Edinburgh Research Explorer

Valuing Alzheimer's Disease drugs

International Journal of
Technology Assessment in Health Care

Valuing Alzheimer's Disease drugs: A health technology assessment perspective on outcomes

Objectives
Due to the nature of the condition, health technology assessment agencies might face considerable challenges in choosing appropriate outcome measures for Alzheimer's disease drugs. To analyse which outcome measures were used in past health technology assessments in three European countries: England, Germany, and The Netherlands; to explore possible reasons for prioritisations, and derive potential implications for future assessments of Alzheimer's disease drugs.
Method
We conducted a literature review of studies that analysed decisions made in health technology assessments (across disease areas) in the three European countries. We then conducted case studies of technology assessments conducted for Alzheimer's disease drugs in these countries.
Results
Overall, outcomes measured using clinical scales dominated decisions or recommendations about whether to fund Alzheimer's disease drugs, or price negotiations. Health technology assessment processes did not always allow the inclusion of outcomes relevant to people with Alzheimer's disease, their carers and families. Processes did not include

	early discussion and agreement on what would constitute appropriate outcome measures and cut-off points for effects. Conclusions To facilitate consistent and timely decisions about the value of new Alzheimer's disease drugs early agreement with various stakeholders about outcomes, outcome measures and cut offs is important to ensure that future AD drugs are appropriately valued.

SCHOLARONE™
Manuscripts

Title

Valuing Alzheimer's Disease drugs: A health technology assessment perspective on outcomes

Running title

Valuing Alzheimer's Disease drugs

Author details

Annette Bauer, MSc, Care Policy and Evaluation Centre, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom.

Raphael Wittenberg, MSc, Care Policy and Evaluation Centre, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom.

Amanda Ly, MSc, Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Nine Edinburgh BioQuarter, 9 Little France Road, Edinburgh, EH16 4UX, United Kingdom.

Anders Gustavsson, PhD, Quantify Research, Hantverkargatan 8, 112 21 Stockholm, Sweden; Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Alfred Nobels allé 23, 141 83 Stockholm, Sweden.

Christin Bexelius, PhD, ROCHE. F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland.

Claire Tochel, PhD, Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Nine Edinburgh BioQuarter, 9 Little France Road, Edinburgh, EH16 4UX, United Kingdom.

1

Martin Knapp, Professor in Health and Care Policy, Care Policy and Evaluation Centre, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom.

Mia Nelson, PhD, Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, 9 Little France Road, Edinburgh EH16 4UX, United Kingdom.

Catherine Sudlow, Professor of Neurology and Clinical Epidemiology, Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, 9 Little France Road, Edinburgh EH16 4UX, United Kingdom.

CORRESPONDING AUTHOR:

Annette Bauer, Care Policy and Evaluation Centre, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom; a.bauer@lse.ac.uk; Tel: +44 207 852 3784

2

Objectives

Due to the nature of Alzheimer's disease, health technology assessment agencies might face considerable challenges in choosing appropriate outcomes and outcome measures for drugs that treat the condition. This study sought to understand which outcomes informed previous health technology assessments, to explore possible reasons for prioritisations, and derive potential implications for future assessments of Alzheimer's disease drugs.

Method

We conducted a literature review of studies that analysed decisions made in health technology assessments (across disease areas) in three European countries: England, Germany, and The Netherlands. We then conducted case studies of technology assessments conducted for Alzheimer's disease drugs in these countries.

Results

Overall, outcomes measured using clinical scales dominated decisions or recommendations about whether to fund Alzheimer's disease drugs, or price negotiations. Health technology assessment processes did not always allow the inclusion of outcomes relevant to people with Alzheimer's disease, their carers and families. Processes did not include early discussion and agreement on what would constitute appropriate outcome measures and cut-off points for effects.

Conclusions

We conclude that in order to ensure that future Alzheimer's disease drugs are valued appropriately and timely, early agreement with various stakeholders about outcomes, outcome measures and cut-offs is important.

3

Key words

Dementia; cholinesterase inhibitors; Alzheimer's disease treatment; technology assessment,

health; priorities, health

Acknowledgements

Main text

Introduction

Many countries face the prospect of rapid increases in expenditure related to Alzheimer's

disease (AD). Governments are faced with the task of making decisions about which drugs

and interventions should be funded. Health technology assessment agencies or other

decision-making bodies are responsible for such decisions based on reviews of clinical and effectiveness or cost-effectiveness evidence through a process called health technology assessment (HTA). Assessing the effectiveness or cost- effectiveness for approved AD drugs has been difficult because AD drugs have historically promised only very small or no effects in functional improvement or modifying disease progression (1,2). Whilst HTA processes vary by country, they have in common that evidence on effectiveness or cost-effectiveness is reviewed by a technical team and interpreted by a group of stakeholders, who present different perspectives such as those of clinicians, drug companies, patient representatives and researchers. Which outcomes and outcomes measures influence final decisions is likely to be based on various criteria including: whether they reflect meaningful changes in a person's life (which is important from the perspective of people living with the condition, their families and carers); whether they are measurable in study designs (which is important from a developer and manufacturer perspective); and whether they are clinically and economically relevant (which is important from a payer perspective). Processes leading to decisions are complex, and are likely to vary between countries. The aims of our study were to understand: (1) which outcomes and outcomes measures are likely to be prioritized in HTAs for AD drugs in different countries, and (2) which processes influence these priorities. This study complemented other work which sought information on outcome prioritization from the perspective of patients, carers and practitioners (3, 4).

Methods

Overall approach and selection of countries

5

We employed two methods: a literature review and case studies. Findings of the literature review informed the design of the case studies. Both methods are explained in more detail below. Researchers with methodological expertise in systematic reviews (CT, AL) and qualitative research interviews (MN) as well as researchers specialized in medicine and neurology (CS), with knowledge of drug reimbursements and of HTA processes (CB, AG), and of dementia policies and economics (MK, RW) were involved in reviewing the methods throughout the research. In addition, the advisory group of the larger research programme of which this study was a part and which consisted of members from HTA or regulatory agencies across the world commented formally on initial findings. Three European countries were selected: England, Germany and The Netherlands. The respective HTA agencies are the National Institute for Health and Care Excellence (NICE) in England, the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany and the Dutch Zorginstituut Nederland (ZIN). The choice was influenced by the size of the economy and roles and responsibilities of HTA agencies with the aim to have multiple perspectives: England and Germany present two large economies in Europe, in which HTA agencies have taken on different roles and responsibilities. For example, whereas in England drugs need to be cost-effective in order to be publicly funded (5,6), in Germany decisions about whether drugs are funded and at what price are primarily based on their added therapeutic benefit (7). The Netherlands, as a relatively small economy in Europe, has taken a middle ground approach in this regard: the cost-effectiveness of drugs needs to be proven if their cost is above a certain threshold (8). An overview of the main features of the HTA agencies in the three different countries is shown in Table 1.

6

--- Table 1 about here ---

Data collection and analysis

First, we conducted a literature review of studies which analyzed how outcomes are

prioritized during HTA processes in the three countries. <mark>For the purpose of the literature</mark>

<mark>review, we pragmatically defined prioritized outcomes (and their measures) as those that</mark>

<mark>informed the final decision about whether the drug gets funded, or about its price. We</mark>

<mark>made this decision based on initial searches, which showed how the issue has been</mark>

<mark>investigated in the literature of HTAs.</mark> Since we expected that there would be limited

evidence from the AD field, we searched for studies across disease areas. Details on search

strategies, review and data extraction methods can be found in Supplement 1. Second, we

gathered data on how outcomes and outcome measures had been prioritized in past HTAs

of AD drugs. We conducted case studies based on information available on HTA websites,

which documented the decision processes from the beginning to final recommendation.

<mark>Here, we conceptualized 'prioritization' as the process of deriving decisions about which</mark>

<mark>outcomes and measures should inform the value of AD drugs. We therefore considered any</mark>

<mark>evidence of how decisions were made including views and opinions expressed by</mark>

<mark>stakeholders about the importance of certain outcomes and measures, and how they</mark>

<mark>thought they should inform the decision about the value of AD drugs.</mark> Information was

extracted on topics relevant to outcomes and outcome measures considered in the

appraisal. The framework for case studies and the data extraction form can be found in

Supplement 2. The analysis was a thematic one, in which we used a mix of inductive and

deductive methods for deriving themes.

7

About the data sources

Our literature review identified a total of 32 studies for the three countries. Thirteen

studies referred to England; fourteen to Germany (this included one study which also

referred to England); and six to The Netherlands. Studies used the following types of

methods: quantitative analysis using statistical methods (n=10); qualitative or mixed

methods (n=16); literature reviews (n=3); opinion papers or editorials (n=4). The main data

sources were HTA reports and documentation of the decision processes from HTA agencies'

websites and interviews. Details of studies can be found in Supplement 3.

The case studies referred to publicly available documentation of HTAs for AD drugs

(cholinesterase inhibitors and memantine) carried out between 2006 and 2010 in each of

the three countries. This included altogether 6 HTAs: England (n=1; covering cholinesterase

inhibitors and memantine together); Germany (n=2; one for cholinesterase inhibitors and

one for memantine); Netherlands (n=3; memantine; donzepil; rivastigmine for Parkinson's

disease). What was documented varied widely between HTAs and countries but covered at

a minimum:

- draft and final scope (including an agreed set of outcomes and outcome

    measures);

- draft and final appraisal of the reviewed evidence (including decisions or

    recommendations);

- consultation comments by stakeholders to draft scope and appraisal.

8

The list of documents that were identified as well as a list of stakeholders involved in the HTA processes can be found in Supplements 4 and 5; no documentation was available for stakeholder consultation in the Netherlands' HTAs.

Results

A range of evidence related to relevant outcomes and outcome measures was collated under eight themes. The purpose of the collation was to have distinguishable themes that reflected the different aspects covered in the case studies and the literature. The themes are related and to a certain degree overlapping. The findings for each will be described in turn, and we refer in brackets to the numbered data source, which can be found in the Supplements 3 and 4.

Cost-effectiveness

In England, decisions about whether to fund AD drugs were based on cost-effectiveness, which in turn was based on health-related quality of life (in the form of quality adjusted life years measured with the EQ-5D) and institutionalization (Suppl. 4: 4.5). No other economic consequences (e.g. for hospital care) were included or discussed. Both, health-related quality of life and institutionalization, were in additional analysis extrapolated from clinical scales for cognition and functioning (Suppl. 4: 4.1; 4.2). In Germany and in The Netherlands, no additional economic analysis and no review of economic evidence was conducted, and there was no mention of cost-effectiveness in the scoping documents (Suppl. 4: 4.12; 4.13; 4.17-1.19). This partly reflects the different approaches in the three countries towards including cost-effectiveness evidence in HTAs (Suppl. Table 1: 3.5): Germany does not include cost-effectiveness in their HTAs. In The Netherlands, the prices of the drugs were considered 'too low' to justify the need for cost-effectiveness considerations i.e. as long as

9

they had additional value and no adverse consequences they would be funded (personal

communication with ZIN representative).

