case. Morbidity related to the condition or co-morbidities, for instance, resorts under the criterion 'impact of the condition on quality of life', whereas side-effects of current treatment can have an unidentifiable effect on patients' quality of life (especially if measured by a generic measure such as the EQ-5D) and still be relevant for the patient. The latter resort under the criterion 'inconvenience of current treatment'. Therefore, it remains very important to describe inconvenience of current treatment separately.

Include potentially differentiating aspects of the current treatment (e.g. sub-population). Reference all statements.

Describe in the **first column** the aspects of the current treatment causing inconvenience and the type of inconvenience created.

Describe in the **second column** for each type of inconvenience mentioned in the first column the estimated magnitude of the inconvenience, with its range of uncertainty (e.g. 95% confidence interval). Multiple estimates can be presented if several studies are available.

Describe in the **third column** the sources used to estimate the magnitude; more specifically mention the study type(s) and the number of patients included.

Inconvenience of current treatment		
Description of aspects of the current treatment causing inconvenience to patients	Magnitude of the inconvenience induced by each of these aspects	Sources used to estimate the magnitude: number of observations and/or studies
For example: Administration route Frequency of use Duration of treatment (per unit of use) Duration of treatment effect Logistics Adverse effects		

Comments:

Impact of the disease on life expectancy

Please describe here the impact of the disease on life expectancy**, given current treatment or care**. Estimates of impact on overall life expectancy is preferred. Do not include information on incidence and prevalence of the disease here. Include differentiating characteristics of sub-groups. Reference all statements.

Describe in the **first column** the dimensions of life expectancy still affected by the disease, despite current treatment. Besides overall survival, which is considered to be the default, disease-free survival and progression-free survival could be mentioned, especially if no evidence on overall survival is available.

Describe in the **second column** for each dimension mentioned in the first column the estimated magnitude of the impact of the disease on that dimension, with its range of uncertainty (e.g. 95% confidence interval). Multiple estimates can be presented if several studies are available.

Describe in the **third column** the sources used to estimate the magnitude; more specifically mention the study type(s) and the number of patients included.

Impact of the disease on life expectancy		
Description of the dimensions of life expectancy affected by the disease	Magnitude of the impact of the disease on each dimension	Sources used to estimate the magnitude: number of observations and/or studies
For example: Overall survival Disease-free survival Progression-free survival		

Comments:

Frequency of the disease

Please describe here the frequency of the disease. Identify, analyse and report the size (epidemiological data) of population affected by the disease/health condition among a specified population at a specific time and its demographic characteristics. Reference all statements.

Describe in the **first column** the prevalence, as a default, and –if available– the incidence of the disease.

Describe in the **second column** the estimated frequency, with its range of uncertainty (e.g. 95% confidence interval).

Describe in the **third column** the data sources used to estimate the frequencies.

Frequency		
Description of the dimensions related to the frequency of disease	Magnitude of the impact of the disease on each dimension	Sources used to estimate the frequencies: number of observations and/or studies
Prevalence		
Incidence		

Comments:

Disease-related public expenditures

Please describe here the impact of the disease on public expenditures. Public expenditures can relate to direct health care expenditures (in terms of drugs, physician visits, hospitalization, surgery), indirect health care expenditures (e.g. reimbursed transportation costs) or disease-related non-healthcare expenditures (e.g. social security benefits (disability, invalidity, as available)). Present evidence on different types of expenditures and their sub-items separately. If no breakdown is available, indicate this as such.

Include potentially differentiating characteristics of sub-groups. Reference all statements.

Describe in the **first column** the types of public expenditures affected by the disease, given current treatment.

Describe in the **second column** for each type of expenditure mentioned in the first column the estimated magnitude of the impact of the disease on that type of expenditure, with its range of uncertainty (e.g. 95% confidence interval). Multiple estimates can be presented if several studies are available.

Describe in the **third column** the data sources used to estimate the magnitude of the expenditures.

Disease-related public expenditures		
Description of the dimensions related to the disease-related public expenditures	Magnitude of each type of expenditures	Sources used to estimate the magnitude of the expenditures: number of observations and/or studies
For example: Healthcare expenditures directly related to the current treatment Healthcare expenditures related to the treatment of side-effects Non-healthcare public expenditures		

Other criteria

If considered relevant and important, information may be provided about other criteria, such as

- **Social/family well-being**
  - Acceptance of disease by society
  - Impact of the disease on the patients' environment (family, informal caregivers)
  - Impact on relationships
  - Development of taboo or stigma
  - Impact on sex life
- **Emotional well-being**
  - Losing hope
  - Ability to cope
- **Functional well-being**
  - Fatigue
  - To sleep well
  - To accept the illness

Reference all statements and include potentially differentiating characteristics of sub-groups.

APPENDIX 5. EXTRACT FROM THE PILOT STUDY PACKAGE

Appraisal of the level of need in several diseases
Pilot test of a multi-criteria approach

Instructions – please read these carefully

YOUR PACKAGE

Your package includes descriptions of 8 diseases, organised as five summary-of-evidence tables per disease. Each table describes the evidence with respect to one criterion:

- The first table describes the impact of the disease on **health-related quality of life**, given the treatment currently already available to patients.
  E.g. to what extent do patients still have problems with usual activities, despite the treatment that is currently provided to them. Also impact on mobility, self-care, pain and anxiety/depression is described if evidence is available.
  For chronic diseases or diseases with long-term consequences, the impact over a patient's lifetime should be considered.
- The second table describes the **inconvenience or discomfort of the current disease management strategy for patients**.
  A management strategy can be a curative treatment, an intervention to relief symptoms or best supportive care. Inconvenience relates to the route of administration (e.g. injection or oral), place of administration (e.g. at the hospital or at home), frequency, or duration of treatment.
  Inconvenience to informal caregivers or others should **not** be taken into account for your score. It can be mentioned in the comment box and discussed afterwards, when the ranking of needs is created.
- The third table describes the impact of the disease on patients' **life expectancy**, given current treatment.
- The fourth table describes the **frequency** of the disease in the population.
- The fifth table describes the **public expenditures** associated with the disease **per patient**.
  Public expenditures encompass direct health care expenditures, but also sickness and invalidity insurance benefits, publicly financed home care services, etc. Your score should be based on the public expenditures per individual patient, and not on what this implies for the entire Belgian population in terms of taxes or contributions.

