D3.1 Initial Overview of potential data sources with RWE data in Europe

116020 - ROADMAP

Real world Outcomes across the AD spectrum for better care: Multi-modal data Access Platform

WP3 – WP Identification, mapping and integration of RWE

Lead contributors	Antje Hottgenroth (17 – Eli Lilly)
	Stephanie Vos (4 – UM)
Other contributors	Pieter Jelle Visser (4 – UM)
	Johan van der Lei (3 – EMC)

Due date	31/01/2017
Delivery date	15/05/2017
Deliverable type	R
Dissemination level	PU

Description of Work	**Version**	**Date**
	V1.0	27/10/2016

Table of contents

Document History

Version	Date	Description
V1.0	04/04/2017	First Draft
V2.0	27/04/2017	Internal formal Peer Review: Josep Garré-Olmo (IDIAP JORDI GOL), Dominic Paes (BIOGEN), Lena Johansson (UGOT)
V2.0	15/05/2017	Comments from Robin Flaig (UK Biobank), final internal review and submission.

1. Introduction

The identification and characterisation of the data sources in order to create a landscape of available real world health care data is the first step to provide an understanding of currently available data for the relevant outcomes and outline potential gaps in currently available information about Alzheimer's disease at various stages. Since the landscape of data should be sustainable information that is available for the current ROADMAP objectives as well as future research on the various stages of Alzheimer's disease it should be documented in an accessible data source catalogue with curation and search features. For that purpose, the existing EMIF AD and EMIF EHR catalogues were selected as the preferred repositories, (https://emif-catalogue.eu/), but at the same time taking into account the different fingerprinting needs for cohort type and EHR/national register type data (see figure 1).

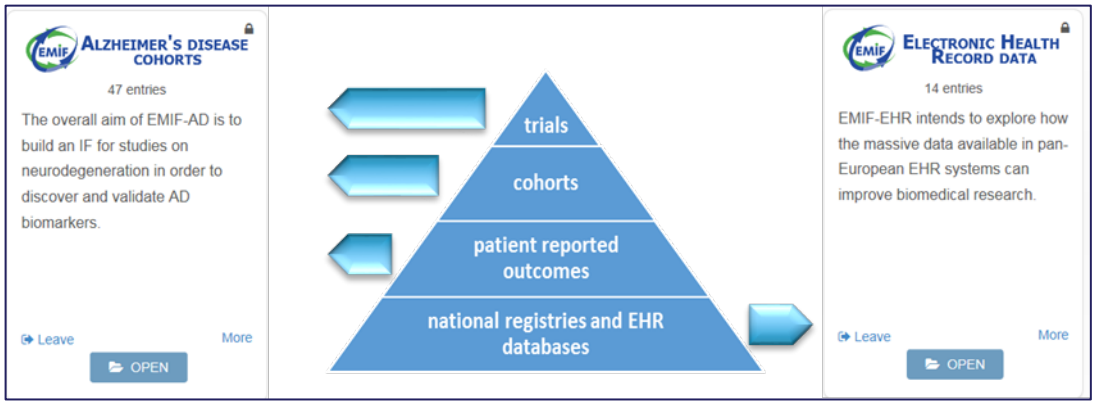

Figure 1. *The Data Pyramid – EMIF catalogue selection*

Both catalogues provide data characterization items, which are already designed for fingerprinting Dementia/Alzheimer's disease related characteristics and can be further adjusted by adding characteristics as identified in ROADMAP work packages; this is ongoing and will be available for the Final Overview.

The Initial Overview was fed mainly by 5 knowledge resources:

- The ROADMAP consortium list of accessible data sources (FPP)
- The EMIF AD + EHR catalogue current fingerprinted data sources
- The DPUK catalogue of data sources
- The EU Dementia Mapping project results
- ROADMAP partner data source landscaping project

These resources will be investigated in depth for the Final Overview of Potential Data Sources in the due course of the project.