Quality of Life (QoL)

There were differences in the ways HTA agencies responded to challenges of measuring

QoL for people with AD: NICE allowed the prediction of QoL in the form of economic

modelling based on surrogate outcomes measured with clinical scales. This approach was in

contrast to the one taken by IQWiG, which does not accept the use of QoL measures like

the EQ-5D and which has been consistently found to rarely accept QoL evidence (Suppl.

Table 3: 3.14 3.21; 3.23; 3.25). Methodological requirements (such as a minimum follow up

rate of 70%) frequently lead to the exclusion of evidence. Based on this and other

methodological requirements not met by studies, IQWiG concluded that the evidence of an

impact of AD drugs on QoL was insufficient (Suppl. Table 4: 4.14-4.16; 4.23-4.25). The

resulting exclusion of QoL outcomes in the appraisal of AD drugs was criticized by some of

the stakeholders (Suppl. Table 4: 4.12; 4.15; 4.16; 4.18; 4.19; 4.21). ZIN, whilst generally

accepting and prioritizing QoL evidence including when measured through the EQ-5D [8],

did not review QoL evidence in their HTAs of AD drugs (Suppl. Table 4: 4.26-4.28). We were

unable to find an explanation.

Outcomes measured with clinical scales (O-CS)

A wide range of outcomes were measured with clinical scales. Table 2 presents an overview

of the scales used in studies reviewed for the technology assessments. Not all scales,

however, informed the advice or decisions about the value of AD drugs equally. In all three

10

countries O-CS such as cognition (measured, for example, with the Alzheimer's Disease Assessment Scale-cognitive subscale; ADAS-cog) or functioning (measured, for example, with Activities of Daily Living scales; ADL) had an important influence on final decisions (Suppl. 4: 4.4; 4.15; 4.24). In the HTAs in England there was less evidence of stakeholder discussion about their relevance to people with AD (Suppl. 4: 4.7-4.10). The surrogate nature of those outcomes was made explicit in NICE's documentation (Suppl. 4: 4.1-4.4). The debate about the relevance of O-CS for people with AD was strongest in German HTAs (Suppl. 4: 4.20; 4.21). Whilst manufacturers argued the importance of clinical outcomes - in particular cognition - as reliable indicators of QoL with good psychometric properties, some stakeholders doubted whether clinical scales measured something that was meaningful to individuals (Suppl. 4: 4.20; 4.21). Both, IQWiG and the Federal Joint Committee (Gemeinsamer Bundesausschuss; G-BA), the body that makes the final and legally-binding decision about which drugs are funded, appeared to treat all O-CS as final health outcomes (Suppl. 4: 4.12-4.14), which meant that they bypassed some of their stricter methodological requirements that would have applied if they had been treated as surrogate outcomes (Suppl. Table 3: 3.24; 3.27). In terms of specific measures, IQWiG did not accept the use of global assessment outcomes (measured for example with the Clinician Interview-Based Impression of Change; CIBIC), which were seen as reflecting the clinician's perspective rather than the perspective of the person with AD (case studies). Instead, they expressed a preference for measures which evaluated personal goal attainment such as the Goal Attainment Scale (Suppl. 4: 4.14). ZIN pragmatically accepted those O-CS that had been accepted by the European Medicines Agency as validated outcome measures (Suppl. Table 3: 3.22; Suppl. 4: 4.26-4.28). This excluded the Mini-Mental State Examination (MMSE) as a measure for cognition, which was in contrast to NICE, which accepted its use as a main

11

outcome measure for modelling final QoL endpoints (Suppl. 4: 4.3;4.4; 4.26-4.28). ZIN

noted that the wide range of outcome measures across different domains made the

comparison of findings from studies difficult (Suppl. 4: 4.26-4.28).

--- Table 2 about here ---

During HTAs of AD drugs in all three countries, stakeholders raised concerns about how to

interpret the identified (often very small) changes on clinical scales (Suppl. 4: 4.7-4.11; 4.16;

4.20). Several stakeholders argued that there was need for greater clarity, from the

beginning, on cut-off points on various scales (Suppl. 4: 4.7-4.11; 4.16; 4.20). They should:

be based on evidence, reflect disease severities, and be relevant to people with dementia.

NICE tried to address the challenge of low effect sizes by giving particular weight to multi-

domain changes (i.e., a simultaneous change in different scales). IQWiG considered every

single outcome separately and as a result came to more conservative conclusions about the

value of drugs (i.e. they concluded more uncertainty about their effectiveness), resulting in

criticism from the drug manufacturers (Suppl. 4: 4.16; 4.20). In the Netherlands, ZIN

expected manufacturers to set out and justify relevant cut-offs before conducting studies

(Suppl. 4: 4.26-4.28). Responding to the uncertainty over clinical relevance and relevance to

people with AD, it decided to make the introduction of the drugs subject to start and stop

criteria and delegated the application of those to clinicians (Suppl. 4: 4.26-4.28).

Adverse effects

In England, benefit-harm considerations were not given much weight during HTAs possibly

because safety concerns were addressed already as part of market authorization and

aspects of adverse effects were thought to be captured in QoL outcomes (Suppl. Table 3:

12

3.11). Some stakeholders felt that the better tolerability of AD drugs, when compared to alternative treatments (such as antipsychotics), was undervalued in this approach (Suppl. 4: 4.11). In contrast, adverse effects were regarded as important outcomes from the perspective of people living with AD in Germany and the Netherlands (Suppl. 4: 4.14; 4.16; 4.26-4.28). Stakeholders of the HTAs carried out by IQWiG criticized the lack of long-term safety data on AD drugs and raised concerns about whether adverse effects had been underestimated (Suppl. 4: 4.11). In The Netherlands, ZIN sometimes left benefit-harm decisions to clinicians, as it concluded that the evidence did not support general conclusions (Suppl. 4: 4.26-4.28).

Outcomes relevant to people with AD

In both England and Germany, stakeholders (mainly patient representatives but also researchers and commissioners) argued that many outcomes relevant to people with AD were not being picked up by the clinical scales (Suppl. 4: 4.2; 4.8; 4.11; 4.21; 4.23; 4.25). They advocated for including more tangible outcomes (e.g. ability for someone to pick up the phone) as well as long-term outcomes (e.g. institutionalization). An ability to maintain aspects of personal identity was seen as another important outcome. Stakeholders highlighted an urgent need for appropriate outcome measures in early stages of AD (Suppl. 4: 4.8; 4.11).

In both countries, stakeholders thought that this required more flexible approaches towards including evidence (Suppl. 4: 4.8; 4.21; 4.23; 4.25). Whilst this need for different and more flexible processes was to a large extent shared by NICE, in Germany IQWiG and

13

G-BA believed that such changes would contradict legislation and reduce the necessary

methodological robustness (Suppl. 4: 4.21; 4.23; 4.25).

In the Dutch HTAs for AD drugs, the challenges of considering outcomes that mattered to

people with AD, their carers and families were not documented but had been - according to

a ZIN representative - discussed at several stages of the process (personal communication).

Carers' outcomes

In NICE's HTAs of AD drugs, carers' QoL was a primary outcome or endpoint (Suppl. Table 4:

4.1; 4.3), reflecting the priority given by NICE to this group. However, final decisions were

based on an economic model, which did not include carers' outcomes, an omission which

was criticised by some stakeholders (Suppl. 4: 4.7; 4.11). In Germany, carers' outcomes

were not viewed as the responsibility of the healthcare system and were given lower

priority relative to outcomes for people with AD (Suppl. 4: 4.21; 4.25). Whilst some

stakeholders argued for including carers' outcomes in its own right, there seemed to be an

overall consensus that carers' outcomes were important mainly because of their impact on

the person with AD (Suppl. 4: 4.21; 4.25). In addition, IQWiG was skeptical about carer-

reported outcomes for the person with dementia, which they argued reflected the needs of

the carer rather than the needs of the person with dementia (Suppl. 4: 4.21; 4.25). Dutch

HTAs for AD drugs did not include carers' outcomes (Suppl. 4: 4.26-4.28).

Institutionalization

Institutionalization for someone with AD was a stated outcome in HTA agencies'

documentation and discussed as important by stakeholders in Germany and England

(Suppl. 4: 4.1; 4.3; 4.12; 4.13; 4.17-4.19). Only English HTAs included institutionalization as

an outcome in the economic modelling although stakeholders discussed whether it was

possible to accurately predict this outcome since there were many other correlated factors

such as the carer's situation and availability of care in the community (Suppl. 4: 4.8) . In

German HTAs institutionalization was viewed by some stakeholders as an outcome that was

primarily important from an economic perspective (Suppl. 4: 4.21; 4.25). Some stakeholders

thought that 'institutionalization' could not be measured separately from 'time spent

caring', and that instead 'hours of care provided' should be measured independently of

whether they were provided by a professional or by an unpaid carer (Suppl. 4: 4.21; 4.25).

Similar to the discussion in English HTAs, stakeholders discussed the lack of evidence on

these outcomes and methodological challenges of including them. In The Netherlands there

was no recorded information on these outcomes.

Table 3 presents an overview of the findings. We applied categories indicating if an

outcome or group of outcomes was prioritized or not prioritized. 'Prioritized' outcomes

were those that informed decisions and 'Not prioritized' outcomes were those that did not

inform decisions.

---- Table 3 about here ----

Discussion

This study assessed the outcomes and outcome measures that dominated HTAs of AD drugs in three European countries, and the processes that influenced those priorities. This is to our knowledge the first study, which examines how outcomes and measures for AD drugs are currently prioritised in technology assessments. This study contributes to an increased transparency about reasons for and challenges of including certain outcomes when assessing the value of AD drugs. Overall, we identified some challenges in the process of how outcomes, outcome measures and cut-off points were defined in technology assessments of AD drugs. This included a lack of early involvement of stakeholders in discussions of appropriate outcomes and outcome measures as well as of cut-off points for appropriate effect sizes. In addition, a narrow focus on evidence from certain types of studies, namely randomised controlled trials, led to a strong focus on outcomes measured with clinical scales to the potential exclusion of (long-term) outcomes relevant for people with AD.

Our study was exploratory in nature, and we chose to conduct two methods to address the gap in evidence about the role of outcomes and outcome measure in HTAs of AD drugs. We first reviewed studies that analysed the influence of outcomes and measures on decisions of the value of drugs in HTAs. Whilst this provided useful knowledge about common decision making patterns in HTAs (and reasons for those), it provided only limited information about the process, by which decisions were made about outcomes and measures, and the process by which they influenced decisions. Whilst we have no affirmative knowledge of the reason

for this missing focus of studies, it is plausible that decisions about outcomes and measures are regarded objective or neutral. It is also likely to reflect a wide acceptance of outcomes measured with clinical scales as patient-relevant. As a result, designers of studies and manufacturers have to make decisions about outcomes and measures without certainty whether those will be accepted by HTAs. In the case of HTAs for AD drugs, this is likely to have contributed to the use of a wide range of measures. With the second method, the case studies, we therefore sought to address the gap in evidence about the process by which outcomes and measures are influencing decisions through in-depth analysis of reports produced for technology assessments. This kind of analysis allowed us to understand the nature of decision processes, and stakeholder viewpoints. Although this study was explorative in nature, we were able to shed new light on the important, currently under-investigated role of outcomes and outcome measures in influencing the value of AD drugs.

