Edinburgh Research Explorer

Real-world evidence in Alzheimer's disease: the ROADMAP Data Cube

Real-world evidence in Alzheimer's disease: the ROADMAP Data Cube

Authors: Olin Janssen[a,*], Stephanie J. B. Vos[a], Glòria García-Negredo[b], Claire Tochel[c], Anders Gustavsson[d,e], Michael Smith[c], Amanda Ly[c,f], Mia Nelson[c], Helen Baldwin[c,g], Catherine Sudlow[c], Christin Bexelius[h], Christoph Jindra[g], Nemanja Vaci[g], Sarah Bauermeister[g], John Gallacher[g], Anna Ponjoan[i,j], Carole Dufoil[k,l], Josep Garre Olmo[j,m], Lars Pedersen[n], Ingmar Skoog[o,p], Antje Hottgenroth[q], Pieter Jelle Visser[a,r], Johan van der Lei[s], Carlos Diaz[b]

Author affiliations:

[a]*Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Alzheimer Center Limburg, Maastricht University, Maastricht, The Netherlands;*

[b]*Synapse Research Management Partners SL, Madrid, Spain;*

[c]*Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, UK;*

[d]*Quantify Research, Stockholm, Sweden;*

[e]*Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden;*

[f]*MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK;*

[g]*Department of Psychiatry, University of Oxford, Oxford, UK;*

[h]*F. Hoffmann-La Roche Ltd., Basel, Switzerland;*

[i]*Vascular Health Research Group (ISV-Girona). Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gorina (IDIAPJGol), Barcelona, Catalonia, Spain;*

jGirona Biomedical Research Institute (IDIBGI), Girona, Catalonia, Spain;

kCHU de Bordeaux, Pole de sante Publique, F-33000 Bordeaux, France;

lCentre Inserm U1219, Institut de Santé Publique, d'Epidémiologie et de Développement (ISPED), Bordeaux School of Public Health, Université de Bordeaux;

mDepartment of Medical Sciences, School of Medicine, University of Girona, Catalonia, Spain;

nDepartment of Clinical Medicine, Aarhus University, Aarhus N, Denmark;

oInstitute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg;

pDepartment of Old Age Psychiatry and Cognitive Disorders, Sahlgrenska University Hospital, Gothenburg, Sweden;

qLilly Deutschland GmbH, Bad Homburg, Germany;

rAlzheimer Centre and Department of Neurology, Amsterdam Neuroscience, VU University Medical Centre, Amsterdam, The Netherlands;

sDepartment of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands

Author names, postal addresses and email:

Olin Janssen, P.O. Box 616, 6200 MD, Maastricht, the Netherlands, olin.janssen@maastrichtuniversity.nl;
Stephanie J.B. Vos, P.O. Box 616, 6200 MD, Maastricht, the Netherlands, s.vos@maastrichtuniversity.nl;
Glòria García-Negredo, Velázquez 94, 1a planta, 28006 Madrid, Spain; ggarcia@synapse-managers.com;
Claire Tochel, Nine Edinburgh Bioquarter, 9 Little France Road, EH16 4UX, Edinburgh, UK, c.tochel@ed.ac.uk;
Anders Gustavsson, Hantverkargatan 8, 112 21 Stockholm, Sweden & Alfred Nobels allé 23, 141 83 Stockholm, Sweden, anders.gustavsson@quantifyresearch.com;
Michael Smith, Nine Edinburgh Bioquarter, 9 Little France Road, EH16 4UX, Edinburgh, UK, Michael.smith@ed.ac.uk;
Amanda Ly, Oakfield House, Oakfield Grove, BS8 2BN, Bristol, UK, Amanda.ly@bristol.ac.uk;
Mia Nelson, Nine Edinburgh Bioquarter, 9 Little France Road, EH16 4UX,

Edinburgh, UK, mia.nelson@ed.ac.uk;

Helen Baldwin, Warneford Hospital, Warneford Lane, OX3 7JX, Oxford, UK & Nine Edinburgh Bioquarter, 9 Little France Road, EH16 4UX, Edinburgh, UK, baldwinhelen22@gmail.com;

Catherine Sudlow, Nine Edinburgh Bioquarter, 9 Little France Road, EH16 4UX, Edinburgh, UK, cathie.sudlow@ed.ac.uk;

Christin Bexelius, Grenzacherstrasse 163, 4057 Basel Switzerland, christin.bexelius@roche.com;

Christoph Jindra, Warneford Hospital, Warneford Lane, OX3 7JX, Oxford, UK, Christoph.jindra@psych.ox.ac.uk;

Nemanja Vaci, Warneford Hospital, Warneford Lane, OX3 7JX, Oxford, UK, Nemanja.vaci@psych.ox.ac.uk;

Sarah Bauermeister, Warneford Hospital, Warneford Lane, OX3 7JX, Oxford, UK, sarah.bauermeister@psych.ox.ac.uk;