2. Methods

The method used for Deliverable 3.1 which was to provide an Overview of potential data sources with RWE data in Europe, was mainly through interrogation of knowledge readily available to ROADMAP consortium members and similar mapping exercises. Further methods will be added as deemed necessary and feasible (e.g. literature review) based on the data needs identified by other work packages in order to provide the Final Overview of potential data sources with RWE data in Europe, which is due at month 18.

Knowledge Resources for the Initial Overview:

- EMIF-AD + EHR catalogue
- DPUK catalogue
- EU dementia cohort mapping project
- Data source mapping project results provided by consortium members (Lilly)
- Networking of consortium members with the scientific community

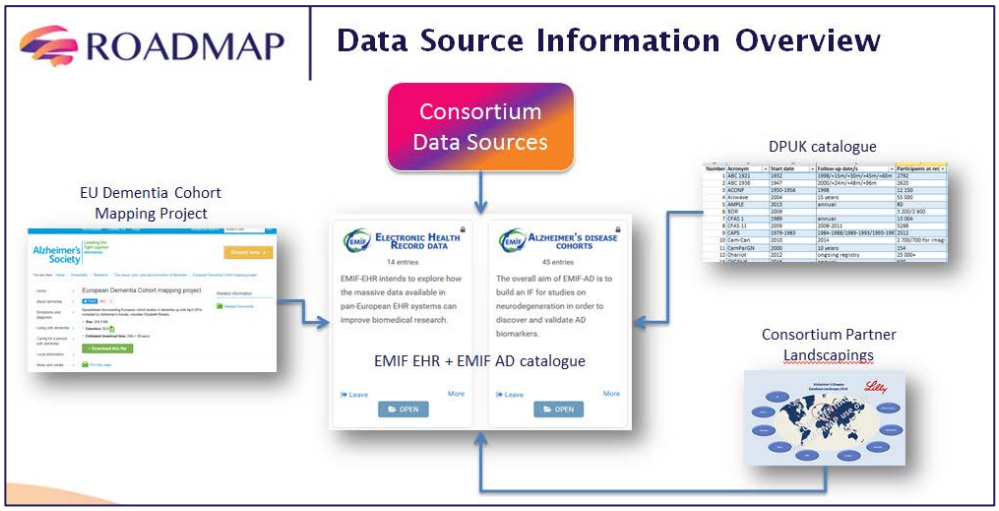

Figure 2. *Data Source Information Overview*

Furthermore, results from literature reviews performed by WP2, WP4 and WP5 will be scanned for further potential data sources.

Data sources are characterized using general descriptors to document the population, disease state and geography covered. Information on the physical location of the datasets, mode/level of access for analysis and data privacy/ethical review requirements is or will be documented for each data source. In addition, there will be a deep dive into the scientific characteristics on collected outcomes, patient characteristics and disease indicators. Source-by-source extracts for ROADMAP partner's data source fingerprints are attached in **Annex I and II**.

3. Results

3.1. Consolidation of Data Source Information from Knowledge Resources

Data sources identified through different Knowledge Resources were consolidated to identify overlap and to provide an initial list of unique sources of data. This identified 250 unique data sources in Europe, which are listed in Annex IV. In addition, Annex IV also includes the sources of information (e.g. DPUK catalogue extract and EU Dementia Cohort Mapping Project Details).

3.2. ROADMAP consortium accessible Data Sources

The ROADMAP consortium does have access to several of the identified data sources already. The level of access governance to these data sources varies and there are 3 access governance scenarios (see Figure 3):

- The ROADMAP consortium partner has direct governance over the data and is the sole/main decision maker to provide access.
- The ROADMAP consortium partner is part of the governance committee or has experience with the application process and will serve as the facilitator to request access to the data.
- ROADMAP has to follow the usual application process for access to the data – no ROADMAP partner will be the single point of contact for facilitation.

