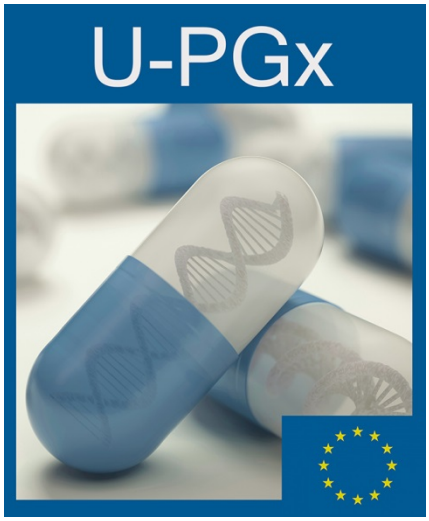
Project Number: 668353

Project Acronym: U-PGx

Project title: Ubiquitous Pharmacogenomics (U-PGx):
Making actionable pharmacogenomic data and effective treatment optimization accessible to every European citizen



DATA MANAGEMENT PLAN

1. Data Summary

For this DMP 'data' are defined as the raw data collected and registered in the eCRF in ProMISe, and all analytical results generated from the collected biomaterials in the randomized control trial PREPARE. The purpose of the data collection is to answer the scientific question whether pre-emptive genotyping of an entire panel of clinically relevant PGx markers for which DPWG guidelines are available (pharmacogenes) can be used in clinical practice, is cost-effective, and results in better outcome for patients. This related to the objectives to implement pre-emptive PGx testing in routine clinical practice in seven clinical sites in the EU taking into account the diversity of European health care settings (WP4); to assess cost-effectiveness of PGx testing in each implementation site (WP5); to develop pharmacometric models to predict expected effects for genetic variant groups on various outcomes, and identify new clinically relevant drug-drug interactions and implement the knowledge on the impact of PGx polymorphisms on drug-drug interactions into the DPWG guidelines to further improve PGx testing in clinical routine (WP9); and to develop and validate a bioinformatics tool to identify the selected panel of PGx markers in existing NGS data and use NGS to identify novel genetic variants in subjects with extreme phenotypes included in PREPARE (WP10). Existing NGS data will be re-used. New data will be collected from patients included in PREPARE at 7 clinical sites between March 2017 and March 2020 (for details see CPT paper 2017, van der Wouden et al.). The size of the data is unknown. The data will be useful for the U-PGx consortium, (clinical) pharmacists, pharmacologists, geneticists, and clinicians with interest in personalized medicine.

2. FAIR data

2.1. Making data findable, including provisions for metadata

The data produced and/or used in the project will be discoverable via the trial register (www.clinicaltrials.gov). We will indicate the existence of the data in our scientific publications and at our website (www.upgx.eu). The identifiability of the data is not applicable. U-PGx has a unique grant agreement number which is 668353. PREPARE has a unique trial register number which is NCT03093818. Search keywords include pharmacogenomics, RCT, personalised medicine, U-PGx, and PREPARE. Once data collection is completed, the eCRF will be frozen. The data set will remain original. Metadata creation is not applicable.

2.2. Making data openly accessible

All data that will be collected will be made openly available. However, patient ID will be anonymised. The data will be documented deposited in the data centre of the U-PGx coordinator, i.e. Leiden University Medical Center, the Netherlands. Re-utilisation of the data will be offered via the U-PGx website upon request. Each request will be discussed by the Executive Board. Software tools that is needed to access the data are SPSS, R, and PLINK.

2.3. Making data interoperable

We will make use of standardized scores such as LCAT for causality, CTC for toxicity, and standardized methodologies for example the star allele nomenclature for genetic variants. We will use standard vocabulary for all data types present in our data set.

2.4. Increase data re-use (through clarifying licences)

We do not intend to licence our data. The quality of the data is assured because all data will be collected according to the GCP standard. The study and specifically data collection is monitored by an independent CRA. In addition, the program in which the data is collected, ProMISe, is ISO-certified. The data will be made available for re-use three years after formal closure of the project. The data can

be re-used upon request by third parties. The Executive Board/Coordinator will evaluate each request before giving permission for re-use. The data will remain re-usable for about five years.

3. Allocation of resources

The estimated costs for making our data FAIR is about 5000 euros per year. In the pilot phase of the DMP, no budget has been allocated for this purpose. During data collection, the PIs of all clinical sites are responsible for the data management. The coordinator is responsible for storage and re-use of the data. We estimate the costs for long term preservation of our data about 25.000 euros. In time, the gene panel and drugs may change. For this reason the biomaterials will be banked. This will require additional funding on top of the estimated 25.000 euros.

4. Data security

All our data will be anonymised. Provisions for secure storage and transfer will be determined later.

5. Ethical aspects

All ethical aspects are covered and described in the deliverable reports for WP12.

6. Other issues

LUMC FAIR principles for data management are currently being developed and will be adopted once completed.

HISTORY OF CHANGES		
Version	**Publication date**	**Change**
1.0	10.08.2017	▪ Initial version