Harnessing the power of big data in prostate cancer research - Data access and sources in the PIONEER project

Susan Evans-Axelsson$_a$, Thomas Abbott$_b$, Louise Fullwood$_c$, Maxim Moinat$_d$, Carl Steinbeisser$_e$, Alex Asiimwe$_e$, Muhammad Imran Omar$_{fg}$, James N'Dow$_g$, **Anders Bjartell**$_a$ and PIONEER Consortium*

a. Lund University, b. Astellas, c. Pinsent Masons, d. The Hyve, e. Bayer AG, f. University of Aberdeen, g. European Association of Urology

Background: Prostate Cancer DIagnOsis and TreatmeNt Enhancement through the Power of Big Data in EuRope (PIONEER) is an Innovative Medicines Initiative 2 pan-European partnership consisting of 32 public and private stakeholders across 9 European countries. PIONEERs goal is to ensure the optimal care for all European men living with prostate cancer by unlocking the potential of Big Data and Big Data analytics.

Work package 3 aims to identify, approach and negotiate appropriate data sharing agreements with a variety of holder of high-quality, real-life prostate cancer-based datasets across European and non-European patient populations. WP3 will collect, standardize and harmonize existing prospective and retrospective data into a single innovative data platform.

Methods: To effectively tackle WP3 work plan, subgroups were formed:

- WP3.1 tasks are to contact the data sources, access their willingness to participate and to collect information about the contents of their database(s)
- WP3.2 negotiates appropriate Data Sharing Agreements (DSA)
  - PIONEER will use a mix of both the Federated (remote data) and Central (importable data) Database Models.
  - WP3 and WP8 (legal, ethics and governance) have established two standardized DSA templates to be used with data contributors.
- WP3.3 will work to convert, harmonize and map the data sets into a common data model similar to other IMI projects
- WP3.4 will establish Data Management Plans to support PIONEER sustainability goals

Results: As part of the initial proposal for the PIONEER consortium, 27 potential Data Providers were identified. WP3 is actively engaged in outreach and this number has now grown to over 60 potential data sources. As of September 2019, **ERSPC Rotterdam and PRIAS (Europe)** datasets are in the process of being mapped to OMOP (Federated sharing model). **Malmö Diet and Cancer/Malmö Preventative Medicine** (Sweden) have also joined and have chosen to use the Central sharing model. Access to clinical trial datasets from industry partners, in most cases, will be provided through **Project DataSphere**. Negotiations are underway for access to databases from Australia/New Zealand, Germany, Finland, the UK, the Netherlands, Denmark, Norway and the USA.

Conclusions: PIONEER has infinite potential thanks to its potential size and content. By applying big data analytics and developing a data platform of unparalleled scale, quality and diversity, PIONEER will be able to answer questions about prostate cancer in a new way that will empower meaningful improvement in clinical practice, prostate cancer disease-related outcomes and health economic outcomes across the European healthcare landscape.

Conflict of Interest: No conflicts of interest related to this presentation