The first column of each table mentions the dimensions of the criteria affected by the disease, e.g. the impact of the disease on self-care as a dimension of quality of life; the second column provides the evidence with respect to these dimensions and the third column specifies the sources of the information.

INSTRUCTIONS

Please **score** each criterion on the scale presented below each table.

You should give **your own appreciation** of the impact of the disease on the selected criteria **for the entire patient population**. If the impact of a disease differs among sub-groups, the impact should be 'spread' over the entire patient population, and your appraisal should be based on **the average impact**.

You should consider the evidence provided as being up-to-date and valid. We do not ask you to assess the accuracy or validity of the information provided.

In the **comments** box below the scoring table you can add any concerns you might have with respect to the quality of the evidence or with respect to the criteria (e.g. concerns related to the impact on informal caregivers, which is not included as a separate criterion but nevertheless important). These comments should be raised during the discussions regarding the ranking of medical needs and might give rise to changes in the ranking.

Please submit your responses via LimeSurvey before 13 March 2016 (you have received a link by e-mail).

Contact person in case questions: Irina Cleemput, Irina.Cleemput@kce.fgov.be

Invasive meningococcal disease

Invasive Meningococcal Disease (IMD) is a severe disease affecting mainly young children and adolescents. IMD clinically presents as meningitis, sepsis or both, sometimes leading to septic shock and in a minority of cases to arthritis or pericarditis. ***When assessing the impact of the disease, you should consider the entire population affected by invasive meningococcal disease, including those patients not affected by any sequelae. Hence, the impact of the sequelae should be distributed over the entire patient population.***

Criterion: Impact of the disease on quality of life (QoL)		
Dimensions	Magnitude of the impact of the disease	Sources of evidence
Patients with sequelae might have problems with: • Mobility • Self-care • Usual activities • Pain/discomfort • Anxiety/depression	Invasive Meningococcal Disease (IMD) may lead to sequelae in 3-19% of surviving cases: • hearing loss: 1.9-7.2% • scar and/or necrosis: 1.2-6.4% • amputation: 0.3-8% • epilepsy or seizures: 1.6-3.3% • renal failure: 2% About 16% of all patients with IMD shows psychological disorders. For survivors with any sequelae, utility losses range from 0.07 to 0.4 on a scale from 0 (=death) to 1 (=perfect health), as measured with the EQ-5D. Values reported for specific sequelae vary considerably: • Utilities for the health state 'mild hearing loss' ranged from 0.49 with a visual analogue scale (VAS*) to 0.91 with EQ-5D*. • Utilities for 'severe hearing loss' (i.e. bilateral hearing loss or deafness) ranged from 0.28 with Health Utilities Index-3 (HUI-3*) to 0.86 with the Standard Gamble (SG*). • Utilities for bacterial meningitis with recovery without severe sequelae range from 0.93 to 0.98 (1=utility of perfect health). *The VAS, EQ-5D, HUI-3 and the SG are instruments used for measuring health state utilities on a 0(=death) to 1(=utility of perfect health) scale.	KCE-study 231 Sequelae: literature review; 7 studies published after 1995 and performed in similar settings (Europe and North America) Quality of life: systematic literature review, 17 studies
Comments		

Scoring "impact of disease on quality of life given current treatment" (please check one box)

The disease has

☐ 0: **no** negative impact on the quality of life of patients, compared to people without the disease

☐ 1: **some** negative impact on the quality of life of patients, compared to people without the disease

☐ 2: a **moderate** negative impact on the quality of life of patients, compared to people without the disease

☐ 3: a **high** negative impact on the quality of life of patients, compared to people without the disease

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Inconvenience of current treatment		
Aspects of current treatment causing inconvenience to patients	Magnitude of the inconvenience induced by each of these aspects	Sources of evidence
	Meningococcal disease requires hospitalization. Following hospital discharge, survivors of meningococcal disease that do not develop sequelae have to have 2 to 3 control visits with a specialist internal medicine. For those with sequelae, the inconvenience of treatment depends on the type of sequelae: • hearing loss requires hearing aid • severe neurological sequelae might require institutional care • severe skin scars or necrosis may require a skin graft • treatment of epilepsy requires lifelong medication intake • renal failure may require dialysis • minor amputations (finger, toe) may require the use of a prosthesis • major amputation (limb) require a prosthesis	KCE report 231
Comments	The description of the inconvenience caused by the treatment of the sequelae in this table is not based on patient reported outcomes, but derived from a description of treatments for potential sequelae.	

