*Liesbet M. Peeters*<sup>1</sup>, Ilse Lamers<sup>1</sup>, Dirk Valkenborg<sup>2</sup>, Peter Feys<sup>1</sup>, Veerle Somers<sup>1</sup>, Annemie Spooren<sup>1,6</sup>, Veronica Popescu<sup>3</sup>, Niel Hens<sup>2</sup>, Paul M. Matthews<sup>4</sup>, Christoph Thalheim<sup>5</sup>, Bart van Wijmeersch<sup>1,3</sup> & Niels Hellings<sup>1</sup>

1 Biomedical Research Institute, Hasselt University, Belgium; 2 Center of Statistics, Hasselt University, Belgium; 3 Revalidation and MS center, Overpelt, Belgium; 4 Data Science Institute, Imperial College Londen, UK OPTIMISE, United Kingdom; 5 European Multiple Sclerosis Platform, Schaarbeek, Belgium; 6 PXL University of Applied Sciences and Arts, Belgium

### MS DATACONNECT - CONNECT DATA, CONNECT PEOPLE -

www.msdataconnect.com

We aim to transform the mostly population based management of MS of today into an individualized, personalized and precision level management. We believe the key to achieve this next level is "FAIR" data.

# FAIR data for next generation management of Multiple Sclerosis **A FOUR C-PLAN FOR SUCCCESS**

### **Step 1: COLLECT**

Develop data collection procedures and tools to create data that

is FAIR (Findable, Accessible, Interoperable and Re-usable)

# F.A.I.K. data

UHASSELT

KNOWLEDGE IN ACTION



#### Findable, Accessible, Interoperable and Re-usable



### **Step 2: CONNECT**

Develop IT solutions to allow (temporarily) pooling and linking of FAIR datasets



### Proof-of-concept is provided by building a *multidisciplinary* repository for MS in Belgium

#### COLLECT

Data is collected 1° in the hospitals or laboratories using the primary systems 2° by patients 3° by independent practitioners using the electronical medical dossier (EMD). The data is registred and captured by optimizing and expanding 3 existing open source IT software (HD4DP, HD4PROM, HD4PrC) (=0 & 1 in fig)

#### CONNECT

### **Step 3: COMPLETE**

Develop statistical methods to define minimal requirements for datasets



## Step 4: CONSTRUCT

Develop new analytical methods for optimal mining of connected

Data is securely transferred, encrypted and pseudo-anonymized by a Trusted Third Party. The use of the National Registry Number as identification number in all data collection steps makes it possible to connect all data belonging to the same patient (2 in fig).

#### COMPLETE

The HD4RES software is used by the researchers for monitoring, follow-up and managing the data collections (3 & 4 in fig)

#### **CONSTRUCT**

Data is stored in a secure data warehouse. Researchers are greanted access and analysis of the data is possible in the data warehouse using standard statistical tools. Data is used for research, for example to build personalized decision support systems and to evaluate and construct composite outcome score measures. The repository is used to provide Health Statistics for all stakeholders involved (5, 6 & 7 in fig)



An online catalogue listing the desired content of the

### and pooled FAIR datasets



An *intuitive* representation of this 4C plan is presented. Data is collected all over the world by different stakeholders (step 1: COLLECT). This results in an extensive amount of data and datasets. Every dataset is represented using a puzzle of a face. Many insights could be discovered when these datasets could be pooled and connected (step 2: CONNECT). Methods to identify the minimal requirements for common datasets are required (=step 3: COMPLETE). When sufficient overlap between the databases involved is secured and powerful analytical methods are developed to cope with the imperfections of datasets featured by different layers of missing data, these data sets can be optimally mined to create new insights for MS management *(=step 4: CONSTRUCT).* 



#### Let's dream... and team up **Dr. Ir. Liesbet M. Peeters**

We are open for collaboration. Please contact us for more info info@msdataconnect.com