John Gallacher, Warneford Hospital, Warneford Lane, OX3 7JX, Oxford, UK, john.gallacher@psych.ox.ac.uk;

Anna Ponjoan, Unitat de Suport a la Recerca - IDIAPJGol, c/ Maluquer Salvador, 11 baixos, 17002 Girona, Catalonia, Spain, aponjoan@idiapjgol.info;

Carole Dufouil,146 rue Léo Saignat, 33076 Bordeaux cedex, France, carole.dufouil@u-bordeaux.fr;

Josep Garre Olmo, Edifici Mancomunitat 1, Parc Hospitalari Martí i Julià, C/ Dr. Castany s/n, 17190, Salt, Catalonia, Spain, josep.garre@ias.cat;

Lars Pedersen, Olof Palmes Allé 43-45, 8200 Aarhus N, Denmark, lap@clin.au.dk;

Ingmar Skoog, Department of Old Age Psychiatry and Cognitive Disorders, Wallinsgatan 6, 431 41 Mölndal, Sweden, Ingmar.skoog@neuro.gu.se;

Antje Hottgenroth, Werner-Reimers-Straße 2-4, 61352 Bad Homburg, Germany, hottgenroth_antje@lilly.com;

Pieter Jelle Visser, P.O. Box 616, 6200 MD, Maastricht, the Netherlands, pj.visser@maastrichtuniversity.nl;

Johan van der Lei, Erasmus MC Medische Informatica, Postbus 2040, 3000 CA Rotterdam, the Netherlands, j.vanderlei@erasmusmc.nl;

Carlos Diaz, Velázquez 94, 1a planta, 28006 Madrid, Spain; cdiaz@synapse-managers.com

Corresponding author:
Olin Janssen
Address: P.O. Box 616, 6200 MD, Maastricht, The Netherlands
Phone: +31 (0) 43 38 84098
E-mail: olin.janssen@maastrichtuniversity.nl

Word count: 3,979

Research in context

1. Systematic review: Articles containing outcomes of Alzheimer's disease (AD) relevant to different stakeholders across the disease spectrum were identified.

2. Interpretation: Availability of real-world outcomes in a variety of data sources across Europe identified through existing catalogs and previous projects, as well as relevance of the outcomes per disease stage, were represented in a Data Cube. Our results provide a first step toward understanding the European real-world AD data landscape in an integrated, visual overview of the available data sources across the disease spectrum.

3. Future directions: Better understanding of the AD real-world data landscape for the development of new AD treatments is needed. The list of AD outcomes considered relevant by all stakeholder groups should be further refined and validated to facilitate the construction of a real-world evidence platform that is both relevant and acceptable to the healthcare community.

Word count: 141 (max. 150 words)

Highlights

- Alzheimer's disease-relevant outcomes were collected for 66 real-world data sources <<85 characters>>

- The ROADMAP Data Cube visualizes available outcomes across different data sources <<83 characters>>

- Stakeholders prioritized quality of life and cognitive and functional abilities <<81 characters>>

- Most data sources captured cognitive ability, comorbidities and treatment outcomes <<84 characters>>

- The least captured outcomes were caregiver and family-related outcomes <<72 characters>>

Abstract

INTRODUCTION: The ROADMAP project aimed to provide an integrated overview of European real-world data on Alzheimer's disease (AD) across the disease spectrum.

METHODS: Metadata were identified from data sources in catalogs of European AD projects. Priority outcomes for different stakeholders were identified through systematic literature review, patient and public consultations, and stakeholder surveys.

RESULTS: Information about sixty-six data sources and thirteen outcome domains was integrated into a Data Cube. Gap analysis identified cognitive ability, functional ability/independence, behavioral/neuropsychiatric symptoms, treatment, comorbidities, and mortality as the outcomes collected most. Data were most lacking on caregiver-related outcomes. In general, electronic health records covered a broader, less detailed data spectrum than research cohorts.

DISCUSSION: This integrated real-world AD data overview provides an intuitive visual model that facilitates initial assessment and identification of gaps in relevant outcomes data to inform future prospective data collection and matching of data sources and outcomes against research protocols.

<<Word count = 149; maximum word count = 150>>

Abbreviations

AD   Alzheimer's disease

DPUK   Dementias Platform UK

EHR   Electronic health record

EMIF   European Medical Information Framework

HCP   Healthcare professional

IPCI   Integrated Primary Care Information database

MCI   Mild cognitive impairment

PPI   Patient and public involvement

ReDeGi   Registry of Dementias of Girona

ROADMAP   Real world Outcomes across the AD spectrum for better care: Multi-

modal data Access Platform

RWE   Real-world evidence

SIDIAP   Information System for Research in Primary Care

SLR   Systematic Literature Review

1. Introduction

There is a diverse landscape of real-world data on Alzheimer's disease (AD) in Europe. Such data are increasingly being used to provide evidence of the effectiveness of new AD treatments [1–3]. Use of real-world evidence (RWE) in research may help to improve the generalizability of findings by accounting for patient factors such as genetics, comorbidities, medication history, and treatment adherence [1,3].