Scoring "inconvenience of current treatment" (please check one box)

☐ 0: Current treatment is **not or only slightly** inconvenient to patients

☐ 1: Current treatment is **somewhat** inconvenient to patients

☐ 2: Current treatment is **highly** inconvenient to patients

☐ 3: Current treatment is **extremely** inconvenient to patients

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Impact of the disease on life expectancy		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Case fatality ratio	Case fatality ratio: 5.4% (95% CI 4.1-7.3). Case fatality ratio is higher among infants (5-6%), lower among older children (4%) and very high in the elderly (14-16%)	NRC data (National Reference Centre) matched with MZG-RHM data (Minimale Ziekenhuisgegevens – Résumé Hospitalier Minimum)

Scoring "impact of disease on life expectancy given current treatment" (please check one box)

The disease has

☐ 0: **no** impact on life expectancy

☐ 1: **some** impact on life expectancy (patient loses a small proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 2: a **high** impact on life expectancy (patient loses a large proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 3: a **very high** impact on life expectancy (patient dies almost immediately or 50% of the patients is dead within one week, despite the best available current treatment or care)

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Frequency of disease		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Incidence	The incidence of Invasive Meningococcal Disease is 1.3/100 000 persons per year (139 cases/year). (=number of acute cases admitted to hospital per year).	Annual numbers of IMD per year and age group: Hospital clinical records (MZG-RHM: Minimale Ziekenhuisgegevens – Résumé Hospitalier Minimum): ICD-9 code 036.

Scoring "frequency of the disease" (please check one box)

The prevalence of the disease is

☐ 0: less than 2000 people in Belgium have the condition (less than 1 per 5500)

☐ 1: between 2000 and 10 000 people in Belgium have the condition (between 1 per 5500 and 1 per 1100)

☐ 2: between 10 000 and 100 000 people in Belgium have the condition (between 1 per 1100 and 1 per 110)

☐ 3: more than 100 000 people in Belgium have the condition (more than 1 per 110)

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Disease-related public expenditures per patient		
Dimensions	Magnitude of each type of expenditures	Sources of evidence
Direct health care expenditures	Costs per patient with IMD in the acute phase (hospitalisation + follow-up care), in €:	Estimates originate from a combination of sources: • Coupled MZG-RHM/AZV-SHA data and NRC data, selection based on ICD-9-CM diagnostic code "36" as principal or any diagnosis indicating IMD • Incidence of sequelae • Nomenclature Calculations in KCE report 231

Parameter	Base case	95% CI	Distribution
*Acute care hospitalization costs – All serogroups**			
– **< 1 year**	7195	1055 – 19 104	Gamma (2.28, 3161.41)
– **1-4 years**	6243	1449 – 14 470	Gamma (3.36, 1856.83)
– **5-9 years**	5960	1869 – 12 382	Gamma (4.77, 1250.40)
– **10-19 years**	7918	1318 – 20 312	Gamma (2.50, 3163.16)
– **20+ years**	10 516	1032 – 30 741	Gamma (1.76, 5962.90)
Cost of follow-up care in those without sequelae	97.14	79.03 – 117.08	Gamma (100, 0.97)

Cost of follow-up care in those with sequelae varies by age:

1. Acute one-off costs:
   - 0-4 year: € 1434 [1044-1885]
   - 5-17 year: € 91 [66-119]
   - 18-19 year: € 45 [33-60]
   - 20-65 year: € 3445 [2508-4529]
2. Lifelong annual cost:

Between € 3215 [2340-4226] and € 3355 [2442-4411], depending of the age of the patient when affected by invasive meningococcal disease.

Scoring "current disease-related public expenditures per patient" (please check one box)

The disease currently has

☐ 0: a **very small** impact on public expenditures per patient

☐ 1: **some** impact on public expenditures per patient

☐ 2: a **high** impact on public expenditures per patient

☐ 3: a **very high** impact on public expenditures per patient

Comments (here you can write down concerns regarding the evidence level or other)

……..

Major depression

Major depression is a common mental health disorder, characterized by the loss of interest or pleasure in ordinary things and experiences, low mood and a wide range of associated emotional, cognitive, physical and behavioral symptoms. The identification and diagnosis of major depression (which is also called 'clinical depression' or just 'depression') is based not only on the severity of symptoms but also on their persistence, the presence of other symptoms, and the degree of functional and social impairment.

There is not a clear 'cut-off' between 'clinically significant' and 'normal' degrees of depression and it is best to consider the symptoms of depression as occurring on a continuum of severity; the greater the severity of depression, the greater the morbidity and adverse consequences. When adding other aspects that need to be considered, including duration, stage of illness and treatment history, there are considerable problems in attempting to classify depression into categories.

The aim of interventions for depression is to relieve symptoms, restore functions and, in long-term, prevent relapse. Treatment continues to be hampered by resistance at the individual level to seek help and the failure, especially in primary care, to correctly identify those who are truly depressed. The most common interventions (treatments) for depression, are psychological and/or pharmacological treatments.(KCE report 230)

Criterion: Impact of the disease on quality of life		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Anxiety/depression	Major depression impacts most the quality of life dimension "anxiety/depression". *Based on the EQ-5D tariff values for Belgium, the following conclusions can be drawn with respect to the impact of major depression on health-related quality of life:* • In Belgium, the EQ-5D score for people with major depression and no other health problems impacting their mobility, self-care, usual activities or pain, is 0.3847 on a scale from 0 (=dead) to 1(=no problem on any of the EQ-5D dimensions). This means that compared to people without major depression, patients with major depression have a utility loss of about 60%.	EQ-5D tariff values for Belgium, derived based on valuations from more than 2000 citizens of several health states described by means of the EQ-5D.
	Based on data of the control arm of randomized controlled trials studying the effect of new interventions compared to "therapy as usual": • Mean EQ-5D score with therapy as usual, acute phase: 0.397 (s.d. 0.097) • Mean EQ-5D at 8 months for people on the waiting list for therapy: 0.75 (s.d. 0.26)	Burns et al. 2013 (cf KCE report 230): 10 observations, women with antenatal depression

	• Mean scores SF-12 mental subscale (s.d.) (scale 0-100) at 6m: 33.7 (12.6), n=209 at 12m: 35.4 (12.8), n=195 • Mean scores SF-12 physical subscale (s.d.) (scale 0-100) at 6m: 42.1 (14.0), n=209 at 12m: 41.1 (13.5), n=195 • Mean QALY (quality-adjusted life year) Netherlands (s.d.) at 6m: 0.32 (0.14) at 12m: 0.65 (0.24)	Kessler et al. 2009: n=91, patients between 18-75 yrs of age, diagnosed with a new depressive episode within 4 weeks preceding referral, Beck depression inventory ≥14 Wiles et al. 2013, patients 18-75 yrs, mean age 49.6 yrs (s.d. 11.7), patients had adhered to adequate dose of anti-depressant medications Van Schaik et al. 2006 N=74, mean age 67.5 yrs (s.d. 9.2), primary care population.

