



EMIF Deliverable 10.1: Initial database fingerprinting

Executive summary

Executive Summary

In the EMIF/Platform network, detailed information describing each database must be available and will provide the context in which data are collected (e.g., EHRs from primary care, enriched with data from pharmacies and hospitals), the source population (e.g., all inhabitants of a certain region), the data elements (e.g., reasons for encounter, diagnosis, laboratory data, images, genetic data, etc.), the local data model, the coding schemes for data elements (e.g., READ, ICD, ICPC, ATC-codes, LOINC, etc.), presence of data elements in free text, data latency and update frequency, and internal data quality standards and verification. Additional possibilities, such as family linkage, will be specifically included in the profile of each database as well. Finally, each organization in the network will be asked to characterize their ability to perform additional specific tasks in the context of re-using the data, such as the ability to locate the patient care provider in order to collect new data (e.g. through questionnaires), or material (e.g., blood samples) for specific studies, or to obtain permission for inclusion in trials.

Fingerprinting is the process of collecting this information. The idea is that a fingerprint should offer a fairly complete and faithful view of a data source so that ultimately end users 'browsing' fingerprints can make decisions about which sources would be suitable for their specific research questions.

As a first step in fingerprinting data sources, we developed an initial questionnaire. This deliverable contains the answers to this initial questionnaire. The data and some organizational/ethical characteristics of each partner in the network are described, as well as a more specific example of future applications, in which each database provided the algorithm to obtain specific secondary information (prevalence of diabetes mellitus). Finally a list of publications was listed by each partner.

Based on this initial questionnaire, the next round of fingerprinting will be developed, which will be subject of the next deliverable D10.2.

Contacts

EMIF-Platform: Johannes van der Lei – Nigel Hughes
j.vanderlei@erasmusmc.nl - nhughes@its.inj.com