Identification and characterization of the different European real-world data sources in AD provides a first step in understanding the European landscape of RWE and identifying data gaps. Several projects, including the European Medical Information Framework (EMIF) (http://www.emif.eu) [4,5] and the Dementias Platform United Kingdom (DPUK) project (https://www.dementiasplatform.uk), have attempted to identify and characterize cohort, electronic health record (EHR), and clinical trial data sources in AD. However, an integrated overview of the different data types available related to real-world outcomes is lacking.

The Real world Outcomes across the AD spectrum for better care: Multi-modal data Access Platform (ROADMAP) project was a 2-year public–private partnership to assess the feasibility of using AD RWE for the development of new AD disease models and treatments (http://roadmap-alzheimer.org) [6]. Among the key aims were the identification of European AD-relevant real-world data sources and AD-relevant outcomes considered a priority by different stakeholder groups (healthcare professionals [HCPs], patients, and caregivers). This article describes the integration of this information into a Data Cube to provide a three-dimensional overview of how different types of European data sources capture key AD outcomes and their relevance across the disease stages. Input from stakeholders was gathered to create an overview of outcome measures that are considered relevant by the different stakeholders. Thus, the Data Cube provides metadata, which are the characteristics of data sources

including the types of collected data, protocol features, and governance information, mapped against stakeholder preferences. The Data Cube itself does not contain the listed databases, neither partly or as a whole. Using the Data Cube, a gap analysis was performed to identify missing or partially reported outcomes in the characterized data sources.

2. Methods

2.1. Identification of data sources

Data sources within the ROADMAP Consortium (see Appendix A) and existing
metadata catalogs from the EMIF project and the DPUK initiative were used as a
starting point to identify European AD-relevant real-world data sources. Both catalogs
include detailed information about characteristics of the data sources and the different
disease-related outcomes that are assessed in each source. The EMIF online catalog
(https://emif-catalogue.eu) contains metadata on both cohort studies and EHR data
sources. Cohort studies typically collect observational disease-specific information from
a group of research participants or patients over a period of time for research
purposes, whereas EHR data sources typically contain registered information resulting
from healthcare processes. Information included in the EMIF-AD catalog was
completed either by data owners or by a member of the EMIF-AD team and
subsequently checked and completed by data owners. The DPUK catalog is based on
EMIF catalog ontology but is customized and accessed through the DPUK Data Portal
(https://portal.dementiasplatform.uk/). Further data sources were also identified through
the "EU dementia cohort mapping project," the results of data source mapping projects,
and knowledge directly provided by ROADMAP Consortium members (including data
from EHRs and clinical trial control groups).

2.2. Identification of relevant outcomes

Real-world AD outcomes of relevance to the different stakeholders were identified from
a systematic literature review (SLR), a series of patient and public involvement (PPI)
consultations, and a stakeholder survey. The SLR included relevant publications
regardless of the language the study was published in, the PPI consultations were
conducted in English, and the survey was undertaken in English. The literature review

identified studies that answered the following research question: "Which AD outcomes across the spectrum are prioritized by patients, caregivers, and HCPs?" Studies published between 2008 and 2017 inclusive were identified across a range of literature types. The whole spectrum of AD was considered, from preclinical AD to severe AD dementia [7].

The PPI consultations involved 29 people diagnosed with different forms of mild to moderate dementia, caregivers, and HCPs who identified key outcomes they considered important indicators of disease progression. People living with dementia (n=9, eight women and one man) and people supporting or providing care for people with dementia (n=7, five women and two men) were approached through the European Working Group of People with Dementia (EWGPWD) with the support of Alzheimer Europe. Carers were current family members including spouses and adult children who acted as facilitators, or carers supporting the people with dementia. HCPs (n=13) were individual professionals working in the field of AD and dementia or with an interest in AD and dementia with varying professional backgrounds (ethics, health economics, clinical medicine and science, advocacy, and industry). These HCPs were recruited through a snowballing approach that originated within the ROADMAP consortium and continued through wider professional networks of the individuals already involved. Nationalities of the different stakeholders varied and included nine European countries and the United States. Outcomes identified in the literature and selected using a consensus-based decision making approach were incorporated into surveys. These surveys gathered information on the relative importance of these outcomes to multiple stakeholder groups, including professionals working in dementia, people with dementia, and caregivers. A mixed-methods analytical approach [8,9] was used to compile and integrate outcomes identified from the systematic literature review, the PPI consultations, and the stakeholder surveys [10] (Table 1). Outcomes at the category level were the connecting point for all three methods. For example, cognitive abilities,

an outcome category, is made of many granular outcome sub-categories, such as memory, language & communication, executive functions, etc. All workstreams raised cognitive abilities as a priority outcome category, however at the granular sub-category level, differences and nuances arose across the workstreams (e.g. memory loss or language difficulties as sub-category). These outcome categories were developed through thematic analysis, consensus approaches, previous deliverables within the project, data synthesis meetings, and the use of clinical nosology. The survey was designed using these specific categories as response options for the questions, and the analysis of the PPI consultations and the SLR was undertaken using the categories as superordinate coding headings. Outcomes at the subcategory level were derived directly from the data and were allocated to a category through data synthesis meetings.