Scoring "impact of disease on quality of life given current treatment" (please check one box)

The disease has

☐ 0: **no** negative impact on the quality of life of patients, compared to people without the disease

☐ 1: **some** negative impact on the quality of life of patients, compared to people without the disease

☐ 2: a **moderate** negative impact on the quality of life of patients, compared to people without the disease

☐ 3: a **high** negative impact on the quality of life of patients, compared to people without the disease

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Inconvenience of current treatment		
Aspects of current treatment causing inconvenience to patients	Magnitude of the inconvenience induced by each of these aspects	Sources of evidence
	Very little information. Antidepressant medication –if used for usual therapy- has known side-effects. David et al. 2008 reported that 9/49 patients receiving antidepressant medication experienced adverse effects: 1 patient had panic attacks, 2 patients had anxiety and insomnia, 1 patient experienced crying and anger, 2 patients had restlessness and 3 had insomnia. *Important note: As the tools we are developing are meant to help public resource allocation decisions, only currently (fully or partly) reimbursed treatments are considered as "current treatments". Non-reimbursed effective treatments –even if available since a very long time – are not necessarily accessible to all patients and can therefore not be considered to be "currently available" to all patients.*	David et al. 2008. n=49, Romania

Scoring "inconvenience of current treatment" (please check one box)

☐ 0: Current treatment is **not or only slightly** inconvenient to patients

☐ 1: Current treatment is **somewhat** inconvenient to patients

☐ 2: Current treatment is **highly** inconvenient to patients

☐ 3: Current treatment is **extremely** inconvenient to patients

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Impact of the disease on life expectancy		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Progression-free survival	Depression used to be viewed as a time-limited disorder, lasting on average 4 to 6 months with complete recovery afterwards. However, incomplete recovery and relapse are common. A WHO study of mental disorders in 14 centres across the world found that 50% of patients still had a diagnosis of depression 1 year later and at least 10% had persistent or chronic depression. At least 50% of people, following their first episode of major depression, will go on to have at least one more episode and, after the second and third episodes, the risk of further relapse rises to 70 and 90%, respectively. People with early onset depression (at or before 20 years of age) and depression occurring in old age have a significantly increased vulnerability to relapse. Thus, while the outlook for a first episode is good, the outlook for recurrent episodes over the long-term can be poor with many patients experiencing symptoms of depression over many years.	Doessel, Williams & Whiteford, 2009
Life years lost	Despite treatment, patients with major depression can still commit suicide. In a study by Strong et al. (2008) 1/99 patients with major depression and cancer committed suicide in a 12 month follow-up period. Suicide rate: 18.6/100 000n inhabitants in Belgium. Suicide accounts for 11% of potential years of life lost in men and 5% of potential years of life lost in women. According to WHO data depression is the fourth ranked cause of life years lived in disability in all age groups and second in the age group between 15 and 44 years.	Ustun, Ayuso-Mateos, Chatterji et al. 2004 ; WHO 2000

Scoring "impact of disease on life expectancy given current treatment" (please check one box)

The disease has

☐ 0: **no** impact on life expectancy

☐ 1: **some** impact on life expectancy (patient loses a small proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 2: a **high** impact on life expectancy (patient loses a large proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment)

☐ 3: a **very high** impact on life expectancy (patient dies almost immediately or 50% of the patients is dead within one week, despite the best available current treatment or care)

Comments (here you can write down concerns regarding the evidence level or other)

……

Criterion: Frequency of the disease		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Prevalence	6% of the adult population (= approximately 517 000 people) 5% year prevalence; 14.9% lifetime prevalence (Belgium)	Self-report WIV-ISP health interview survey 2008 Van Audenhove, Scheerder
Incidence	2,05/1000	Boffin et al. (cfr KCE report 230 on major depression)

Scoring "prevalence of the disease" (please check one box)

The prevalence of the disease is

☐ 0: less than 2000 people in Belgium have the condition (less than 1 per 5500)

☐ 1: between 2000 and 10 000 people in Belgium have the condition (between 1 per 5500 and 1 per 1100)

☐ 2: between 10 000 and 100 000 people in Belgium have the condition (between 1 per 1100 and 1 per 110)

☐ 3: more than 100 000 people in Belgium have the condition (more than 1 per 110)

Comments (here you can write down concerns regarding the evidence level or other)

……

Criterion: Disease-related public expenditures per patient		
Dimensions	Magnitude of each type of expenditures	Sources of evidence
	No information for Belgium A US study estimated that in 2003 the societal cost of depression, excluding the cost of productivity losses, was $20 billion[g]: direct healthcare costs and treatment costs accounted for $12 billion and costs related to mortality for $8 billion.	Wang, Simon et Kessler 2003
	For the Netherlands, the estimated direct healthcare cost per patient with major depression was estimated at €773, direct non-medical costs at €584, leading to a total direct cost per patient of €1357 in 2003.	Smit et al. 2006

Scoring "current disease-related public expenditures per patient" (please check one box)

The disease currently has

☐ 0: a **very small** impact on public expenditures per patient

☐ 1: **some** impact on public expenditures per patient

☐ 2: a **high** impact on public expenditures per patient

☐ 3: a **very high** impact on public expenditures per patient

Comments (here you can write down concerns regarding the evidence level or other)

g 1 Billion = 1 miljard / 1 milliard

Alzheimer

Clinical Alzheimer's disease (AD) is preceded by a slowly progressing accumulation in the brain of amyloid plaques and neurofibrillary tangles with hyperphosphorylated tau protein. In many cases AD is present in combination with some degree of cerebrovascular damage (mixed dementia). With the exception of some genetically well-defined forms of AD, at present a definitive diagnosis of AD still requires histopathological confirmation of a probabilistic clinical diagnosis.