In terms of methodological robustness, the literature review, although pragmatic, applied systematic search strategies and involved detailed data extraction. Researchers with a high and diverse level of methodological and clinical expertise were involved in and contributed to the robustness of the research process. Approval of the research methods and interpretation of the findings was provided by experts in the field. In terms of limitations, for the case studies, we were reliant on publicly available information, which was limited, especially for the Dutch case studies. Furthermore, by focusing on HTAs with the most comprehensive information and those that were most comparable between countries, we might have missed some aspects of more recent updates of HTAs. Overall, our findings need to be interpreted in the context of a rapidly evolving field. Considerations that decision makers need to take into account today may very well change in the future, for example in light of new evidence and new technologies.

17

The findings from this study suggest that there are substantial challenges in including outcomes relevant to people with AD when assessing the value and cost-effectiveness of AD drugs. Those challenges are not only relevant to existing AD drugs but to other types of treatment and interventions, which seek to prevent or alter the progression of AD. Unless there is an agreed set of outcomes, outcome measure and cut-offs that define a meaningful diversion from the path without intervention, it will be challenging to assess the value of a drug or an intervention (in particular in relation to other interventions). In the future, this is likely to be relevant to pricing or investment decisions for disease-modifying treatments, which may need to be offered at pre-dementia stages, and which would require measuring surrogate outcomes such as imaging or other biomarkers (9). Without outcome measures that are acceptable to relevant stakeholders – including patients, carers and the wider public - and agreed before HTAs are conducted or preferably even before studies are being developed, there is a risk of delays in the appropriate evaluation of, and access to, new treatments (10). Clear methodological guidance on accepted outcome measures in fields such as prevention and diagnostics is therefore needed (11).

This includes the need to consider patient-relevant outcomes in HTAs in addition to clinical outcomes (12). Whilst in early stages, innovative methods have been developed (and tested) that allow HTA agencies to consider patient preferences over different outcomes when developing methodological guidance (12). Knowledge is also becoming increasingly available about how to best include patient and carers' perspectives in HTAs (13). Decisions about the value of drugs in HTAs in some countries (including England) have shown to be substantially influenced by aspects of value not captured by clinical and economic evidence (14). Whilst this is a reflection of including patient and stakeholders perspectives it also raises questions about transparency and consistency of decisions (14). Therefore, including

18

outcomes, measures and cut-offs that are more patient-relevant (and agreeing on those in advance) is likely to contribute to more consistent decision making as it reduces the need for additional considerations that in effect address the issue of evidence not being sufficiently relevant to what matters to patients, carers and the wider public.

Furthermore, the challenges we identified suggest a need for collaborative approaches between multiple stakeholders to enable decisions on outcomes and measures to be made early in the process. Some of the required processes are already in place, to varying degrees in different countries, whilst others still need to be developed.

Such multi-stakeholder approaches should go hand in hand with including wider sets of evidence, often referred to as real-world evidence (15). This requires an investment in data that can be used to demonstrate long-term impact on costs and outcomes (1). This might include data on the costs associated with different rates of disease progression so that cost savings linked to a delay in disease progression can be estimated. Findings from a study that modelled the likely cost-effectiveness of disease-modifying treatments (should they become available) showed that in England the benefit from deferring onset by one year would be substantial at about £28,000 (in 2012/13 prices) (16). This highlights the importance of including such data in decision making. Unless the impacts on disease progression, QoL, need for care and costs over time are considered, there is a risk that that future AD drugs and interventions are not valued in line with patient, carers and wider public interests.

Conclusions

This study investigated the role of outcomes and outcome measures in HTAs of AD drugs in three European countries. The findings highlight the strong priority placed on outcomes measured with clinical scales as well as the challenges of considering measures that capture

changes in disease progression that are potentially relevant from the perspective of people living with the condition, their families and carers. We conclude that there is an urgent need to reform HTA processes to appropriately assess the value of AD drugs.

Conflicts of interest

Acknowledgements

References

[1] **Bradley P, Akehurst R, Ballard C, et al.** Taking stock: A multi-stakeholder perspective on improving the delivery of care and the development of treatments for Alzheimer's disease. *Alzheimers Dement.* 2015;**11**, 455-461.

[2] **Goldman DP, Fillit H, Neumann P.** Accelerating Alzheimer's disease drug innovations from the research pipeline to patients. *Alzheimers Dement*. 2018;**14**:833-836.

[3] **Tochel C, Smith M, Baldwin H, et al.** ROADMAP consortium. What outcomes are important to patients with mild cognitive impairment or Alzheimer's disease, their caregivers, and health-care professionals? A systematic review. *Alzheimers Dement. 2019;* **11**:231–247.

[4] **Nelson M, Smith M, Ly A, et al.** [Internet] ROADMAP WP2 partners. D2.3 and D2.4 Stakeholder generated lists of priority Real World Evidence relevant outcomes and Disease progression and outcomes classification matrix. [cited 2020 March 8]. Available from:

https://roadmap-alzheimer.org/wp-content/uploads/2018/07/ROADMAP_D2.3D2.4.pdf

[5] **Drummond M.** Clinical guidelines: A NICE way to introduce cost-effectiveness considerations. *Value Health.*2016;**19**:525-530.

[6] **National Institute for Health and Care Excellence**. Social Value Judgements: Principles for the Development of NICE Guidance (2nd ed.). London: National Institute for Health and Care Excellence. 2008.

[7] **Paris V, Belloni A**. Value in pharmaceutical pricing. OECD Health Working Paper. Paris: Organisation for Economic Cooperation and Development. 2013.

[8] **Nederland Zorginstituut**. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. Diemen: Zorginstituut Nederland. 2015.

[9] **Foster NL, Hackett JSM, White G, et al**. Justifying reimbursement for Alzheimer's diagnostics and treatments: Seeking alignment on evidence. *Alzheimers Dement*. 2014;**10**: 503-508.

[10] **Ruof J, Knoerzer D, Dunne AA, et al.** Analysis of endpoints used in marketing authorisations versus value assessments of oncology medicines in Germany. *Health Policy.* 2014;**118**:242-254.

[11] **Versteegh M, Knies S, Brouwer W**. From Good to Better: New Dutch Guidelines for Economic Evaluations in Healthcare. *Pharmacoeconomics*. 2016;**34**:1071-1074.

[12] **Kinter ET, Schmeding A, Rudolph I, et al**. Identifying patient-relevant endpoints among individuals with schizophrenia: An application of patient-centered health technology assessment. *Int J Technol Assess Health Care*. 2009;**25**: 35-41.

[13] **Perfetto E, Boutin M, Reid S, et al.** OP38 Improving The Patient Centricity Of Value Assessments: A Rubric. Int J Technol Assess Health Care. 2017;32:17-18.

[14] **Nicod E, Kanavos P.** Scientific and social value judgments for orphan drugs in health technology assessment. *Int J Technol Assess Health Care* 2016;**32**:218-232

[15] **Makady A, de Boer A, Hillege H et al.** What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. *Value Health*. 2017;**20**:858–65

[16] **Anderson R, Knapp M, Wittenberg R et al.** Economic Modelling of Disease-Modifying Therapies in Alzheimer's Disease. London: Care and Policy Evaluation Centre (formerly: Personal Social Services Research Unit). 2018.

22

Legend

Table 1: Features of health technology assessment (HTA) agencies in England, Germany, and The Netherlands

Table 2: Clinical scales used in studies identified in health technology assessments (HTAs) of Alzheimer's disease drugs in England, Germany, and The Netherlands

Table 3: Prioritization of outcomes and outcome measures in health technology assessments (HTAs) of Alzheimer's disease (AD) drugs

23

Supplementary file

Supplement 1: Details on search strategies, inclusion and exclusion criteria, data extracted and quality appraisal

The following databases were searched for articles published between 2007 and November 2017: CINAHL, MEDLINE, SocScience, EconLit, Elsevier Science Direct. The following subject headings and keywords were used: outcome-related term (i.e. outcome OR benefit OR effect OR endpoint) AND country-related term (i.e. Germany OR Netherlands OR England) AND a technology assessment-related term (i.e. benefit assessment OR technology assessment). If the number of results was particularly large, we added an additional search term for decision making process (i.e. process OR decision making).

In addition, the smart search (CINAHL), recommendations based on previously read articles (Elsevier Direct) and articles frequently cited together (PubMed) functions of databases were used. Additional searches in journals of particular relevance such as 'Value in Health' and 'Medical Decision Making' were also carried out. A few reference searches for key articles were carried out to test if all relevant articles were captured in the searches.

Articles were included that referred to decisions, processes and standards of health technology assessments if they made reference to the role of outcomes. Excluded were articles, which were critical discussions about the use of specific methods - such as: the quality-adjusted life years (QALY) measure; social discount rates in economic evaluation methods; multi-criteria decision making – or which related to personalised medicine, described the influence of HTA processes on market access to drugs or focused on price setting mechanisms and negotiations.

The following information was extracted for each study: study ID; setting; purpose; design; type of data and analysis method; further details about methods (where required); results; conclusions and limitations stated by author(s). For each study a rating was generated to reflect its relevance for our research questions.

In a next step, information was summarised for each country using the following headings (which were identified during the initial analysis of information):

- Responsibilities of HTA and other relevant agencies in regards to HTA or reimbursement process;
- HTA process and requirements;
- Decision making process and criteria;
- Price negotiations and status of listing decision;
- Stakeholder involvement in process;
- Surrogate and composite outcomes;
- Quality of life (and quality-adjusted life years);
- Cost-effectiveness;
- Sub groups.

Supplement 2: Case study framework and data extraction form

Case study framework

Case studies were carried out for health technology assessments / appraisals (HTA) carried out in the dementia/ AD field in three countries: England, Germany and Netherlands.

For England the case study referred to one Multiple Technology Appraisal for donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (AD)[1], which was published 23 March 2011 with last update 11 May 2016. Relevant publicly available documents were included to inform the case studies, such as:

- Guidance and appendices;
- Research recommendations information;
- Documents produced as part of the guidance development such as:
  - Background information (includes review decision documents, press releases);
  - Assessment report documents;
  - Draft and final protocol documents;
  - Draft and final matrix documents;
  - Draft and final scope documents (including consultation comments);
  - Appraisal consultation documents (e.g. Assessment reports; Consultee and commentator comments on the assessment report; Manufacturer and Non-manufacturer Submissions; Expert written personal statements;
  - Final appraisal determination documents (including comments on appraisal consultation

In Germany case studies referred to the following 3 single drug benefit assessments: Memantine in AD; Cholinesterase inhibitors in AD; Ginkgo compounds. There were no technology appraisals in form of early benefit assessments carried out for dementia/ AD drugs since introduction of the new legislation (AMNOG) in 2011. Instead, all appraisals refer to drug assessments carried out before AMNOG. Relevant information included the following documents from the IQWiG website:

- Final and preliminary reports
- Documentation and appraisal of comments on the preliminary report
- Report plan (different versions) and amendments
- Documentation and appraisal of comments on the report plan
- Executive summary of the working paper 'Memantine in Alzheimer's disease: Results of the unpublished studies IE2101 and MEM-MD-22 as well as unpublished responder analyses'
- Press releases

---

[1] https://www.nice.org.uk/guidance/ta217

We also looked at the G-BA website for manufacturers' value decisions, G-BA value decisions (Tragende Gruende). Further information about decisions and the role of clinical endpoints in those decisions were also available online[2] [3] [4] [5].