Outcome prioritization across the workstreams at the category and subcategory for individual stakeholder populations was undertaken for the data cube using an inclusive approach such that if it was prioritized by any of the workstreams for that stakeholder population it was included in the data cube filter for that stakeholder group. Data was not equally available across the workstreams for filtering the data cube by both stakeholder population and specific stages of disease progression. Stage specific prioritization for people with dementia and carers is derived from the PPI consultations and the SLR, whereas stage specific prioritization for professionals is derived from the SLR and the survey.

2.3.   Data integration into a Data Cube

The list of outcomes from the mixed-methods analytical approach was reviewed and refined by ROADMAP consortium members to develop a final consensus-based list consisting of main domain headings together with related outcomes. To improve user-friendliness and ease of completion of the Data Cube, a final selection of outcome

domains and outcomes to be included in the Data Cube was made. This selection process was supported by the expertise of different Consortium members. Outcomes excluded from the Data Cube were those indirectly captured by another outcome (e.g. "life as a whole" was excluded since this cannot be directly captured in most cohort studies, is not registered in EHR databases or clinical trial data, and is indirectly captured by other quality-of-life measures), or those that were difficult to identify and conceptualize and therefore unlikely to be directly captured in any of the data source types.

Information about the identified AD data sources, the prioritized AD outcomes with measurement tools/scales, and information on their relevance in each disease stage were integrated into a three-dimensional Data Cube (Figure 1). The disease stages considered were mild cognitive impairment (MCI), mild AD, and moderate-to-severe AD. After the outcomes of interest had been identified, a questionnaire was sent to custodians of the data sources to determine which of the outcomes were assessed in their data source and the measurement tools/scales used. For most data sources, the ROADMAP team could extract this information directly from the EMIF-AD or DPUK metadata catalogs without contacting the data custodians, since the information in these catalogs was provided or checked by the data custodians. Data extraction was completed by a ROADMAP Consortium member with expertise in the outcomes assessed in the different studies. In the questionnaire, data custodians or ROADMAP team members ranked the availability of outcomes in answer to the question: "Are the following AD outcomes reflected by any of the measures available in your database?" The presence or absence of each outcome for each data source was visualized in the form of a "heat map," which formed the first facet of the Data Cube (Figure 2).

The second facet of the Data Cube included information on the level of importance of each outcome for the different disease stages, derived from HCP ratings. The importance of each outcome for each disease stage was again visualized as a "heat

map." The third facet consisted of another "heat map" combining the data sources and the different stages of AD. This facet was used as an indicator of how well a data source collected information for each disease stage. The user interface for the Data Cube allows the user to zoom in on each facet of the rotating cube to view the two-dimensional heat maps in detail. The Data Cube can be accessed via https://datacube.roadmap-alzheimer.org.

2.4. Data gap analysis

Gaps in the availability of outcomes across the different data sources were evaluated by calculating the percentage of data sources containing information on at least one of the outcomes within each outcome domain. This was performed separately for the three different types of data source and for all data sources combined.

3. Results

3.1. Data sources

A total of 300 unique data sources in Europe were identified. Of these, questionnaire information on outcomes was available from 66 to complete the Data Cube. The 66 data sources consisted of four clinical trials (placebo data), four EHR databases (one each from Denmark and the Netherlands and two from Spain) and 58 cohort studies (Appendix B). Of the 58 cohort studies, 30 were identified from the EMIF-AD catalog, 25 from the DPUK catalog, and three from Consortium data owners (Registry of Dementias of Girona [ReDeGi; Spain], Memento [France], and the Gothenburg population study H70 [Sweden]).

3.2. Relevant outcomes and stakeholder priority ratings

A number of outcomes across all three stakeholder groups (HCPs, patients, and caregivers) were identified from the systematic literature review. A greater variety of outcomes pertaining to caregivers than to HCPs were identified, and there was a significant amount of overlap between outcomes relevant to patients and those relevant to caregivers (Figure 3).