Alzheimer's disease is an incurable brain disease, wherein cells in some parts of the brain cease to function and in some parts of the brains to function and die. Alzheimer's disease is irreversible and despite extensive research, the causes of the disease and treatment methods are still unknown. The symptoms of the disease are forgetfulness, personality changes, disorientation, and loss of speech. Alzheimer's disease is the most common form of dementia. The risk of dementia increases with age. A high age is the most important known risk factor for Alzheimer's dementia and for the most other forms of dementia.

Criterion: Impact of the disease on quality of life		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Self-care Usual activities Anxiety/depression	Memory impairment is usually one of the first symptoms of AD. As the disease progresses cognitive deficits start to interfere with activities of daily living (ADL) and behavioural problems may appear. These behavioural and psychological signs and symptoms in dementia commonly include depression, apathy, agitation, disinhibition, psychosis, wandering, aggression, incontinence and altered eating habits. They contribute significantly to caregiver burden, institutionalization (placement in an elderly home), and decreased quality of life for patients with dementia. Informal caregivers pointed out that the early stage of the disease is often the most difficult one for the patient. Before the diagnosis is made, the patient realises that there is something wrong, without knowing what exactly. Mostly patients react by trying to hide the malfunctioning. They start avoiding the activities causing them trouble and making excuses for things going wrong. Mostly this results in conflicts with family members and more isolation. Patients can react very anxious or even become violent, as they lose control, feel frustrated and depressed. As the disease progresses, quality of life deteriorates. Communication becomes difficult to impossible, verbally, but also reading and writing skills are lost. Mobility worsens to bedridden. Intellectual capacities disappear gradually.	KCE report 111 Interview representatives Alzheimer Liga
Comments	Caregiver burden The burden for informal caregivers is huge. They too get more isolated as the care for the patient takes all their time and energy. Caregivers claim that as the disease progresses the caregiver is more in need of emotional, practical and informational support than the patient. The step to residential care is very difficult. Informal caregivers experience it as a personal failure and it takes time to admit that they can no longer take care of the patient on their own.	Interview representatives Alzheimer Liga

Scoring "impact of disease on quality of life <u>of patients</u> given current treatment or care" (please check one box)

The disease has

☐ 0: **no** negative impact on the quality of life of patients, compared to people without the disease

☐ 1: **some** negative impact on the quality of life of patients, compared to people without the disease

☐ 2: a **moderate** negative impact on the quality of life of patients, compared to people without the disease

☐ 3: a **high** negative impact on the quality of life of patients, compared to people without the disease

Comments (here you can write down concerns regarding the evidence level or other)

……

Criterion: Inconvenience of current treatment or care		
Aspects of the current treatment or care causing inconvenience to patients	Magnitude of the inconvenience induced by each of these aspects	Sources of evidence
	Non-pharmacological interventions There is lack of standardization of non-pharmacological interventions in Alzheimer's disease. Many variations and combinations of interventions are evaluated in studies. The fact that few studies address the same issues in comparable ways is a major methodological problem, limiting the level of evidence which can be associated with such unique interventions. Non-pharmacological interventions include, for instance, cognitive stimulation or training, reminiscence therapy, "snoezelen" multisensory stimulation, music therapy, massage and touch, exercise or activity therapy, ergotherapy, etc. There is no evidence on the inconvenience of these interventions. Among these interventions, promising are cognitive stimulation (alone or add-on to therapy with inhibitors of acetylcholinesterase), ADL-rehabilitative care, music therapy, massage/touch, physical activity, and maybe interventions aiming at improving communication/interaction. For all the other interventions, insufficient high quality evidence is available to support or reject the therapy; including environmental changes like group living, intensified care programs or special dementia care units. Many patients with Alzheimer's disease are taken care off at home. Care at home for as long as possible is often preferred over institutionalization both by the patient and the family-caregiver. However, caring for a person with dementia at home is intensive and burdensome. Caregivers are at high risk of psychosocial morbidity and associated breakdown in care. Ample evidence is available on this impact. For interventions involving formal or informal caregiver(s), some moderate level of evidence was found for a positive effect of several forms of psychosocial interventions and psychoeducation on informal caregiver depression and stress. Education and training of staff were found to be promising interventions. Support measures preventing caregivers from becoming overburdened and depressed result in a delay of institutionalisation, as shown in a meta-analysis of 13 support programs. One example from this meta-analysis is the study by Mittelman et al., for which the longest follow-up period is available. This large randomized controlled trial (RCT) studied over a 9.5-year period 406 spouse caregivers of community dwelling AD patients in New York City. Enhanced counselling and support consisted of six sessions of individual and family counselling, support group participation, and continuous availability of ad-hoc telephone counselling. This intervention was associated with a delay in median time to placement of 557 days. In addition, self-rated health in intervention group caregivers was significantly better than control group caregivers.	KCE report

	Pharmacological treatment Drugs are not curative, but claim to slow down the deterioration of patients with dementia. Donepezil, Galantamine, Rivastigmine and Memantine are all drugs approved in Belgium for the treatment of dementia (apart from Memantine which is approved for moderate to severe dementia, all drugs are approved for mild to moderately severe dementia). Inhibitors of acetylcholinesterase (ChEIs) (donepezil, galantamine and rivastigmine) have dose-dependent adverse effects, including vomiting, nausea, diarrhoea and anorexia. These gastrointestinal adverse event rates limit dose increases of ChEIs. They can be partly avoided by means of a slower dose titration rate. Also cardiac side-effects (bradycardia, AV-block) have been reported but these are less frequent. It is well established that the use of both typical and atypical antipsychotics in patients with dementia is associated with an increased mortality rate. Their use should be restricted, e.g. to hostile, aggressive patients.