In The Netherlands, case studies referred to short pharmacotherapeutic reports for donezepil (for the indication and symptomatic treatment of mild to moderately severe Alzheimer's dementia)[6]; rivastigmin for people with Parkison's disease and memantine.

Across case studies, the following information were extracted with respect to the following questions:

Study endpoints (mortality/ morbidity/ quality of life)	Which endpoints were set out during scoping?
	Which primary endpoints were used in studies that supported the recommendation? (This might include information about categorised clinical endpoints and clinical scales)
	Which surrogate endpoints were used that supported recommendations, which methods of validation were used? Did the Committee discuss the appropriateness of surrogate endpoint as validated indicators of endpoints?
	How were patient preferences (satisfaction, adherence, complaints) and patient reported outcomes considered?
	Which endpoints were considered in cost-effectiveness analysis that supported recommendations? → How was clinical evidence mapped to final endpoint quality of life (in cost-effectiveness analysis)?
	Were aspects of meaningful delay and disease progression considered in endpoints?
Stakeholder views and influence	Which suggestions were made in regards to clinical endpoints?
	Which challenges around including relevant clinical endpoints were discussed by stakeholders?
	Were any clinical endpoints were considered differently as a result of stakeholder involvement?
	Which clinical endpoints were identified as relevant for future research?
Uncertainty	How did uncertainty in data influence discussions about outcome? Were there any criteria or rationales that made an uncertain outcomes more acceptable?
Threshold	Which thresholds were applied in regards to clinical measures and/or cost-effectiveness?
	How were benefit harm ratios considered?

Data extraction form

[2] https://www.iqwig.de/en/press/press-releases/press-releases/long-struggle-for-appropriately-processed-manufacturer-data-leads-to-a-new-assessment-of-memantine.2216.html

[3] https://www.iqwig.de/en/press/press-releases/press-releases/galantamine-and-rivastigmine-patches-positive-influence-on-cognition-possible.2461.html

[4] https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0034004/

[5] https://www.ispor.org/News/articles/Oct06/german_policy.asp

[6] https://english.zorginstituutnederland.nl/publications/reports/2013/06/24/donepezil-hydrochloride-aspen-for-the-indication-symptomatic-treatment-of-mild-to-moderately-severe-alzheimer%E2%80%99s-dementia

For each HTA, information was extracted from publicly available documentation relating to the HTA using a range of categories that were derived from headings used in analysis of data from the literature review and from an initial analysis of the information. The categories were as follows:

Outcomes included:

- Outcomes set out during scoping
- Outcomes considered during review
- Outcomes considered differently as a result of stakeholder involvement (Suggestions made by stakeholders in regards to outcomes)
- Outcomes identified as relevant for future research
- Outcomes used in studies that supported the recommendation
- Outcomes considered in cost-effectiveness analysis

Challenges around including outcomes:

- Types of evidence considered
- Surrogate outcomes and methods of validation
- Patient preferences and patient-reported outcomes
- Aspects of meaningful delay and disease progression
- Influence of data uncertainty on outcomes
- Thresholds in regards to outcomes measures or cost-effectiveness

Supplement 3: Details of studies identified in literature review

Study ID (Relevance)	Purpose	Setting	Method	Data sources	Details	Results	Limitations
England (N=13)							
3.1 Allen et al 2017 (Low)	To compare initial Canadian national HTA recommendations with the initial decisions of the other HTA agencies, and to identify factors for differing national HTA recommendations between the four HTA agencies.	Australia, Canada, England, Scotland	Medicines that were reviewed by all four agencies and received a negative recommendation from only one agency were selected as case studies. Statistical analysis of HTA recommendations classified as positive or negative (numerically coded); percentage agreement was calculated	Information from websites of HTA and bodies responsible for final reimbursement decision	Process map using a previously developed mapping methodology; this enabled identification and relationship between HTA agencies and responsible body for reimbursement decision	HTA bodies considered clinical efficacy; adverse effects; cost-effectiveness; all have implicit or explicit quality-adjusted life-year threshold; factors influencing decisions were: uncertainties surrounding a range of factors including: cost-effectiveness; comparator choice; clinical benefit; safety; trial design; submission timing	Use of publicly available sources; inclusion criteria limited to products listed on Controlled Drug Regulation, which resulted in exclusion of cancer medicines
3.2 Carroll et al 2017 (Medium)	To explore the type of additional exploratory analyses conducted by Evidence Review Groups and their impact on the recommendations made by NICE	England	A content analysis of relevant documents was undertaken to identify and extract relevant data, and narrative synthesis was used to rationalize and present these data.	100 most recently completed single technology appraisals since 2009 with published guidance were selected for inclusion	Categories for exploratory analyses developed with research team; this was used to inform coding; all data extraction were double checked by two researchers	The additional analyses undertaken by Evidence Review Groups in the appraisal of company submissions are highly influential in the policy-making and decision-making process; clear influence on 47% of final appraisal determinations	No limitations stated by author(s)
3.3 Cerri et al 2013 (High)	This study examined the impact of evidence, process and context factors on NICE decisions; to assess which of factors best explains the pattern of NICE decisions	England	With multinomial logistic regression, the relative contribution of explanatory variables on NICE decisions was assessed	A data set of NICE decisions 2004-2009 in HTAs was created, including 32 variables extracted from published information. A three-category decision outcome variable was created	A total of 65 technology appraisals (118 technologies) were analysed	Results showed significant associations ($p<0.10$) between NICE decision outcome and four variables: (i) demonstration of statistical superiority of the primary endpoint in clinical trials by the appraised technology; (ii) the incremental cost-effectiveness ratio (ICER); (iii) the number of pharmaceuticals appraised within the same appraisal; and (iv) the appraisal year.	No limitations stated by author(s)

3.4 Clement et al 2009 (High)	To assess how committees use evidence on effectiveness and cost-effectiveness (including any barrier to such use) and what additional factors have influenced decisions	Australia, Canada, England	Descriptive analysis of retrospective data from HTA bodies; 3 case studies: diabetes mellitus, ranibizumab for age-related macular degeneration, and teriparatide for osteoporosis	All publicly available documents as of 31st December 2008	Primary endpoint used in the supportive clinical studies and categorised end points as clinical endpoints, or surrogate endpoints; for surrogate end points, authors determined whether the committee felt the surrogate was a valid predictor of changes in the relevant clinical end point	Factors that influenced decisions: The differences in listing decisions often appeared less about the interpretation of the clinical or economic evidence and more about differences in agency processes In terms of outcomes: More than 50% of submissions reviewed by NICE used clinical end points (rather than clinical scales or surrogates), and if surrogate outcome were used they were more likely to be judged valid by committee	Use of publicly available sources; there may be subtle issues that were not captured, particularly in the deliberation process; surprisingly few common drugs across the 3 systems, making comparisons across committees less conclusive
3.5 Dakin et al 2014 (Medium)	To investigate the influence of cost-effectiveness and other factors on NICE decisions and whether NICE's decision-making has changed over time	England	Logistic regression to predict whether a technology was recommended or not; NICE's decisions as binary choices for/ against a technology in a specific patient group	Data on all NICE decisions published by December 2011 were obtained from HTAinSite [www.htainsite.com].	Independent variables comprised of the following: clinical and economic evidence; characteristics of patients, disease or treatment; and contextual factors potentially affecting decision-making.	Cost-effectiveness was main driver for NICE decisions; past decisions appear to have been based on a higher threshold than £20 000–£30 000/QALY; this may reflect consideration of other factors that cannot be easily quantified.	No limitations stated by author(s)
3.6 Drummond & Sorenson 2009 (Medium)	Opinion paper that explains NICE activities, achievements, challenges and lessons learnt	England	Opinion paper	-	-	No direct conclusions; issues discussed around QALY, ICER, sub group analysis and stakeholder involvement	N/A
3.7 Drummond et al 2013 (Medium)	Opinion paper that explores HTA approaches, in both methods and policy, to help bring about reconciliation between	Europe	Opinion paper	-	-	HTA initiatives are likely to give manufacturers an incentive to more closely align their research and development with social objectives; adequate stakeholder involvement is needed to ensure	N/A

	different parties and focus on social values and patient perspective					that the values incorporated in HTA processes adequately encompass social values
3.8 Fischer 2012 (High)	To structure empirical evidence of coverage decisions made in practice based on the components 'methods and evidence', 'criteria and standards', 'decision outcome' and 'processes'	Focus on England, scope international	Literature review	Electronic databases, journals and HTA websites were searched for publications between 1993 and June 2011. Included were analysis of past decisions and application of quantitative methods.	Each study was categorized by the scope of decision-making and the components covered by the variables used in quantitative analysis.	Important influence of therapeutic value where decision makers did not explicitly account for cost-effectiveness; the ICER had significant influence on decisions in Canada, Australia and the UK, but usually in combination with other aspects such as burden of disease or health condition. Budget considerations were significant influences in Australian and Dutch decision-making.	No limitations stated by author(s)
3.9 Kreis and Schmidt 2013 (Low)	This article explores operational processes and underlying rationales of public engagement at HTA agencies	France, Germany, United Kingdom	Authors explored qualitatively public engagement processes and underlying rationales	The analysis is based on website information, legal framework documents, published and grey literature, and semi structured, in-depth interviews with top officials at these agencies	Authors used the term public as the broadest generic term to include engagement of individual citizens, patients, consumers (or users), laypeople, or formal or informal representatives of groups of these	Engagement processes differed across agencies, particularly regarding the areas in which the public is involved, which groups of the public are involved, what weight they have in influencing decisions, how they are recruited and supported, and how potential conflicts of interests are addressed.	No limitations stated by author(s)
3.10 Nicod and Kanavos 2012 (Medium)	To identify diverging HTA recommendations across five countries, understand the rationale for decision-making, and suggest ways forward to minimize inter-country differences	England, Scotland, Sweden, Canada, and Australia	Comparative statistical analysis of HTA recommendations for 287 drug-indication pairs appraised by countries between 2007 and 2009, including an in-depth analysis of two case studies	Appraisal reports from each agency	Agreement levels were measured using kappa scores. Associations between the HTA recommendations and the HTA body issuing the recommendation were explored	Substantial disparities in recommendations for/ against drugs; HTA processes potentially influenced by: different priorities in different settings; different perception of benefit and value, and use of different tools of addressing uncertainty; patient preferences and characteristics	No limitations stated by author(s)