The outcomes considered most important across all stakeholder groups were cognitive ability, functional ability and independence, and patient quality of life. By disease stage, the most important outcomes as rated by HCPs were cognitive ability, functional ability/independence, and patient quality of life for MCI; patient quality of life, functional ability/independence, cognitive ability, neuropsychiatric symptoms, and quality of caregiver's and family's lives for mild AD; and quality of caregiver's and family's lives, patient quality of life, neuropsychiatric symptoms, and behavioral symptoms for moderate-to-severe AD. For MCI, cognitive ability was considered very important by a large proportion of respondents, but the importance of this outcome domain decreased

as disease severity increased. Functional ability/independence showed a similar pattern. Results of medical investigations also declined in importance with disease severity. Conversely, behavioral and neuropsychiatric symptoms and quality of caregiver's and family's lives increased in importance with disease progression. Patient quality of life and significant disease-related life events received a large share of votes at all disease stages. There was little evidence addressing the opinions of patients and caregivers about important outcomes in the later stages of AD.

3.3. Data gap analysis

In general, most of the data sources had at least some information on the different outcome domains (Figure 2). When the three types of data source were combined, outcomes collected the most (>80% of the data sources) were cognitive ability, comorbidities, and therapeutic treatment. Outcomes collected the least (<30% of data sources) were functional and structural neuroimaging biomarkers, caregiver use of health and social care services, and quality of the caregiver's and family's lives. Among clinical trial placebo data sources, information on all outcomes was available except for quality of the caregiver's and family's lives. All EHR registries collected information on comorbidities, mortality, significant disease-related life events, treatment, patient use of health and social care, and behavioral and neuropsychiatric symptoms. Outcomes not collected by the EHR registries were cognitive ability, apolipoprotein E4 genotype, functional and structural neuroimaging biomarkers, physical and neurological examination, and caregiver use of health and social care services. Among the cohort studies, the outcomes most frequently (>75% of studies) collected were cognitive ability, comorbidities, treatment, and physical examination. Outcomes collected the least (<25% of studies) were quality of the caregiver's and family's lives, caregiver use of health and social care services, and functional and structural neuroimaging

biomarkers. The percentage of data sources with data available for the different

outcomes identified is presented in Appendix C.

4. Discussion

Overall, the outcome domains most commonly captured were cognitive ability, functional ability and independence, behavioral and neuropsychiatric symptoms, treatment, and comorbidities and mortality. Fewer data sources collected information on significant disease-related life events, medical investigations, use of health and social care, and patient quality of life. Data were most lacking on caregiver-oriented outcomes, particularly those related to the quality of caregiver's and family's lives.

4.1.   Data sources

Different data sources capture different types of information to varying degrees of richness and coverage, and no single data source type contains information on all of the outcomes across the full AD spectrum. Evidence generation should therefore cover several different types of data sources to provide a more complete picture of the data landscape.

In general, the four included EHR databases captured a broad spectrum of data and documented actual clinical care, but this information often lacked detail specifically for AD and was less suitable for investigation of the causes of AD. By contrast, cohort data were well structured and detailed, focusing on disease-relevant items; however, they covered a limited proportion of the AD patient population. Clinical trial placebo data covered most outcome domains, possibly because these studies are usually designed to cover outcomes considered important for evaluation of the efficacy and safety of a drug. In our work, which covered approximately 25% of over 300 identified data sources, the number of clinical trial placebo data sources and EHR databases was lower than the number of cohort studies. Mapping of broad EHR databases to the AD-outcomes questionnaire required a deep knowledge of the data in order to identify outcomes present in less obvious measures than the established scales and tests used

in research cohorts and randomized controlled trials (RCTs). This limited the number of mapped EHRs in our project. Nevertheless, we believe that our work provides a good reflection of the availability of outcomes in the different data sources.

4.2.   Outcomes

The outcomes considered most important across stakeholder groups were cognitive ability, functional ability and independence, and patient quality of life. These outcomes were assessed to some degree in the four clinical trial data sources included in our analysis, but mapping of additional clinical trial data sources is needed to verify this. In a recent review by Webster et al. [11], patient quality of life was measured in only 16 of 125 included clinical trials, and functional ability was measured in less than half of the trials. Data capture of these priority outcomes should therefore be encouraged. Several of the identified priority outcomes from a stakeholder perspective were qualitative in nature (e.g. emotional issues) and therefore differed from those captured in research studies and registry databases.

Although our work has produced a comprehensive list of AD-related outcomes, this is only the first step towards understanding these outcomes in the context of different stages of the disease for patients, caregivers, and HCPs. Future studies should consider the differences between stakeholder-reported priority outcomes and outcomes captured in previous studies when choosing tools for outcome assessment. Coordinated national and regional efforts may be needed to change the reporting of outcomes in routine clinical care to ensure that missing outcomes are documented in EHRs and registries, and international collaboration is needed to identify opportunities for harmonization.

4.3.   Data Cube and data gap analysis

The Data Cube was developed to create a visual, comprehensive overview of the European RWE landscape in AD and to provide a first step towards an integrated overview of the AD data environment. The evolution from a two-dimensional data source catalog to a third dimension addressing the relevance of outcomes across the disease spectrum should facilitate research studies that are more relevant to those impacted by AD.