Scoring "inconvenience of current treatment or care <u>to patients</u>" (please check one box)

☐ 0: Current treatment is **not or only slightly** inconvenient to patients

☐ 1: Current treatment is **somewhat** inconvenient to patients

☐ 2: Current treatment is **highly** inconvenient to patients

☐ 3: Current treatment is **extremely** inconvenient to patients

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Impact of the disease on life expectancy		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
	The average life expectancy after diagnosis of AD is 5 to 6 years, or 8 years after the first symptoms. Life expectancy depends on the age at the moment of the diagnosis: as high as 10.7 years for the youngest patients (65-69 years) to a low of 3.8 years for the oldest (90 or older at diagnosis). Roughly, this is about half of the life expectancy of a person with the same age but without AD.	KCE report

Scoring "impact of disease on life expectancy given current treatment" (please check one box)

The disease has

☐ 0: **no** impact on life expectancy

☐ 1: **some** impact on life expectancy (patient loses a small proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment or care)

☐ 2: a **high** impact on life expectancy (patient loses a large proportion of his remaining life expectancy due to the disease, even if he/she gets the best available current treatment or care)

☐ 3: a **very high** impact on life expectancy (patient dies almost immediately or 50% of the patients is dead within one week, despite the best available current treatment or care)

Comments (here you can write down concerns regarding the evidence level or other)

……..

Criterion: Frequency of the disease		
Dimensions	Magnitude of the impact of the disease on each dimension	Sources of evidence
Prevalence	In 2005, the proportion of people with dementia in the European Union was estimated at 1.14% to 1.27% of the population. For Belgium the estimate was 1.22% to 1.35%, corresponding to 127 174 and 140 639 subjects.(www.dementia-in-europe.eu, consulted May 14, 2008). The number of Alzheimer's Disease patients in Belgium is not well documented but can be estimated at about 75 000 patients in 2008 (presentation Prof Patrick Santens, Ghent, 2008). This corresponds to about 55% of all dementia patients.	KCE report

Scoring "prevalence of the disease" (please check one box)

The prevalence of the disease is

☐ 0: less than 2000 people in Belgium have the condition (less than 1 per 5500)

☐ 1: between 2000 and 10 000 people in Belgium have the condition (between 1 per 5500 and 1 per 1100)

☐ 2: between 10 000 and 100 000 people in Belgium have the condition (between 1 per 1100 and 1 per 110)

☐ 3: more than 100 000 people in Belgium have the condition (more than 1 per 110)

Comments (here you can write down concerns regarding the evidence level or other)

……..

APPENDIX 6. VADEMECUM CONCEPTS AND TERMINOLOGY EVIDENCE TABLES

Astigmatism	Abnormal curvature of the cornea, where the eye does not focus light evenly onto the retina
Case fatality ratio	The proportion of deaths within a designated population of "cases" (people with a medical condition), over the course of the disease.
Certainty of the evidence	'Certainty of the evidence', 'quality of the evidence' and 'confidence in the estimate' are alternative expressions for an assessment of how good an indication the research provides of the likely effect; i.e. the likelihood that the effect will be substantially different from what the research found. "Substantially different" means a large enough difference that it might affect a decision. This assessment is based on an overall assessment of reasons for there being more or less certainty.
Coelomic cavity	The coelom refers to the main body cavity and is positioned inside the body to surround and contain the digestive tract and other organs.
Compartment syndromes	Conditions in which increased pressure within a limited space compromises the blood circulation and function of tissue within that space. Some of the causes of increased pressure are trauma, tight dressings, haemorrhage, and exercise. Sequelae include nerve compression; paralysis and ischemic contracture.
Data saturation	Data saturation is the point in qualitative data collection when no new or relevant information emerges with respect to the research question. This point of closure is arrived at when the information that is being shared with the researcher becomes repetitive and contains no new ideas, so the researcher can be reasonably confident that the inclusion of additional participants is unlikely to generate any new ideas.
Debridement	The removal of foreign material and devitalized or contaminated tissue from or adjacent to a traumatic or infected lesion until surrounding healthy tissue is exposed.
Edema	Abnormal fluid accumulation in tissues or body cavities. Most cases of edema are present under the skin in subcutaneous tissue.
EQ-5D	Descriptive system of health-related quality of life states consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) each of which can take one of three (of five) responses. In the three-level version, the responses record three levels of severity (no problems/some or moderate problems/extreme problems) within a particular EQ-5D dimension. In the five-level version, the responses record five levels of severity (no problems/slight problems/moderate problems/severe problems/extreme problems) within a particular EQ-5D dimension.

