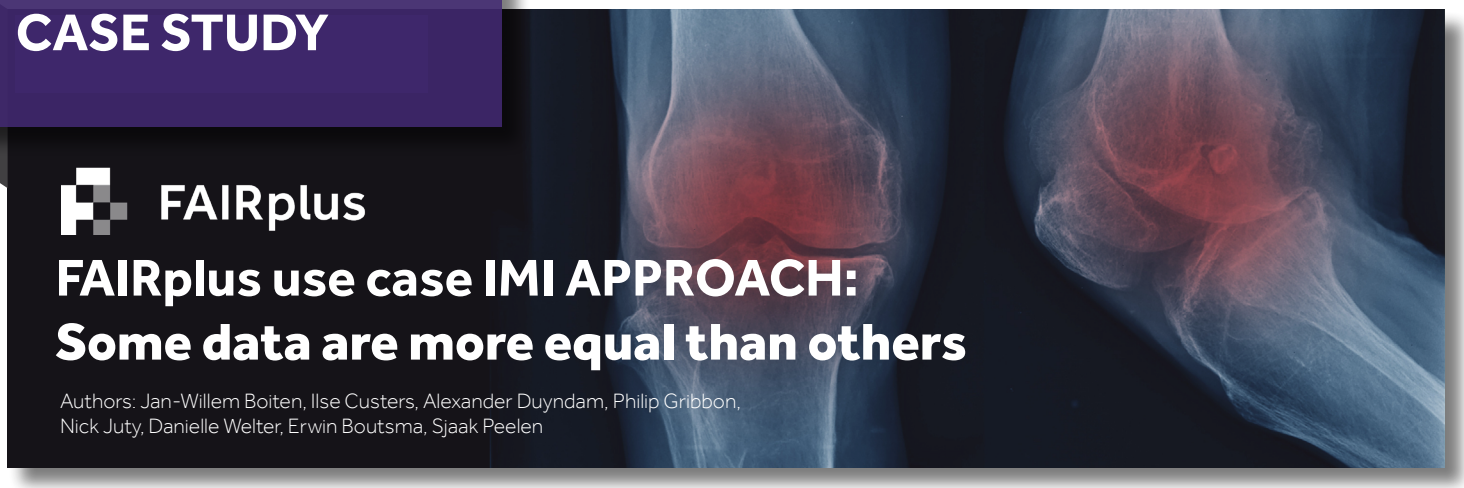
CASE STUDY



FAIRplus

FAIRplus use case IMI APPROACH: Some data are more equal than others

Authors: Jan-Willem Boiten, Ilse Custers, Alexander Duyndam, Philip Gribbon, Nick Juty, Danielle Welter, Erwin Boutsma, Sjaak Peelen

Challenge

Although having a major impact on the quality of life of those involved, for a disease like osteoarthritis there is no universally accepted ontology yet to describe the needed clinical data. Hence data cannot easily be shared and evaluated across clinical research projects.

Solution

The granular clinical data harvested in the APPROACH sites have been mapped by the FAIRplus squad to the standard CDISC-SDTM data model. This improves the interoperability as mapping to other ontologies is available.

Overview

Osteoarthritis, a chronic inflammation joint disease, is an unmet medical need, causing moderate to severe disability in over 40 million people worldwide. Many patients struggle to carry out ordinary daily activities. Developing effective treatments for this debilitating condition is extremely challenging. The IMI APPROACH project is creating a platform comprising of data from thousands of patients and healthy people to identify groups of patients with similar profiles, assuming these subgroups would have a better chance to respond well to specific treatments. This patient stratification is supported by newly identified and validated biomarkers to measure disease progress.