					through correspondence analysis	seem to weigh more heavily in certain disease areas than other	
3.11 Nicod et al 2017 (High)	To better understand the reasons for differences in reimbursement decisions for orphan drugs in four European countries	England, Scotland, Sweden, France	Semi structured interviews with representatives of HTA bodies	Semi-structured interviews; eight representatives from the four HTA bodies were interviewed between March and June 2015	An interview topic guide was developed on the basis of findings from a systematic comparison of HTA decisions for 10 orphan drugs. Qualitative thematic data analysis using the framework approach	Decisions regarding orphan drugs made in context of lower quality evidence; threshold of acceptable uncertainty varied by country; NICE more likely to accept surrogate endpoints for orphan drugs; NICE always prefers overall survival to progression-free survival; HRQOL data were considered as a hard end point by NICE. Safety only implicitly considered because already part of marketing authorisation.	No limitations stated by author(s)
3.12 Oyebode et al 2016 (Low)	To determine the aspects of expert advice that decision-makers find most useful in the development of evidence-based guidance and to identify the characteristics of experts providing the most useful advice	England	(1) Interviews examined the usefulness of expert advice during guidance development. (2) Associations between usefulness score and characteristics of the expert advisor were investigated using univariate and multivariate analyses	(1) Semi-structured interviews with 17 members of the Interventional Procedures Advisory Committee of NICE. (2) Data were extracted from 211 experts' questionnaires for 41 consecutive procedures.	(1) Transcripts were analysed inductively to identify themes; (2) Usefulness of advice was scored using an index developed through the qualitative work.	Values and challenges of using expert opinion in HTA processes are analysed	Authors reflect on their own potential bias due the researchers' previous experience at NICE and working in public health and medical roles; concept of 'usefulness' was potentially problematic
3.13 Spinner et al 2013 (Medium)	To assess whether different clinical evidence bases may have influenced listing recommendations	Australia, Canada, England and Wales	Authors reviewed the evidence considered for each listing recommendation, identified the similarities and differences, and evaluated the extent to which different clinical	Appraisal reports between 2007 and 2010 (including manufacturers' submissions) for nine drugs for which the three agencies had	Not provided	Decisions across HTA bodies associated with differences in the clinical evidence base considered. NICE considered indirect and/or mixed-treatment comparisons; in some cases, NICE excluded trials from review if the drug and/or the comparator	Small number of case studies; only publicly available documents were considered

			evidence bases could have contributed to different decisions	provided listing recommendations		were not administered according to the relevant marketing authorization.
Germany (n=14)						
3.14 Blome et al 2017 (High)	To determine methodological requirements for QoL measurement and data presentation in early benefit assessment (EBA)	Germany	Qualitative content analysis based on documents of all EBAs completed by 2014; analysis included information extraction, coding, critical discussion an consensus building	Documents publicly available on the G-BA website including: manufacturer dossier; dossier evaluation and benefit assessment by IQWiG or Federal Joint Committee (G-BA); protocol of the oral hearing; rationale of the G-BA decision ("Tragende Gruende"=main justifications)	Documents were searched for the term QoL; Relevant passages of all EBAs of 2011–2013 were independently extracted and reduced to key content by two researchers. Recurring patterns were identified and verified through comparison with EBAs of 2014.	No association between the inclusion of QoL data in benefit dossiers and the G-BA's rating decision might be explained by non-compliance with the various methodological requirements found in our analysis, so that in most cases, the mere inclusion of QoL data in the dossier did not lead to a positive evaluation of QoL benefit. In addition, many EBAs did include QoL outcomes, but there were no statistically or clinically significant effects	No limitations stated by author(s)
3.15 Fischer and Stargardt 2014 (Medium)	To explain the decisions made in early benefit assessments (EBAs), clarify the roles of manufacturers, IQWiG, German Federal Joint Committee (G-BA), and guide manufacturers in developing future submissions	Germany	Authors evaluated differences in rating decisions by manufacturers, the IQWiG, and the G-BA with regard to each pharmaceutical's added benefit. Authors used Cohen's kappa to analyze agreement between rating decisions; chi-square test and bivariate regression were used to identify associations between components of the EBA	Data extracted for EBAs for which the G-BA made a rating decision between 2011 and 2013. Authors developed a variable list including: rating decisions of manufacturers, IQWiG, G-BA; characteristics of the process; types of evidence submitted; methods used to	Two independent reviewers extracted data. Once completed, the worksheets were compared to identify any deviations. Interrater reliability was good, with an average Cohen's kappa coefficient of 0.63 (range, 0.28 to 1.00) for categorical variables and an average Pearson's correlation coefficient of 0.80	While the G-BA tended to disagree with the rating of benefit by manufacturers, it softened IQWiG's decisions, potentially to make the final outcome more acceptable. Concerns voiced that the G-BA might be exceeding its statutory authority by taking cost or procedural considerations into account appear to be unfounded. Choosing appropriate evidence to submit for each endpoint remained a challenge, as submission of	No limitations stated by author(s)

			process and the rating decisions of the G-BA	generate evidence; and pharmaceutical's maximum possible budget impact.	for continuous variables (range, – 0.18 to 1.00). Any disagreement was resolved through discussion between the authors.	health outcomes evidently influenced decisions.	
3.16 Fischer et al 2016 (Medium)	To analyse whether decisions of the German regulatory agency (G-BA) deviate from decisions from HTA or regulatory agencies in England (NICE), Scotland (SMC) and Australia (PBAC).	Focus on Germany, compared with England, Scotland, Australia	Authors analysed statistically decisions made for comparable patient subgroups by the four agencies between 2011 and 2014. First, decisions were compared (a) by their final outcome, i.e. whether a health benefit was identified, and (b) by the agencies' judgement on comparative effectiveness. Subsequently, they partially explored reasons for differences between HTA agencies.	All early benefit assessments of G-BA completed between January 2011 and December 2014; for G-BA decisions and information on the corresponding EBAs, the database developed by the Hamburg Center for Health Economics (HCHE) was used; otherwise documents available from HTA websites	For each comparison, authors analysed the agreement between G-BA and each of the other HTA agencies. Agreement was quantified by calculating Cohen's Kappa, to determine whether agreement between two raters was by chance	G-BA deviated considerably in decisions compared to other HTA agencies; G-BA tends to appraise stricter than NICE. HTA Agencies differed in accepting endpoints such as recognising the surrogate endpoint progression-free-survival. Another example is to prefer disease-specific mortality over over-all mortality as endpoint or vice versa. Other factors in which agencies were different: choice of comparator(s); differences in handling lack of evidence Agreement in endpoints between the agencies was highest for adverse events and quality of life followed by mortality; for morbidity, G-BA and the other agencies agreed least often	No limitations stated by author(s)
3.17 Ivandic et al 2014 (High)	To explore to which extent methodological requirements of HTA agencies differ between Germany and England	Germany, England	The following aspects were examined: guidance texts on methodology and information sources for the assessment; clinical study design and methodology; statistical analysis, quality of evidence base, extrapolation of results (modelling), and generalisability of study	Not stated; publicly available information on methods from legal and guidance documents from HTA websites	The findings are presented separately for the two HTA systems and thus may serve as stand-alone references. A concise, integrated comparison follows to highlight the main similarities and	Methodological requirements differed mainly in the acceptance of low-level evidence, surrogate endpoints, and data modeling. Some of the discrepancies may be explained, at least in part, by differences in the health care system and procedural aspects (e.g. timing of assessment).	No limitations stated by author(s)

			results; and categorisation of outcome		differences in the methodological requirements.		
3.18 Griffith and Griffith 2015 (Low)	Analysis of past decisions of German HTA to inform future submissions	Germany	All IQWiG decisions from January 2011 to May 2015 were assessed, and the effect of the clinical evidence base on the submission outcome was examined.	Completed single drug appraisals from Jan 2011 to May 2015	Recommendation ('added benefit' or 'no added benefit'), indication, rationale, and evidence base were extracted	Over half of drugs appraised by IQWiG since 2011 have been given 'no added benefit' status, and direct evidence against an appropriate comparator remains a priority for a favourable decision	No limitations stated by author(s)
3.19 Kohler et al 2015 (Medium)	To determine the information gain from AMNOG documents compared with non-AMNOG documents for methods and results of studies available at market entry of new drugs.	Germany	Authors assessed reporting quality for each study and each available document for eight methods and 11 results items For each document type they calculated the proportion of items with complete reporting for methods and results, for each item and overall, and compared the findings.	Dossier assessments conducted by IQWiG between 1 Jan 2011 and 28 Feb 2013; European public assessment reports, journal publications, and registry reports.	Not provided	Concludes that AMNOG documents provide a considerably higher proportion of complete information than European public assessment reports; this includes information on methods, results and patient relevant outcomes. The information gap was most striking when the drug was approved only in a certain subpopulation.	Small sample
3.20 Kvitkina et al 2014 (Low)	To describe the feasibility of the early benefit assessment on the basis of patient-relevant outcomes by systematically characterising the outcomes available in manufacturers' dossiers and comparing the companies' and IQWiG's evaluations regarding patient relevance and surrogate validity	Germany	Dossier assessments were used for data extraction; the outcomes available and the respective evaluations were extracted and compared. 12 out of 22 submitted dossiers contained sufficient data to assess outcomes; all 12 assessable dossiers provided data on patient-relevant outcomes.	Publicly available manufacturers' dossiers; published between October 2011 and June 2012	Not provided	Data on mortality and adverse events were available in almost all dossiers; data on morbidity and health-related quality of life available in 8 and 7 dossiers, respectively. Of a total of 214 outcomes extracted by IQWiG, 124 patient-relevant and 3 surrogate outcomes were included in IQWiG's assessment (companies: a total of 183 outcomes included, of which 172 were patient-relevant and 11 were surrogates outcomes partly deviated from each other.	No limitations stated by author(s)

3.21 Lauenroth and Stargardt 2017 (High)	To analyze how value is determined within the scope of the German Pharmaceutical Restructuring Act	Germany	Generalized linear model regression to analyze impact of added benefit on difference between negotiated prices and prices of comparators	All pharmaceuticals that had undergone assessment, appraisal, and price negotiations in Germany before June 30, 2016	Data were extracted from G-BA databases; added benefit was defined in various ways; in all models, they controlled for additional criteria such as size of patient population, European price levels, and whether the comparators were generic.	Authors conclude that price premiums were driven by health gain, the proportion of people benefitting from a pharmaceutical, European price levels, and whether the comparator was generic. QoL did not play a role in current decision making	No limitations stated by author(s)
3.22 Leverkus and Chuang-Stein 2016 (Medium)	To investigate requirements of benefit assessment with special attention on: choice of the comparator, patient relevant endpoints, subgroup analyses, extent of benefit, determination of net benefit, primary and secondary endpoints, and uncertainty of the additional benefit.	Germany	Authors state they contrast the approaches taken by the G-BA and IQWiG with those of the European Medicines Agency (EMA).	Authors referenced IQWiG's General Methods paper, German Social Code Book, and G-BA's Rules of Procedure.	For principles underlying regulatory decisions, they reference primarily the International Conference on Harmonization (ICH) E9 (Statistical Principles for ClinicalTrials, 1998) document.	Provides comprehensive overview and opinion on methodological requirements and issues in German HTA process, with particular focus on the role of outcomes and evidence types	No limitations stated by author(s)
3.23 Lohrberg et al 2016 (High)	To analyse how QoL is defined in early benefit assessment (EBA) and which role does it play	Germany	Qualitative analysis all benefit assessments completed by the end of 2013 were processed. Additionally, data on the decision outcomes were collected and analysed	Publicly available dossiers (summaries), dossier evaluations, protocols of the oral hearings, the final resolutions of the Federal Joint Committee (G-BA) and main	Documents were imported to software and searched for QoL terms; resulting paragraphs were reduced and summarized by two researchers; coding was performed on the basis of summaries	QoL has not been well defined in HTA processes and does not inform final decisions; they identified the absence or the inappropriate presentation of QoL data; at the same the stakeholders saw the value and importance of including QoL in EBA	No limitations stated by author(s)