EQ-5D Tariff values	EQ-5D tariff values are utility values for all health states that can be described with the EQ-5D descriptive system for health states. The tariff values are obtained from the general public and range from 0 (=death) to 1(=perfect health)
Escharotomy	A surgical procedure used to treat full-thickness (third-degree) circumferential burns.
Evidence profile	A table summarising key information about all relevant outcomes of an intervention or option, including a summary of the effects and judgements about the certainty of the evidence
Frontotemporal degeneration	The most common clinical form of frontotemporal lobar degeneration, this dementia presents with personality and behavioural changes often associated with disinhibition, apathy, and lack of insight.
Healthy life expectancy	The number of remaining years that a person of a certain age with this disease is still supposed to live without disability
Hyperopia	Farsightedness
Incidence	Incidence is the rate of new (or newly diagnosed) cases of the disease. It is generally reported as the number of new cases occurring within a period of time (e.g., per month, per year). It is more meaningful when the incidence rate is reported as a fraction of the population at risk of developing the disease (e.g., per 100,000 or per million population).
Impact of disease on life expectancy	The consequences/impact of the disease/ health condition with regard to symptoms and clinical presentation, comorbidities, associated risks, healthy life expectancy, mortality and morbidity
Impact of disease on quality of life	the degree to which the quality of life is affected by this disease/health condition in terms of physical well-being, social/family well-being, emotional well-being and functional well-being
Intervention	An action or option, including preventive, therapeutic, rehabilitative, diagnostic, public health and health system interventions.
Lifetime prevalence	The number of individuals in a statistical population that at some point in their life have experienced a "case" (e.g. a disorder), compared to the total number of individuals (i.e. it is expressed as a ratio or percentage).
Main outcomes	All outcomes that are critical for a decision and those that are important, but not critical. Outcomes are good or bad events or developments that can happen after an intervention (option) or comparison
Meningitis	Inflammation of the coverings of the brain and/or spinal cord
Mobility	Ability to walk around
Myopia	Near-sightedness

Presbyopia	Age-related type of farsightedness, due to the hardening of the eye lens
Prevalence	The proportion of individuals in a population having a disease or characteristic. Prevalence is a statistical concept referring to the number of cases of a disease that are present in a particular population at a given time, whereas incidence refers to the number of new cases that develop in a given period of time. Prevalence is most meaningfully reported as the number of cases as a fraction of the total population at risk.
Self-care	Ability to wash and dress self
Sepsis	Systemic inflammatory response syndrome with a proven or suspected infectious aetiology.
Tracheostomy	Surgical formation of an opening into the trachea through the neck, or the opening so created.
Tunica vaginalis testis	The serous covering of the testis. It is a pouch of serous membrane, derived from the processus vaginalis of the peritoneum, which in the foetus preceded the descent of the testis from the abdomen into the scrotum.
QALY	Quality adjusted life year, a measure of the value of health outcomes that combines the length of life and quality of life into a single number e. It assigns a weight corresponding to health-related quality of life (e.g. based on health state utility values) to each year of life.
p-value	In statistics, the ***p*-value** is a function of the observed sample results (expressed as a test statistic) that is used for testing a statistical hypothesis. More specifically, the *p*-value is defined as the probability of obtaining a result equal to or "more extreme" than what was actually observed, assuming that the null hypothesis is true.
QLC-LC13	The European Organisation for Research and Treatment (EORTC) of Cancer Quality of Life Questionnaire Lung Cancer 13. It is an instrument to assess the quality of life of lung cancer patients.
QLQ-C30	Questionnaire developed by the EORTC to assess the quality of life of cancer patients.
Quality of Life (QoL)	Quality of life (QOL) is a broad multidimensional concept that includes subjective evaluations of both positive and negative aspects of life. Health Related Quality of Life (HRQoL) is narrower. It refers to the individual's satisfaction or happiness with domains of life insofar as they affect or are affected by health related aspects.
Quality of evidence (grading)	The quality of evidence indicates the extent to which one can be confident that an estimate of effect, as reported in a scientific paper, is correct. The quality/level of evidence is related to the study design, the study quality (details of methodology and execution), consistency (similarity of estimates across studies on the same topic), and directness (the extent to which the people, interventions, and outcome measures of a study are similar to those of interest). The strength of a recommendation depends on the level of evidence.
Qualitative research/studies	Qualitative Research is primarily exploratory research. It is used to gain an understanding of underlying reasons, opinions, and motivations. It provides insights into the problem or helps to develop ideas or

	hypotheses for potential quantitative research. Qualitative Research is also used to uncover trends in thought and opinions, and dive deeper into the problem. Common methods include focus groups, interviews, participation and observation.
Quantitative research/studies	Quantitative research is used to quantify a problem by way of generating numerical data or data that can be transformed into useable statistics. It is used to quantify processes, reactions, attitudes, opinions, behaviours, and other defined variables – and generalize results from a larger sample population. Quantitative Research uses measurable data to formulate facts and uncover patterns in research.
Relative effect	The ratio of outcome measures in one intervention (option) comparison group compared to another
Research evidence	Findings from research that has used systematic and explicit methods to address questions
RSCL	A Toll-like receptor signalling antagonist
Standard deviation (s.d.)	In statistics, the standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values. A standard deviation close to 0 indicates that the data points tend to be very close to the mean, while a high standard deviation indicates that the data points are spread out over a wider range of values.
Sequelae	A pathological condition resulting from a prior disease, injury, or attack.
SF-12	The SF-12 is a shorter version of the SF-36 that uses just 12 questions to measure functional health and well-being from the patient's point of view.
SF-36	A 36-item, patient-reported survey of patient health consisting of eight scaled scores, which are the weighted sums of the questions in their section. Each scale is directly transformed into a 0-100 scale on the assumption that each question carries equal weight. The lower the score the more disability. The higher the score the less disability i.e., a score of zero is equivalent to maximum disability and a score of 100 is equivalent to no disability. The eight sections are: vitality physical functioning bodily pain general health perceptions physical role functioning emotional role functioning social role functioning

	mental health
Standard Gamble	Respondents are asked to choose between remaining in a state of ill health for a period of time, or choosing a medical intervention which has a chance of either restoring them to perfect health, or killing them. The standard gamble results in a value between 0 and 1 reflecting the severity of the state of ill health. The lower the value, the worse the health state.
Subgroup *	A subset of a population who share one or more specified characteristics; e.g. women or children, for whom different judgements, and/or a different recommendation may be made
Systematic review	A summary of studies addressing a clear question, using systematic and explicit methods to identify, select, and critically appraise relevant studies, and to collect and analyse data from them
Target audience	People for whom the recommendation is intended or people affected by the decision
Usual activities	Work, leisure, study
Utility value *	A measure of how much people value an outcome (health state) in relation to other outcomes, where 0.00 represents death and 1.00 represents perfect health
Values	How much the people affected by an intervention or option value each of the outcomes. How much people value outcomes in relation to each other needs to be considered when weighing up the desirable effects of a treatment against the undesirable effects. Utility values are sometimes used as a measure of how much people value outcomes
VAS: visual analogue scale	A subjective psychometric response scale used to measure distinct behavioural or physiological phenomena based on linear numerical gradient or yes/no alternatives.