By combining information from different data sources, we were able to identify gaps in the current collection of AD outcomes. The identification of data gaps that are relevant to the various stakeholders and further evaluation of the Data Cube heat maps form an important part of data landscape evaluation and should help to inform plans to develop supplementary data collection platforms and prospective research protocols to collect missing data in the future. Currently, data on outcomes relating to the caregiver and family are often lacking and should be assessed in future research. The analysis of outcomes available across the different data sources showed that most of the data sources had at least some information on the different outcomes. However, data sources differ in their suitability for answering specific research questions, and it is important to understand the limitations of each type of data source as described above in terms of level of detail and specificity. Since it is unlikely that a single data source will cover the full spectrum of AD, from pre-symptomatic disease to severe dementia, the combining of different data sources may be necessary.

4.4.   Strengths and limitations

An important strength of the work reported here was the starting base of an evidence-based outcomes list that incorporated multiple stakeholder perspectives. In addition, the Data Cube combined disparate information from a wide range of European data sources. A limitation of the Data Cube is that it does not include patient-level data, a feature that would bridge the gap between the current metadata overview and a future

RWE data platform. Another limitation is the reliance on indirect methods to reflect the availability of outcomes per disease stage, and further reiterations of the Data Cube should consider whether each data source actually records data from individuals in specific disease stages.

With regards to data sources, the main limitation was the under-representation of EHR registries and clinical trials. Thus, current data do not provide a representative sample of these data sources. However, the few EHR registries included in the Data Cube represent several European countries, and variation between different clinical trial data sources is often limited by the length of follow-up or high-cost outcomes. A limitation of capturing relevant outcomes was the lack of evidence from patients with severe AD. Gaining such evidence would be practically and ethically difficult, but it is a gap that should be addressed and could be partly bridged by the use of AD-specific registry data. Furthermore, although metadata included in the Data Cube were from several countries within Europe, the Data Cube might not be generalizable to the whole of Europe or countries outside of Europe. However, the Data Cube does provide a clear overview of the data outcomes that are generally available and accessible within Europe. Challenges of developing the Data Cube were ensuring a balance between compactness and sufficient level of detail, creating a user-friendly interface, and visualizing the Data Cube in a straightforward and concise manner. Further challenges were uncertainty over the accuracy of the information obtained from metadata catalogs and variability in responses to questionnaires, with some respondents possibly interpreting questions in a slightly different manner.

4.5. Future use of the Data Cube

Criteria for what constitutes an outcome or outcome measure will require clarification, and priority outcomes will need to be further investigated in terms of their validity, reliability, and ability to detect meaningful change in the progression of AD. Since

current outcome prioritization per disease stage was based on HCP ratings only and PPI consultations did not gather information on the MCI stage of AD, this will also need to be addressed. Considerations for further development and future use of the Data Cube include expansion of the number of data sources included, a greater balance between different data source types, more direct integration of disease-stage information, maintenance of the Data Cube so it remains current, consideration of how the Data Cube could be used in future initiatives, exploration of how the Data Cube could be combined with subject-level data into a RWE platform, and making the Data Cube publicly available. To realize further developments and sustainability, a Data Cube Monitoring Committee has been established. We plan to integrate the Data Cube with the existing European data sharing initiatives EMIF-AD and DPUK in order for researchers to obtain access to the individual-level data of data sources of interest to answer potential research questions. These data sharing initiatives have data governance and harmonization structures set in place. More information can be found on their respective websites (https://emifad.bmd-software.com; https://portal.dementiasplatform.uk).

The Data Cube provides an overview of the AD data landscape and can be used as a starting point to select appropriate data sources for different types of research questions. Questions that could be addressed with data sources included in the Data Cube are for example related to cognitive ability or comorbidities, whereas questions related to caregiver-oriented outcomes would less likely be able to address given the lower availability of the latter in the Data Cube data sources. With that pre-selection, researchers are encouraged to visit the EMIF-AD or DPUK Catalogs as well as engaging with the data source custodians to find out more detailed information and discuss access to the data. Moreover, the Data Cube could be used to identify gaps in availability of priority outcomes in different data sources in order to drive future data

collection.

4.6.   *Future RWE generation*

Future data-landscaping and RWE platform development should focus on enhancing global collaboration to create a common technical, legal, and ethical model for real-world research and evidence generation. International, national, and local efforts to harmonize database/cohort protocols should also be encouraged. The list of outcomes considered relevant by all stakeholders should be further refined and validated to facilitate data collection for a platform of European RWE that is both acceptable and relevant to the healthcare community and provides information about AD and appropriate intervention pathways. Data custodians and patients should be integrated into RWE research, not only as data contributors but also as research partners. In addition, measures should be implemented to inform them about research results and the immediate benefits. Finally, EHR databases and registries should be encouraged to include information relevant to specific research programs.