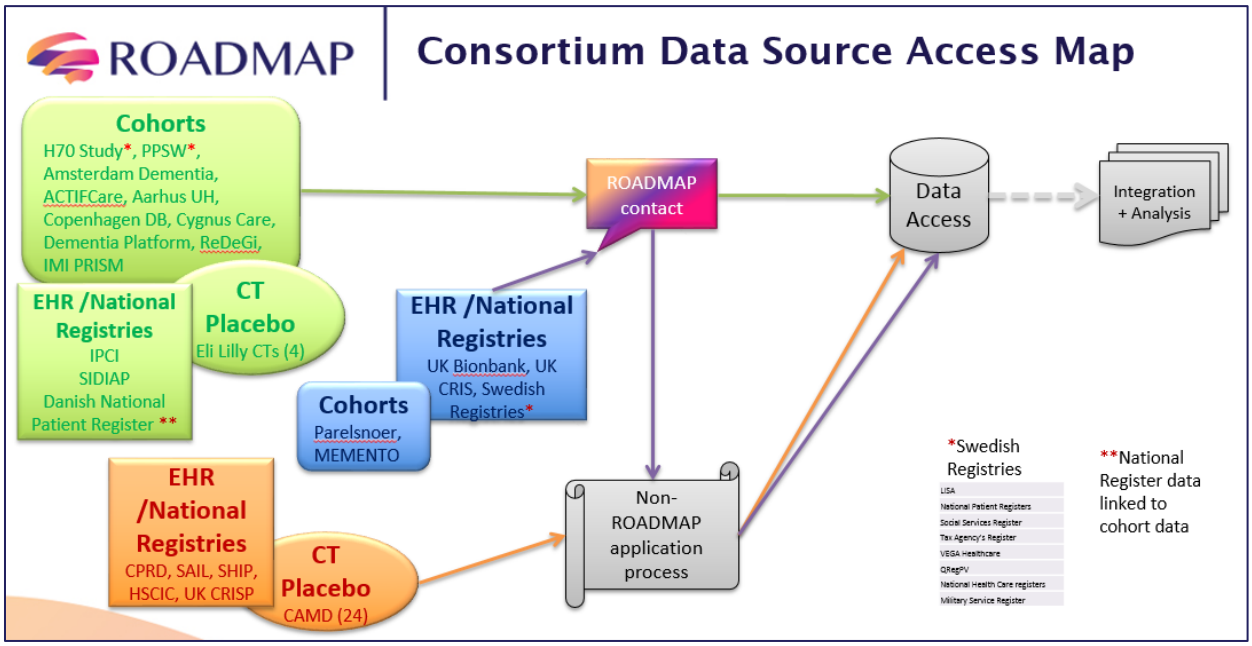

Figure 3. *Consortium Data Source Access Map*

A contact list for all data sources was created including high level information on expected timelines for access application and approval and is now integrated into the ROADMAP general contact list. It

should be noted, that requests for access to the data have to be submitted with a study/project outline and access governance processes of each individual data source have to be followed.

3.3. National Registries and EHR databases

National Registries, national claims and EHR databases are available through ROADMAP consortium members in several countries and others might be approached subsequently. Most of the national registers are well described and in the public domain and information will not be transferred to the EMIF-EHR catalogue in these cases (e.g. various Swedish national registries). EHR database information will go into the EMIF EHR catalogue and for disease-specific registries it will be determined based on the general characteristics, whether a fingerprinting in the EHR or the AD catalogue is of most benefit. See **Annex I** – Fingerprinted consortium EHR.

3.4. ROADMAP consortium accessible Cohort Overview

Information about existing cohort data is already available in the EMIF AD catalogue including timelines for data access. Several consortium members are providing access to cohort data, which are either already characterized in the EMIF AD catalogue or fingerprinting is in progress.

For the cohort data in the DOW, which are readily accessible through consortium members, the following table provides a high-level overview on data characteristics (green=available, red=not available, S=subset of data, CT= Computed Tomography). A detailed fingerprint is attached in **Annex II** and could be accessed through the EMIF AD catalogue. Fingerprinting of the SveDem and the Cygnus care cohort is in progress. The identification and fingerprinting of additional cohorts of interest is ongoing.