REFERENCES

1. Cleemput I, Devriese S, Kohn L, Devos C, van Til J, Groothuis-Oudshoorn K, et al. Incorporating societal preferences in reimbursement decisions - Relative importance of decision criteria according to Belgian citizens Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2014 22/12/2014. KCE Reports 234 Available from: http://kce.fgov.be/sites/default/files/page_documents/KCE_234_reimbursement_decisions_Report.pdf
2. le Polain M, Franken M, Koopmanschap M, Cleemput I. Drug reimbursement systems: international comparison and policy recommendations. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2010 2011-03-07 (2nd print; 1st print: 2011-01-11). KCE Reports 147C (D/2010/10.273/90) Available from: https://kce.fgov.be/sites/default/files/page_documents/KCE_147C_Drug_reimbursement_systems_4.pdf
3. Christiaens W, Kohn L, Léonard C, Denis A, Daue F, Cleemput I. Models for citizen and patient involvement in health care policy - Part I : exploration of their feasibility and acceptability. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2012. KCE Reports 195 (D/2012/10,275/105) Available from: https://kce.fgov.be/sites/default/files/page_documents/KCE_195C_citizen_and_patient_participation.pdf
4. Koninklijk Besluit van 21 december 2001 tot vaststelling van de procedures, termijnen en voorwaarden inzake de tegemoetkoming van de verplichte verzekering voor geneeskundige verzorging en uitkeringen in de kosten van farmaceutische specialiteiten, B.S. 29 december 2001. Arrêté royal de 21 décembre 2001 fixant les procédures, délais et conditions concernant l'intervention de l'assurance obligatoire soins de santé et indemnités dans le coût des spécialités pharmaceutiques, M.B. le 296 décembre 2001., 2001.
5. Raeymaekers P, Vandenbroeck P, Denis A, Van Steenberge J. Reimbursements in health care: An agenda for change. Brussels: King Boadouin Foundation, ; 2015.

6. Devlin N, Sussex J. Incorporating multiple criteria in HTA. Methods and processes. London: Office of Health Economics; 2011.
7. Klein R. Dimensions of rationing: who should do what? BMJ. 1993;307(6899):309-11.
8. Marsh K, IJzeman M, Thokala P, Devlin N. The ISPOR MCDA Task Force: How best to use it in health care decision making. [Presentation at the 20th Annual International ISPOR Meeting]. Philadelphia, PA, USA; 2015.
9. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and Value: Impact on DEcisionMaking--the EVIDEM framework and potential applications. BMC Health Serv Res. 2008;8:270.
10. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. Med Decis Making. 2012;32(2):376-88.
11. Raeymaekers P, Teller M. De mens achter de ziekte, de burger achter de beslissing. Welke zorg voor welke prijs: de burgers geven een leidraad voor terugbetalingen in de gezondheidszorg. Brussels: Koning Boudewijnstichting; 2015. D/2893/2015/03
12. Patient and Citizen Involvement Group. Patient and Citizen Involvement. [Web page].2015 [cited 11 December]. Available from: http://www.htai.org/interest-groups/patient-and-citizen-involvement.html
13. Facey K, Stafinski T. Beyond guidelines - tools to support patient involvement in Health Technology Assessment;. In: G-I-N, editor. G-I-N Public Toolkit: Patient and Public Involvement in Guidelines.; 2015. Available from: http://www.g-i-n.net/document-store/working-groups-documents/g-i-n-public/toolkit/toolkit-2015
14. Patient and Citizen Involvement Group.2015 [cited 11 December]. Available from: http://www.htai.org/interest-groups/patient-and-citizen-involvement/resources.html
15. Terwee CB, Roorda LD, de Vet HC, Dekker J, Westhovens R, van Leeuwen J, et al. Dutch-Flemish translation of 17 item banks from the patient-reported outcomes measurement information system (PROMIS). Qual Life Res. 2014;23(6):1733-41.
16. Becher K, Rohden C, Vandenbroeck P, Van den Hende W. Een patiëntgedreven innovatie-agenda. Vlaams Patiëntenplatform; 2014. Available from: http://vlaamspatientenplatform.be/_plugin/ckfinder/userfiles/files/Eindrapport%20-%20Pati%C3%ABntgerichtheid%20bij%20zorginnovatie.pdf
17. Grondin J, Blais J-G. A Rasch analysis on collapsing categories in item's response scales of survey questionnaire: Maybe it's not one size fits all. In: Annual meeting of the American Educational Research Association. Denver, CO, April 30 - May 4, 2010; 2010. p. 29.
18. Weijters B, Cabooter E, Schillewaert N. The effect of rating scale format on response styles: the number of response categories and response category labels. International Journal of Research in Marketing. 2010;27:236-47.
19. EVIDEM Collaboration. Decision Criteria. Conceptual background, definitions, design and instructions. 2013. v2.3 Available from: https://www.evidem.org/docs/2014/EVIDEM-v2-3-Decision-criteria-2013-12-24.pdf
20. Angelis A, Kanavos P. Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment. Pharmacoeconomics. 2016;34(5):435-46.