

Deliverable D8.1 Report on basic model for data sharing and access within the consortium

A central element of the PIONEER Project will be the identification of datasets relevant to Prostate Cancer and, having identified as many appropriate datasets as possible, arranging for the data therein to be accessible to the PIONEER Project. It is thus a necessary step to identify an appropriate model for such data sharing and access.

WP8 has considered available data models and analysed their features, advantages and disadvantages. It has also sought to learn from other IMI BD4BO projects and in particular, the helpful recommendations made by DO-IT.

We have also assessed the challenges arising from data protection legislation (particularly following GDPR) where data is personally identifiable patient data.

Our conclusion has been to go forward on the basis of a mixed model, with initial data processing being undertaken under a Federated Database Model, with a tailored and de facto anonymised Centralised Database Model being utilised where dataset providers are willing to share data on this basis. At the heart of the strategy is the avoidance of processing personally identifiable data, but seeking to maximise the usefulness of data by pursuing this de facto anonymised Centralised Database model where practicable.