5. Conclusions

A diverse landscape of real-world data is available across Europe for research into AD. ROADMAP investigated this landscape and provided an integrated overview of the different AD data sources and priority outcomes from the perspectives of different stakeholder groups in relation to the different stages of AD, identifying data gaps with respect to relevant outcomes. This integrated data overview should enable the matching of data sources and relevant outcomes against research protocols, allow the development of research studies that are most relevant to those impacted by AD, and inform prospective data collection to address outcome evidence gaps.

Acknowledgements

Declaration of interest:

O. Janssen, S.J.B. Vos, G. Garcia-Negredo, M. Smith, A. Ly, M. Nelson, H. Baldwin, C. Sudlow, C. Jindra, N. Vaci, S. Bauermeister, J. Gallacher, A. Ponjoan, C. Dufouil, J. G. Olmo, L. Pedersen, P.J. Visser, J. van der Lei and C. Diaz reported no potential conflicts of interest. C. Tochel presented ROADMAP results at Alzheimer's Disease International conference 2018 for which flights were paid by Roche. A. Gustavsson is a partner of Quantify Research, providing consultancy services to pharmaceutical companies and other private and public organizations and institutions. AG contributes to ROADMAP on behalf of Roche Pharmaceuticals. C. Bexelius is an employee of F. Hoffmann-La Roche. I. Skoog is an advisor and speaker for Takeda. A. Hottgenroth is an employee of Lily Deutschland GmbH.

Role of the funding source

The ROADMAP project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 116020 ("ROADMAP"). This joint undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. The research described and preparation of this article were performed within this framework.

References

[1]   Eichler HG, Abadie E, Breckenridge A, Flamion B, Gustafsson LL, Leufkens H, et al. Bridging the efficacy–effectiveness gap: a regulator's perspective on addressing variability of drug response. Nat Rev Drug Discov 2011;10(7):495–506. https://doi.org/10.1038/nrd3501

[2]   Berger ML, Lipset C, Gutteridge A, Axelsen K, Subedi P, Madigan D. Optimizing the leveraging of real-world data to improve the development and use of medicines. Value Health 2015;18(1):127–30. https://doi.org/10.1016/j.jval.2014.10.009

[3]   Makady A, de Boer A, Hillege H, Klungel O, Goettsch W; on behalf of GetReal Work Package 1. What is real-world data? A review of definitions based on literature and stakeholder interviews. Value Health 2017;20(7):858–65. https://doi.org/10.1016/j.jval.2017.03.008

[4]   Bastião Silva LA, Días C, van der Lei J, Oliveira JL. Architecture to summarize patient-level data across borders and countries. Stud Health Technol Inform 2015;216:687–90. https://doi.org/10.3233/978-1-61499-564-7-687

[5]   Vaudano E, Vannieuwenhuyse B, Van Der Geyten S, van der Lei J, Visser PJ, Streffer J, et al. Boosting translational research on Alzheimer's disease in Europe: the Innovative Medicine Initiative AD research platform. Alzheimers Dement 2015;11(9):1121–2. https://doi.org/10.1016/j.jalz.2015.02.002

[6]   Gallacher J, de Reydet de Vulpillieres F, Amzal B, Angehrn Z, Bexelius C, Bintener C, et al. Challenges for optimizing real-world evidence in Alzheimer's disease: the ROADMAP project. J Alzheimers Dis 2019;67(2):495–501. https://doi.org/10.3233/JAD-180370

[7]   Tochel C, Smith M, Baldwin H, Gustavsson A, Ly A, Bexelius C, et al. 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. https://doi.org/10.1016/j.dadm.2018.12.003

[8]   Greene JC. Is mixed methods social inquiry a distinctive methodology? J Mix Methods Res 2008;2(1):7–22. https://doi.org/10.1177/1558689807309969

[9]   O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. BMJ 2010;341:c4587. https://doi.org/10.1136/bmj.c4587

[10]   Ly A, Wilkinson T, Sudlow C, Bexelius C, Reed C, Gallacher J, et al. First list of priority real world evidence relevant outcomes for AD. https://roadmap-alzheimer.org/wp-content/uploads/2017/04/116020_ROADMAP_D2.1_First-list-of-priority-RWE-outcomes-for-AD.pdf; 15 Mar 2017. Accessed 05 Feb 2019.

[11]   Webster L, Groskreutz D, Grinbergs-Saull A, Howard R, O'Brien JT, Mountain G, et al. Core outcome measures for interventions to prevent or slow the progress of dementia for people living with mild to moderate dementia: systematic review and consensus recommendations. PLoS One 2017;12(6):e0179521. https://doi.org/10.1371/journal.pone.0179521

Table 1   AD-related outcome domain categories and related outcomes included in the ROADMAP Data Cube.