			justifications completed by 2013 (n = 66)				
3.24 Riedel et al 2014 (Low)	To explain some fundamental concepts in Health Economic Evaluations (HEE) and how these concepts are adapted in different countries, notably in Germany	Germany focused, international	Bibliographic search to identify existing methods of health economic evaluation of new drugs used by HTA agencies in 11 countries and comparison with German HTA agency	Published literature	-	Although the core principles of HEE are very similar worldwide, there is a lack of harmonization. Overcoming the fourth hurdle (the reimbursement hurdle) is likely to be increasingly challenging for new drugs.	No limitations stated by author(s)
3.25 Ruof et al 2014 (a) (High)	To analyse the outcomes 18 months after introduction of the new AMNOG legislation on early benefits assessments (EBA)	Germany	All EBAs commenced prior to June 2012 were included and analysed (proportions were calculated; no statistical analysis was carried out)	The G-BA website (http://www.g-ba.de/ informationen/ nutzenbewertung) was used to obtain manufacturers' benefit dossiers, IQWiG assessments, and G-BA decisions	27 EBAs were analysed in regards to: additional benefit, appropriate comparative therapy (ACT), patient-relevant endpoints, and adverse events	Considerable variance was observed in additional benefit reported by manufacturers, IQWiG and G-BA. Areas of disagreement included comparator selection, definition of subgroups and patient-relevant endpoints, and classification and balancing of adverse events.	No limitations stated by author(s)
3.26 Ruof et al 2014 (b) (High)	To compare endpoints and related benefit categories used in marketing authorisation to those considered by G-BA in the field of oncology	Germany	Evaluation of early benefit assessments (EBAs) in oncology commencing prior to 31 December 2013	The Summary of Product Characteristics (SPC) for the respective marketing authorisations was derived from the website of the EMA.	Clinical trial endpoints that supported the marketing authorisation and the benefit assessment were derived from (i) the SPCs, (ii) manufacturers' value dossiers and (iii) the G-BA value decisions	Inconsistencies in acceptance of morbidity and QoL outcomes between G-BA and EMA; EMA accepted well established and clinically relevant morbidity endpoints (e.g. progression-free survival and response rate), which were mostly excluded by G-BA; final decisions by G-BA mostly driven by mortality outcomes	No limitations stated by author(s)
3.27 Staab et al 2016 (High)	To evaluate the acceptance of clinically	Germany	Medicines for oncological, metabolic and infectious diseases with EBAs	Manufacturer's dossiers, regulatory	Documents were analysed to determine patient	Inconsistencies were identified in patient relevance of morbidity-related PEPs as well as in	No limitations stated by author(s)

	acknowledged primary endpoints (PEPs) from regulatory trials in early benefit assessments (EBAs) conducted by the Federal Joint Committee (G-BA)		finalised before 25 January 2016 were evaluated.	assessments, G-BA appraisals and oral hearing minutes were reviewed, and PEPs	relevance of outcomes from G-BA perspective; acceptance of symptomatic vs. asymptomatic outcomes were also analysed	acceptance of asymptomatic endpoints by the G-BA	
Netherlands (N=6)						
3.28 Angelis et al 2017 (High)	To study the practices, processes and policies of value-assessment for new medicines across eight European countries and the role of HTA beyond economic evaluation and clinical benefit assessment	France, Germany, England, Sweden, Italy, Netherlands, Poland and Spain	A systematic (peer review and grey) literature review was conducted using an analytical framework examining: (1) 'Responsibilities and structure of HTA agencies'; (2) 'Evidence and evaluation criteria considered in HTAs'; (3) 'Methods and techniques applied in HTAs'; and (4) 'Outcomes and implementation of HTAs'	Two electronic databases (MEDLINE—through PubMed resource—and the Social Science Citation Index—through the Web of Science portal) were searched up to January 2014; with article searches taking place in February 2013 in the first instance and update taking place at the end of January 2014	Systematic literature review method based on the Centre for Reviews and Dissemination (CRD) guidance Feedback from the Advance-HTA consortium partners was provided in August 2014. Additional input, including the most recent updates on national HTA processes, was collected from HTA experts and national competent authorities between March and May 2016.	Debates about health utilities/ preferred health gain; for example, while NICE favours the use of the QALY, IQWiG strongly opposes its use on the grounds that it does not reflect patient-level utilities Increasing use of incorporating real world data; considerable subjectivity in the criteria selection used to interpret evidence and determine product value; increasing realisation by many HTA agencies that value is multi-dimension; move away from only relying on 'scientific value judgments' (safety/ efficacy/ effectiveness); need for methodological approaches that encompass multiple evaluation criteria explicitly.	No limitations stated by author(s)
3.29 Cerri et al 2014 (Medium)	To examine the factors that influence decisions made by the Dutch HTA agency (CVZ) to recommend, restrict or not recommend pharmaceutical	Netherlands	Descriptive statistics for each variable, stratified by outcome group (recommended, restricted or not recommended): chi-squared test for categorical variables; ANOVA test for continuous variables;	CVZ decisions in 2004–2009. A data set of CVZ decisions pertaining to pharmaceutical technologies was created, including 29 variables	Technologies included in list 1A/1B or on the expensive drug list considered recommended; those included in list	The multinomial model showed significant associations (p B 0.10) between CVZ outcome and several variables, including: (1) use of an active comparator and demonstration of statistical superiority of the primary endpoint in clinical trials, (2) pharmaceutical budget impact	Reliance on publicly available data sources; data extraction performed by single researcher (under supervision from senior researchers)

	technologies for use in the Netherlands		Kruskal-Wallis for not normally distributed indicators. A multinomial logit regression was used in the analysis to model the probabilities associated with the three types of technology appraisal outcome.	extracted from published information.	2 were considered restricted;	associated with introduction of the technology, (3) therapeutic indication and (4) prevalence of the target population. Results confirm the value of a comprehensive and multivariate approach to understanding CVZ decision-making.	
3.30 Franken et al 2013 (Medium)	To investigate the role of pharmacoeconomic evidence in drug reimbursement decision making; and (ii) to determine the extent to which appraising the importance of full economic evaluations relative to other evidence is a transparent process	Netherlands , Sweden	Authors investigated all reimbursement dossiers published in the period January 2005 to July 2011.	Data sources included all Dutch and Swedish drug reimbursement information published in the period January 2005 to July 2011	The analysis started in 2005 because that was the first year in which pharmacoeconomic evidence was required for reimbursement decision making in The Netherlands.	Therapeutic value appeared to be the most decisive criterion; the relative importance of full economic evaluations is more modest than would generally be expected, especially in The Netherlands; both countries could make the appraisal process more transparent by more explicitly showing the role of different criteria.	Reliance on publicly available data sources
3.31 Le Polain et al 2010 (Medium)	To describe and critically evaluate drug reimbursement decision processes, to identify their strengths and weaknesses and to formulate general policy recommendations.	Austria, Belgium, France, the Netherlands and Sweden	Comparative study (1) for the description of drug reimbursement decision processes, authors used the Hutton framework; (2) systems were evaluated using accountability for reasonableness framework by Daniels and Sabin.	Literature, policy documents and interviews with stakeholders	-	The paper provides a wide range of information on assessment and appraisal processes of Dutch HTA, and draws conclusions about criteria: For example, although there is no formal hierarchy in assessment criteria, most interviewees stated that effectiveness, efficacy and side effects were often the most important criteria determining the therapeutic value. Interviewees also acknowledged that the majority of time in a meeting of the Dutch HTA is devoted to determining the therapeutic value, less time is spent on	Analysis took place in supply-driven context; it was beyond the scope of this study to explore opportunities to move towards demand-driven system, where the societal needs drive the industry's strategic plan

						assessing cost-effectiveness evidence.	
3.32 Stolk et al 2009 (Medium)	To review the current approach to HTA used in The Netherlands in medical specialist care; the authors seek to provide a basic understanding of the strengths and weaknesses of the specific practices and processes	Netherlands	Opinion paper	-	-	Authors explore trends in future of (Dutch) HTA: What can be expected is a growing incentive for all parties to generate HTA data; increasing trend for conditional reimbursement linked to requirements for data collection and further study; further work is needed to understand how assessments and procedures jointly affect decision-making and to develop best practice guidelines; broader appraisals might be needed where the assessment will also cover optimal positioning of a service amongst the variety of services available to patients	N/A
3.33 Versteegh et al 2016 (Medium)	In this editorial, the authors highlight the distinguishing features of the new Dutch guidelines for economic evaluation; and highlight which developments, in their opinion, are desirable in coming updates, but are still in development or controversial	Netherlands	Editorial	-	-	New guidelines set preference for QALYs measured with the EQ-5D if appropriate but also offer alternative approaches for areas in which QoL might not be appropriate such as: prevention; diagnostics; medical devices; long-term care; forensics; reference is also made to multi-criteria decision making	N/A

References

Allen N, Walker SR, Liberti L, Salek S. Health Technology Assessment (HTA) Case Studies: Factors Influencing Divergent HTA Reimbursement Recommendations in Australia, Canada, England, and Scotland. *Value Health*. 2017;**20**:320-328.

Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. *Eur J Health Econ*. 2018;**19**:123-152.

Blome C, Augustin M, Metin H, Lohrberg D. Four years of early benefit assessment of new drugs in Germany: a qualitative study on methodological requirements for quality of life data. *Eur J Health Econ*. 2017;**18**:181-193.

Cerri KH, Knapp M, Fernandez JL. Public funding of pharmaceuticals in The Netherlands: investigating the effect of evidence, process and context on CVZ decision-making. *Eur J Health Econ*. 2014;**15**:681-695.

Cerri KH, Knapp M, Fernandez JL. Decision making by NICE: examining the influences of evidence, process and context. *Health Econ Policy Law*. 2014;**9**:119-141.

Carroll C, Kaltenthaler E, Hill-McManus D, Scope A, Holmes M, Rice S et al. The Type and Impact of Evidence Review Group Exploratory Analyses in the NICE Single Technology Appraisal Process. Value Health. 2017;**20**:785-791.

Clement FM, Harris A, Li J, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: A comparison of Britain, Australia, and Canada. *Jama*. 2009;**302**:1437-1443.

Dakin H, Devlin N, Feng Y, RiceN, O'Neill, Parkin D. The influence of cost-effectiveness and other factors on NICE decisions. *Health Econ*. 2015;**24**:1256-1271.

Drummond M, Sorenson C. Nasty or Nice? A Perspective on the Use of Health Technology Assessment in the United Kingdom. *Value Health*. 2009;**12**(2 Suppl):S8-S13.