Cohort	Clin info	Dem rating scale	Neuropsy scale	QoL	Caregiver	Health resources	Cogn screening test	Neuropsy testing	Phys exam	Blood	CSF	MRI	PET
Amsterdam													
Parelsnoer													
LeARN													S
ActifCare													
The H70 Study											S	S (CT)	
PPSW											S	S (CT)	
SveDem										S	S	S	S
Dementia platform UK										S	S	S	S
Cygnus care cohort										S	S	S	
Memento													
ReDeGi										S		S	

Figure 4. *Data Characteristics Overview for Consortium Cohorts*

3.5. Clinical Trial Placebo Data Overview

Through the CAMD Institute, clinical trial placebo data are available for analysis for Alzheimer's and MCI patients. Consortium members have conducted clinical trials as well and are expected to share placebo data insights. The characteristics of the data, which will be available, will be added to the EMIF AD catalogue and more detailed descriptions of all placebo data items will be available for the Final Overview.

For the clinical trial placebo data, which are readily accessible through consortium members the table provides a high level overview on data characteristics (green=available, red=not available) – see also Annex III.

Name	Patient Population	Duration (weeks)	N	Access Type	Data Location	Finger-print Location	Demo-graphics	Vital Signs	Diagnosis	Functional Scales (item level)	Genotype	Labs	AD Biomarkers	Conmeds	Adverse Events
CAMD Alzheimer's Clinical Trial Placebo Data 24 trials	MCI (3) AD- mild to moderate AD- unspecified (1)	12-104	6500	Application process as of CAMD	CAMD CODR	EMIF AD (TBD)	Age, gender, race, ethnicity, country	BP, Heart Rate, Height, Weight, BMI, Temp, Resp. Rate	Primary, Family History, General Medical History	MMSE, ADAS-COG, others as collected	APOE (MTHFR)	All labs collected; CTS	Imaging - No EEG - No Biofluids - No Expanded Genetics - No	Memantine, AChEi, General Medications	Event, Severity, Duration
Eli Lilly 1 - Placebo Data	AD – mild to moderate	76	501	Consortium	TBD	EMIF AD									
Eli Lilly 2 - Placebo Data	AD – mild to moderate	76	553	Consortium	TBD	EMIF AD									
Eli Lilly 3 - Placebo Data	AD – mild to moderate	80	506	Consortium	TBD	EMIF AD									
Eli Lilly 4 - Placebo Data	AD – mild to moderate	80	519	Consortium	TBD	EMIF AD									

CAMD: https://c-path.org/programs/camd/ (Tools and Teams)
Eli Lilly 1-4: Henley et al. Alzheimer's Research & Therapy (2015) 7:43

Figure 5. *Data Characteristics Overview for Consortium Cohorts*

3.6. Patient Reported Outcomes

Patient Reported data are available through ROADMAP consortium members and characteristics are directly available (e.g. IMI PRISM and The Cygnus Study). Information gathering is in progress to be transferred to the EMIF AD catalogue.

4. Summary

A list of relevant data sources is available and several are accessible through consortium members. These should be used to populate the data cube of evidence. The characterization and documentation is currently ongoing, as well as adjusting the fingerprinting information needs based on the relevant outcomes and markers identified by work packages 2, 4 and 5.

Furthermore, searches for relevant data sources will be initiated based on identified gaps during the progress of the ROADMAP project and included in the Final Overview.

ANNEXES

ANNEX I. EMIF Fingerprinted Consortium EHRs


EHR AHU


EHR IPCI


SIDIAP

ANNEX II. EMIF Fingerprinted Consortium Cohorts


Actifcare


Amsterdam cohort


H70


LeARN


Memento


PPSW


PSI


ReDeGi

ANNEX III. Clinical Trial Placebo Data

Clinical Trial Placebo Data

ANNEX IV. Data Source Landscape Grid

Data Source Landscape Grid