Outcome domain category and related outcomes
Clinical diagnosis
Subjective memory complaint
Subjective cognitive complaint
MCI
Amnestic MCI
Non-amnestic MCI
Dementia
AD
Preclinical AD
Prodromal AD
Other dementia subtypes
Vascular dementia
Mixed dementia
Frontotemporal dementia
Dementia with Lewy bodies
Dementia in other diseases classified elsewhere
Disease severity and progression
Staging severity of dementia
Global improvement/stability/decline
Cognitive abilities
Memory
Language and communication
Visuospatial ability
Attention/executive functions
Intelligence
Getting lost in own home
Not recognizing family
Spatial awareness
Conscious awareness
Repeated questions
Losing the sense of who you are
Other
Functional ability and independence
Clinical judgment on dependence/independence
Hygiene/bathing
Dressing
Toileting
Transferring
Walking
Mobility
Eating
Ability to use telephone
Shopping capacity
Food preparation
Housekeeping
Laundry

Transportation/driving capacity

Responsibility for own medications

Financial capacity

Management of everyday technology

Supervision (preventing dangerous events)

Communication

Social engagement

Behavioral and neuropsychiatric symptoms

Sleep disturbance/change in sleep–wake patterns/night-time behaviors

Appetite/eating disorders

Change in physical activity

Aggression

Agitation

Irritability

Disinhibition

Motor disturbances

Euphoria

Delusions

Hallucinations

Depression

Anxiety

Dysphoria

Apathy/indifference

Sight

Emotional issues

Sensory changes and challenges

Swallow reflex

Self-efficacy

Medical investigations

Genetic biomarkers

ApoE4

Other variants

Amyloid biomarkers

CSF

PET

Plasma

Tau biomarkers

CSF

PET

Plasma

Functional neuroimaging biomarkers

FDG-PET

Structural neuroimaging biomarkers

Whole-brain volume change/atrophy rate

Hippocampal/medial temporal lobe volume change/atrophy rate

Ventricular volume change/atrophy rate

Others

Other genetic, amyloid, tau, neuroimaging biomarkers

Physical examination

Neurological examination

Assessment by HCP

Date and frequency of healthcare appointments Date and frequency with which tests are administered
Use of healthcare and social services Health service costs Access to information related with the disease *Patient-related items* Living accommodation Patient healthcare resource use Copayment patient *Caregiver-related items* Caregiver work status, impact, work days missed Caregiver sleep Caregiver healthcare resource use Personal financial items Formal caregiver time Informal caregiver time
Therapeutic treatment Starting medication for symptomatic AD Starting antipsychotic medication Other medications and dosage Side effects Medical device use Other therapeutic interventions
Significant disease-related life events Ability to drive Hospitalization Institutionalization Need for assistance at home Need for full-time care Safety Premature loss of paid employment Respite care take-up Sick leave Welfare support (monetary support) Guardianship measures
Patient QoL Self-reported QoL/health utility Proxy-reported QoL/health utility Impact on relationships and marriage strain/break-up Maintaining ability to participate in hobbies
Quality of the carer's and family's lives Caregiver perceived burden Caregiver comorbidities Caregiver's QoL Spouses' "duty" to care Quality of the patient–caregiver relationship Caregiver social support
Mortality and comorbidities *Mortality* *Comorbidities*

Stroke
Cerebrovascular disease
Coronary heart disease
Diabetes
Other vascular diseases
Obesity
Parkinson's disease
Epilepsy
Other neurological disorders
Cancer
Musculoskeletal diseases
Respiratory diseases
Genitourinary diseases
Renal diseases
Eye-related impairments
Hearing impairments
Fractures and other trauma
Delirium
Neuropsychiatric disorders
Frequent infections
Frailty

Abbreviations: AD, Alzheimer's disease; ApoE4, apolipoprotein E4; CSF, cerebrospinal fluid; FDG-PET, fluorodeoxyglucose positron emission tomography; HCP, healthcare professional; MCI, mild cognitive impairment; PET, positron emission tomography; QoL, quality of life.

Fig. 1.	The ROADMAP Data Cube, integrating AD data sources, outcomes, and disease stages.

Description: The Data Cube provides a three-dimensional overview of identified AD data sources, prioritized AD outcomes, and relevance of the outcomes for the different disease stages. The different heat maps visualise the presence or absence of each outcome per data source, the importance of each disease stage for the different outcomes, and the presence of the different outcomes in each data source and their relevance for the different disease stages. Abbreviations: AD, Alzheimer's disease.

Fig. 2.	Heat map showing the availability of outcome domains and selected outcomes in the different data sources.

Description: Shading represents available outcomes, white sections represent an outcome that is not available. Abbreviations: ApoE4, apolipoprotein E4; HCP, healthcare professional; EHR, electronic health record.

Fig. 3.	Outcomes identified from the systematic literature review, consultations, and surveys in relation to their relevance to stakeholder groups.

Description: Venn diagrams showing the identified outcomes relevant to people with dementia (top), caregivers (bottom left), and healthcare professionals (bottom right). Three different methods were used to identify these prioritized outcomes: systematic literature review, consultations, and surveys.