Drummond M, Tarricone R, Torbica A. Assessing the added value of health technologies: reconciling different perspectives. *Value Health*. 2013;**16**(1 Suppl):S7-13.

Fischer KE, Heisser T, Stargardt T. Health benefit assessment of pharmaceuticals: An international comparison of decisions from Germany, England, Scotland and Australia. *Health Policy*. 2016;**120**:1115-1122.

Fischer KE, Stargardt T. Early benefit assessment of pharmaceuticals in Germany: manufacturers' expectations versus the Federal Joint Committee's decisions. *Med Decis Making*. 2014;**34**:1030-1047.

Franken M, Nilsson F, Sandmann F, de Boer A, Koopmanschap M. Unravelling drug reimbursement outcomes: a comparative study of the role of pharmacoeconomic evidence in Dutch and Swedish reimbursement decision making. *Pharmacoeconomics*. 2013;**31**:781-797.

Griffiths EA. The German NICE or the German nasty? An analysis of IQWIG decisions and requirements for an 'added benefit'. *Value Health*. 2015;**18**:A335.

Ivandic V. Requirements for benefit assessment in Germany and England - overview and comparison. *Health Econ Rev*. 2014;**4**:12.

Kohler M, Haag S, Biester K, et al. Information on new drugs at market entry: retrospective analysis of health technology assessment reports versus regulatory reports, journal publications, and registry reports. *BMJ*. 2015;**350**:h796.

Kreis J, Schmidt H. Public engagement in health technology assessment and coverage decisions: a study of experiences in France, Germany, and the United Kingdom. *J Health Polit Policy Law*. 2013;**38**:89-122.

Kvitkina T, ten Haaf A, Reken S, McGauran N, Wieseler B. [Patient-relevant outcomes and surrogates in the early benefit assessment of drugs: first experiences]. *Z Evid Fortbild Qual Gesundhwes*. 2014;**108**:528-538.

La Polain M, Franken MG, Koopmanschap M, Cleemput I. Drug reimbursement systems: international comparison and policy recommendations. Brussels: Belgian Health Care Knowledge Centre (KCE); 2010. Report No.: KCE Reports 147C.

Lauenroth VD, Stargardt T. Pharmaceutical Pricing in Germany: How Is Value Determined within the Scope of AMNOG? *Value Health*. 2017;**20**:927-935.

Leverkus F, Chuang-Stein C. Implementation of AMNOG: An industry perspective. *Biom J*. 2016;**58**:76-88.

Lohrberg D, Augustin M, Blome C. The definition and role of quality of life in Germany's early assessment of drug benefit: a qualitative approach. *Qual Life Res*. 2016;**25**:447-455.

Nicod E, Kanavos P. Commonalities and differences in HTA outcomes: A comparative analysis of five countries and implications for coverage decisions. *Health Policy*.**108**:167-177.

Nicod E. Why do health technology assessment coverage recommendations for the same drugs differ across settings? Applying a mixed methods framework to systematically compare orphan drug decisions in four European countries. *Eur J Health Econ*. 2017;**18**:715-730.

Oyebode O, Patrick H, Walker A, Campbell B, Powell J. The Ghost in the Machine? The Value of Expert Advice in the Production of Evidence-Based Guidance: A Mixed Methods Study of the Nice Interventional Procedures Programme. *Int J Technol Assess Health Care*. 2016;**32**:61-68.

Riedel R, Repschlager U, Griebenow R, et al. International standards for health economic evaluation with a focus on the German approach. *J Clin Pharm Ther*. 2013;**38**:277-285.

Ruof J, Knoerzer D, Dunne AA, et al. Analysis of endpoints used in marketing authorisations versus value assessments of oncology medicines in Germany. *Health Policy*. 2014;**118**:242-254.

Ruof J, Schwartz FW, Schulenburg JM, Dintsios CM. Early benefit assessment (EBA) in Germany: analysing decisions 18 months after introducing the new AMNOG legislation. *Eur J Health Econ*. 2014;**15**:577-589.

Spinner DS, Birt J, Walter JW, Bowman L, et al. Do different clinical evidence bases lead to discordant health-technology assessment decisions? An in-depth case series across three jurisdictions. *Clinicoecon Outcomes Res*. 2013;**5**:69-85.

Staab T, Isbary G, Amelung VE, Ruof J. Inconsistent approaches of the G-BA regarding acceptance of primary study endpoints as being relevant to patients - an analysis of three disease areas: oncological, metabolic, and infectious diseases. *BMC Health Serv Res*. 2016;**16**:651.

Stolk EA, de Bont A, van Halteren AR, Bijlmer RJ, Poley MJ. Role of health technology assessment in shaping the benefits package in The Netherlands. *Expert Rev Pharmacoecon Outcomes Res*. 2009;**9**:85-94.

Versteegh M, Knies S, Brouwer W. From Good to Better: New Dutch Guidelines for Economic Evaluations in Healthcare. *Pharmacoeconomics.* 2016;**34**:1071-1074.

Supplement 4: HTA documents analysed for case studies

National Institute for Health and care Excellence (NICE), England

<u>Draft documents for consultation</u>

4.1 Health Technology Appraisal Donepezil, galantamine, rivastigmine and memantine for the treatment of mild to moderate Alzheimer's disease (Part review of TA 111) Draft scope

4.2 Alzheimer's disease - donepezil, galantamine, rivastigmine and memantine (review): appraisal consultation document (online)

<u>Final documents</u>

4.3 Health Technology Appraisal Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (Review of TA 111) Final Scope

4.4 Final Appraisal Determination Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of NICE technology appraisal guidance 111)

<u>Reports by the Assessment group</u>

4.5 The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of TA111): a systematic review and economic model, Produced by: Peninsula Technology Assessment Group (PenTAG), University of Exeter [Note that this includes a revised section on results]

4.6 Overview Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease   (Review of NICE technology appraisal guidance 111)

<u>Comments to Technology Assessment Report (TAR)</u>

4.7 Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (Review of TA 111), Responses by various stakeholders including: Eisai/Pfizer; NHS Quality Improvement Scotland; NHS West Kent and NHS Islington; Novartis; Shire Pharmaceuticals; Alzheimer's Society; RICE (The Research Institute for the Care of Older People); Lundbeck

<u>Responses by Assessment Group</u>

4.8 NICE Health Technology Appraisal, Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of TA 111), Response to consultee and commentator comments on the draft remit and draft scope

4.9 NICE, Health Technology Appraisal, Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of NICE technology appraisal guidance 111)   Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)

4.10 AChEIs and memantine for Alzheimer's Disease, PenTAG responses to Consultee comments 17th August 2010

<u>Submissions</u>

4.11 Various submissions including by manufacturers and other stakeholders e.g. Alzheimer's Society Report; British Geriatrics Society; Royal College of Psychiatrists (Faculty of old age psychiatry); NHS Quality Improvement Scotland

Institut für Qualität und Wirtschaflichkeit im Gesundheitswesen (IQWiG), Germany

<u>Cholinesterase Inhibitors: Donezepil, Galantamin, Rivastigmin</u>

4.12 Berichtsplan zum Bericht „Cholinesterasehemmer bei Alzheimer Demenz" , [Auftrag A05-19A], Version 1.0 Stand: 02. Juni 2005; Report plan. Last accessed 10th January 2018

4.13 Amendment 1 zum Berichtsplan „Cholinesterasehemmer bei Alzheimer Demenz", [Auftrag A05/19A] , 12.06.2006; Amendment 1 to Report Plan version 1.0. Last accessed 10th January 2018

4.14 IQWiG. Cholinesterasehemmer bei Alzheimer Demenz. Vorbericht A05/19-A. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG); September 2006. [Preliminary report] Last accessed 10th January 2018

4.15 IQWiG: Cholinesterase inhibitors in Alzheimer's disease. Final report A05-19A. Cologne: Institute for Quality and Efficiency in Health Care (IQWiG); February 2007. Last accessed 10th January 2018

4.16 IQWiG. Cholinesterasehemmer bei Alzheimer Demenz. Abschlussbericht A05-19A. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG); Februar 2007. Last accessed 10th January 2018 [→ this is the German version of 4.15; in addition to 4.15 it includes the documented stakeholder involvement through meeting and written consultation]

Memantine

4.17 Berichtsplan zum Bericht „Memantin bei Alzheimer Demenz" [Auftrag A05-19C] Version 1.0 Stand: 24. August 2005, Report plan version1. Last accessed 10th January 2018

4.18 Amendment 1 zum Berichtsplan „Memantin bei Alzheimer Demenz" [Auftrag A05/19C], 12.06.2006; Amendment 1 to the report plan version. Last accessed 10th January 2018

4.19 Amendment 2 zum Berichtsplan Memantin bei Alzheimer Demenz, Auftrag A05-19C Version 1.0 Stand: 06.08.2007; Amendment 2 to the report plan version. Last accessed 10th January 2018

4.20 Memantin bei Alzheimer Demenz, Dokumentation and Würdigung der Stellungnahmen zum Berichtsplan, Auftrag A05-19C Version 1.0 Stand: 11.02.2008 ; documentation and appraisal of comments on the report plan version 1.0. Last accessed 10th January 2018

4.21 Memantin bei Alzheimer Demenz, Berichtsplan, Auftrag A05-19C Version 2.0 Stand: 11.02.2008 ; Report plan version 2.0. Last accessed 10th January 2018

4.22 Memantin bei Alzheimer Demenz Vorbericht (vorläufige Nutzenbewertung), Auftrag A05-19C Version 1.0 Stand: 01.08.2008 ; Preliminary report. Last accessed 10th January 2018

4.23 Memantin bei Alzheimer Demenz, Dokumentation und Würdigung der Stellungnahmen zum Vorbericht, Auftrag A05-19C Version 1.0 Stand: 28.04.2009.; Documentation and appraisal of comments on the preliminary report. Last accessed 10th January 2018

4.24 IQWiG-Berichte – Jahr: 2009 Nr. 59 Memantin bei Alzheimer Demenz, Abschlussbericht, Auftrag A05-19C Version 1.0 Stand: 08.07.2009 Final report. Last accessed 10th January 2018

4.25 Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Einleitung eines Stellungnahmeverfahrens zur Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage III – Übersicht der Verordnungseinschränkungen und –ausschlüsse Memantin, Vom 10. August 2010. Last accessed 10th January 2018

Zorginstituut Nederland, previously: College voor Zorgverzekeringen (CVZ), Netherlands

4.26 CFH rapport 07/11 memantine (Ebixa®), (2e)herbeoordeling, Op 2 april 2007 uitgebracht aan de minister van Volksgezondheid, Welzijn en Sport

4.27 GVS-rapport 13/11 donepezil (hydrochloride) Aspen® Vastgesteld op 24 juni 2013, College voor zorgverzekeringen, Diemen.