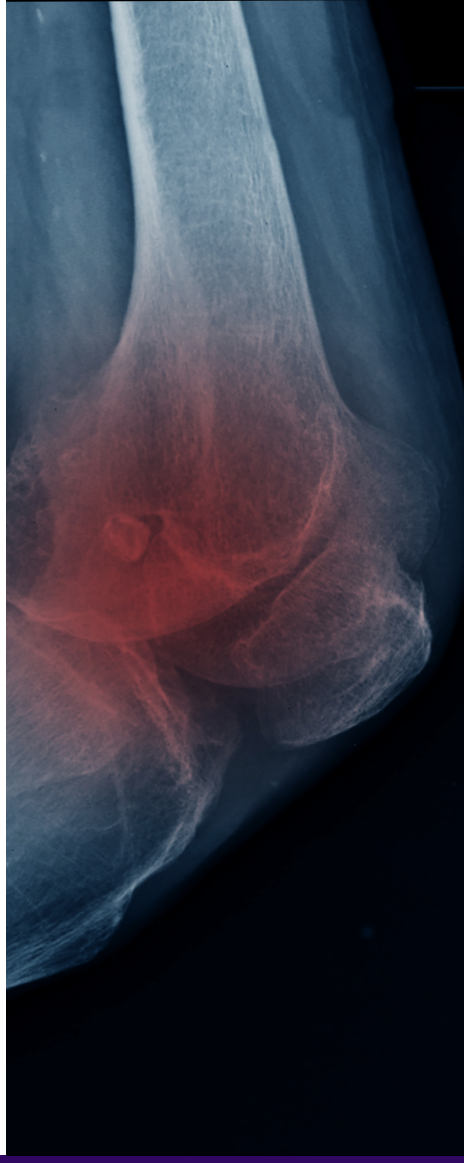
The IMI FAIRplus project aims to develop tools and guidelines for making life science data FAIR (Findable, Accessible, Interoperable, Reusable). In the past year, the so-called 'squad teams' from FAIRplus, consisting of experts working in universities and pharmaceutical companies, have been actively working to FAIRify data sets from other large IMI projects such as APPROACH. The developed tools and methods are subsequently added as 'recipes' to the FAIR cookbook, enabling projects and companies with similar FAIR data challenges to apply this consolidated know-how to increase the FAIRness of their data.

Aims and impact

The IMI projects to be supported by FAIRplus are selected in an unbiased process, based on a recorded set of criteria, among others societal impact. FAIRifying the data would facilitate sharing the data outside the consortium to conduct further research. The first set of projects the FAIRplus squad teams worked on consisted of pre-clinical data (eTox: in vivo and in vitro toxicology data, RESOLUTE: proteomics and transcriptomics). With the selection of APPROACH, a second round started since it involves also clinical data.

The APPROACH project aims to combine the clinical data from thousands of patients and healthy people. To have a proper license is critical when sharing the data in public registries as planned for these APPROACH data. A dedicated recipe in this respect in the FAIRplus Cook Book provides guidance on licensing [FCB035].

The Data Steward of APPROACH, Sjaak Peelen (Lygature), indicated that the data were already standardized within APPROACH between the five clinical centers, but not for use outside the consortium. Initially working with just the APPROACH data dictionary and a small set of illustrative data, Danielle Welter (University of Luxembourg), together with the FAIRplus squad, had a hard time to understand the context of some of the collected variables due to their highly specific nature and the lack of standard ontologies for



osteoarthritis. Peelen then provided a more complete dataset, fully anonymized to protect the privacy of persons involved. The FAIRplus squad was then able to map the data fields to the standard data model of CDISC-SDTM and further onwards to other standards, such as SNOMED-CT and RxNORM. This mapping improves the Interoperability within the FAIR framework.

Analyzing the data collected in APPROACH is still work in progress for researchers in the APPROACH consortium. The next step will be to outline the results in scientific publications and publish the data in registries and data catalogs. When these data fields – as part of the FAIRification process – are mapped to a recognized data model, this will promote future use of the APPROACH data.

Additionally, aggregated metadata for the main APPROACH patient cohort was submitted to the IMI Data Catalog and the Biosamples database. This aids the Findability of APPROACH data as it enables researchers looking for existing data for meta-analyses to discover the existence of the APPROACH data through these catalogs and potentially reuse the data in further research.

What are the lessons learned of the involvement of the FAIRplus squad? *'**FAIR by design** should have been the way to go at the start'*, Peelen says. *'Also, researchers should get educated about FAIR. Another issue is that a consortium, such as APPROACH, is not a separate legal entity. This means that often separate agreements with multiple partners are needed. Instead, a single agreement should be put in place at the time when the consortium agreement is being set up and not at a later stage. It is important to stress for researchers that FAIRifying their data helps them with their scientific publications as their data will be more mature.'*

Conclusions

FAIRifying data is a very useful exercise increasing the value of the data collected, in particular for the wider research community. Unfortunately, FAIRification of the APPROACH data was done after most data already had been collected. For future projects it is advised to include FAIR into the design of the data collection. The prerequisite for this is to have project resources knowledgeable on FAIRification. However, even with a post-data collection FAIRification, additional value was added to the APPROACH data, which is a good example for future projects. Because if all future projects ensure their data are FAIR, we will have achieved a huge benefit.

The benefit

The FAIRification of the APPROACH data has increased the shareability and interoperability of the data set and it has enriched the FAIR Cookbook, herewith facilitating future research that combines data of multiple (osteoarthritis) studies. It is a benefit also for the partners that contributed in APPROACH.

FAIRplus

FAIRplus Project
FAIRplus-PM@elixir-europe.org

The FAIRplus project has received funding from the Innovative Medicines Initiative Joint Undertaking under grant agreement No 802750. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA Companies. This communication reflects the views of the authors and neither IMI nor the European Union, EFPIA or any Associated Partners are liable for any use that may be made of the information contained herein.