4.28 Farmacotherapeutisch rapport rivastigmine (Exelon®) bij Parkinsondementie, 2006

Supplement 5: List of stakeholders involved in HTAs as identified in case studies

ENGLAND	
Stakeholder group: Manufacturers	
Accord Healthcare (donepezil)	Novartis (rivastigmine)
Aspire Pharma (galantamine, rivastigmine)	Pfizer (donepezil)
Actavis UK (all four drugs)	Ranbaxy (donepezil)
Consilient Healthcare (galantamine, memantine)	Sandoz (all four drugs)
Dr Reddy's Laboratories (all but galatamine)	Shire (galantamine)
Eisai (donepezil)	Teva UK (all four drugs)
Lundbeck Ltd (memantine)	Wockhard UK (donezepil)
Mylan (galantamine, memantine)	Zentiva UK (all but rivastigmine)
Stakeholder group: Patient/ carer groups	

Afiya trust	Mental Health Foundation
Alzheimer's Society	Muslim Council of Britain
Carers UK	Muslim Health Network
Disability Rights UK	Neurological Alliance
Equalities National Council	Neurosupport
Innovations in Dementia	South Asian Health Foundation
Leonard Cheshire Disability	Specialised Healthcare Alliance
Stakeholder group: Professional associations	
Association of British Neurologists	Royal College of General Practitioners
Association of Directors of Adult Social Services	Royal College of Nursing
British Geriatrics Society	Royal College of Pathologists
British Neuropathological Society	Royal College of Physicians
British Neuropsychiatry Association	Royal College of Psychiatrists
College of mental health Pharmacy	Royal Pharmaceutical Society
Dementia Action Alliance	Royal Society of Medicine
Institute of Neurology	United Kingdom Clinical Pharmacy Association
Primary Care Neurology Society	
Others	
Department of Health	NHS South Eastern Hampshire CCG
NHS England	Welsh Government
NHS Somerset CCG	
GERMANY	
Stakeholder	**English translation or description**
Technology Assessment for Memantine	
Bundesverband für Gesundheitsinformation und Verbraucherschutz e. V.	Association for health information for the public and consumer protection (charity)
Deutsche Alzheimer Gesellschaft e.V	German charity for Alzheimer
Deutsche Gesellschaft für Neurologie; Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde	Professional association for psychiatry, psychotherapy and neurology
Hirnliga e.V.	Charity for the brain; refers to Alzheimer
Institut für angewandte Statistik	Institute for applied statistics
Institut für Arzneimittelsicherheit in der Psychiatrie	Institute for safety of psychiatric drugs
Karolinska Institutet	Swedish medical university
Kompetenznetz Demenz	Network for researchers, clinicians, people living with Alzheimer and their families
Lundbeck GmbH	Pharma company
Merz Pharmaceuticals GmbH	Pharma company
Novartis Pharma GmbH	Pharma company
The Research Institute for the Care of Older People (RICE)	/
Verband Forschender Arzneimittelhersteller e. V. (VFA)	Association of pharma companies involved in research

Verein zur Förderung der Forschung auf dem Gebiet der experimentellen Neurologie	Association to promote research in neurology
Technology Assessment for Cholinesterase inhibitors	
Eisai GmbH	Pharma
Novartis GmbH	Pharma
Pfizer GmbH	Pharma
Janssen-Cilag GmbH	Pharma
Merz Pharmaceuticals GmbH	Pharma
Verband Forschender Arzneimittelhersteller e.V.	Association of Pharma Companies involved in Research
University of Manchester	University, England (UK)
Alzheimer-Ethik e.V.	Charity for Alzheimer, founded by carers
Universitätsklinikum Freiburg	University
Deutsche Gesellschaft für Gerontologie und Geriatrie	German Society of Gerontology and Geriatrics
Arznei-Telegramm	News magazine about drugs
Deutsche Gesellschaft f. Gerontopsychiatrie und –psychotherapie (DGGPP) e. V.	German Psychogeriatric Association
Deutsche Alzheimer Gesellschaft e. V.	German Alzheimer Association (charity)
Universitätsklinikum Hamburg-Eppendorf	Medical University in Hamburg, Germany
Kompetenznetz Demenzen	Network for dementia
Hirnliga e.V.	Charity for the Brain, specifically Dementia
Bezirkskrankenhaus Günzburg	Hospital
Institut für Klinische Pharmakologie, Klinikum Bremen-Mitte	Pharmacological Institute, Medical university

Table 1: Features of health technology assessment (HTA) agencies in England, Germany, and

The Netherlands

COUNTRY HTA Agency's name	ENGLAND National Institute for Health and Care Excellence (NICE)	GERMANY Institute for Quality and Efficiency in Health Care (IQWiG)	NETHERLANDS Zorginstituut Netherlands (ZIN)
Roles & responsibilities	Advice about which drugs get funded based on cost-effectiveness (=fourth hurdle)	Advice on added value used to inform price negotiations	Advice informs whether drug gets on positive list (which influences price)
Value assessment	Cost-effectiveness	Clinical benefit	Clinical benefit; cost-effectiveness for drugs above certain threshold
Enforcement of methods	Guidance (manual) and reference case	Legislation and guidance	Guidance and reference case
Stakeholder involvement	Public consultations, Committee of experts (including patient representatives) co-develops guideline	Consultation with experts (including patient representatives)	Consultation with experts (including patient representatives)

Table 2: Clinical scales used in studies identified in health technology assessments (HTAs) of Alzheimer's disease drugs in England, Germany and The Netherlands

Drug covered in HTAs	**Memantine**	**Donezepil**	**Galantamine**	**Rivastigmine**	**Different cholinesterase inhibitors**
England					
Cognition	SIB, ADAS-cog	ADAS-cog, MMSE, SIB scales	ADAS-cog	ADAS-cog, MMSE	MMSE, SIB
Functioning	ADCS-ADL, FAST	ADL	ADCS-ADL, DAD, GAS	PDS, ADCS-ADL, ADL	ADL
Behavioural functioning and psychological symptoms	NPI, NPI nursing home version, BGP	NPI, in particular aggression and agitation	NPI	NPI, Hamilton DS	NPI
Global assessment	CIBIC-plus	CDR, CDR-SB, CIBIC-plus, GBS	CIBIC-plus	CIBIC-plus, GDS, ADCS-CGIC	GDS
Germany					
Cognition	SIB, MMSE, BGP-C	ADAS-cog, MMSE, SIP (for people with Down syndrome	ADAS-cog	ADAS-cog	ADAS-cog, SIB, MMSE
Functioning	ADCS-ADL-sev, BGP, BGP-D	IDDD, PSMS-plus, IADL-plus, DAD, CMCS, PDS	DAD, ADCS-ADL, GAS	PDS	BADLS, ADCS-AL, PMS, Blessed-Roth Dementia Scale
Behavioural functioning and psychological symptoms	NPI	NPI, NPI nursing home	NPI	NOSGER, BEHAVE-AD	NPI, NPI-D, BEHAVE-AD
Global assessment (information provided in supplement)	CIBIC-plus, ADCS-CGIC, GDS, FAST	CIBIC-plus, J-CGIC, GBS, CDR, CDR-SB	CIBIC-plus	CIBIC-plus	GDS
The Netherlands					
	Memantine	**Donezepil vs. Rivastigmine**	**Donezepil vs. Galantamine**	**Rivastigmine for Parkinson's dementia**	
Cognition	ADAS-cog, SPI	SIB, MMSE	ADAS-cog, MMSE	ADAS-cog, ADCS-CGIC MMSE	

Functioning	ADCS-ADL	ADCS-ADL	BrADL	ADCS-ADL
Behavioural functioning and psychological symptoms	NPI	NPI	NPI	NPI-10
Global Assessment	CIBIC-PLUS	GDS	/	/

Index of abbreviations used in Table: Activities of daily living (ADL); Alzheimer's Disease Assessment Scale – cognitive subscale (ADAS-cog); Alzheimer's Disease Cooperative Study – Activities of Daily Living (ADCS-ADL); the mild cognitive impairment ADL scale (ADCS-MCIADL); Allocation of caregiver time burden (ACTS); Behavioural Rating Scale for Geriatric Patients (BGP); Behavioral Pathology in Alzheimer's Disease Rating Scale (BEHAVE-AD) and the Behavioral Rating Scale for Dementias (BRSD); Behavioral Rating Scale for Geriatric Patients – Cognitive Subscale (BGP-C); Bristol Activities of daily Living Scale (BrADL); Clinical dementia rating (CDR); CDR sum of boxes (CDR-SB); Clinician's Interview-based Impression of Change (CIBIC)-plus; Caregiver Activity Survey (CAS); Caregiving burden scale (CBS); Caregiver stress Scale (CSS); Caregiver-rated Modified Crichton Scale (CMCS); Disability Assessment for Dementia (DAD); Functional Assessment Staging (FAST); Goal Attainment Scale (GAS); Global Deterioration Scale (GDS); Gottfried, Brine and Steen scale (GBS); Interview for Deterioration in Daily Living Activities in Dementia (IDDD), subscale; Instrumental Activities of Daily Living (IADL)-plus; Japanese-Clinical Global Impression of Change (J-CGIC); Mini Mental State Examination (MMSE), Neuropsychiatric Inventory (NPI); NOSGER (Nurses' Observation Scale for Geriatric Patients); Neuropsychiatric Inventory Caregiver Distress Scale (NPI-D); Neuropsychiatric inventory (NPI); Progressive Deterioration Scale (PDS); Severe Impairment Battery (SIB) scales; Physical Self-Maintenance Scale (PSMS)-plus; Severe Impairment Battery (SIB)

Table 3 Prioritization of outcomes and outcome measures in health technology assessments (HTAs) of Alzheimer's disease (AD) drugs

Outcome	ENGLAND	GERMANY	THE NETHERLANDS
Cost-effectiveness	**Prioritized** Economic modelling conducted	**Not prioritized** No economic modelling conducted	**Not prioritized** No economic modelling conducted
Quality of life	**Prioritized** Derived from outcomes measured with clinical scales to inform the economic modelling	**Not prioritized** Method requirements (e.g. 70% follow up rate) prevented influence of outcome on decisions	**Not prioritized** No assessment
Outcomes measured with clinical scales	**Prioritized** Cognition (including Mini-Mental State Examination; MMSE), Activities of daily living (ADL) and multi-domain change prioritized	**Prioritized** Global assessment outcomes less prioritized	**Prioritized** Mini-Mental State Examination (MMSE) less prioritized / not accepted
Adverse effects	**Not prioritized** Formally included but not referred to in decision	**Prioritized** Included and referred to in decision	**Prioritized** Included and referred to in decision
Outcomes relevant to people with Alzheimer's Disease	**Not prioritized** Focus on evidence from randomised controlled trials that measure outcomes with clinical scales	**Not prioritized** Focus on evidence from randomised controlled trials that measure outcomes with clinical scale	**Not prioritized** Focus on evidence from randomised controlled trials that measure outcomes with clinical scale
Carers' outcomes	**Prioritized** Included as primary endpoint but lower priority than outcomes for person with AD; methodological challenges prevent influence of outcome on decisions	**Not prioritized** Not considered responsibility of health care system; relevant only if impact on patient outcomes	**Not prioritized** Potential role in economic modelling but not carried out so far for AD drugs
Institutionalization	**Prioritized** Included in economic modelling (by linking outcomes with clinical scales to non-AD data sets)	**Not prioritized** Absence of trial evidence; method requirements also prevent influence of outcome on decisions	**Not prioritized** Not included due to absence of trial evidence