

Coordination of TB diagnostics research:  
Enabling standards and sharing of data on the molecular basis of drug resistance

A workshop jointly organized by the New Diagnostics Working Group and CPTR

**Senate House, London, UK**  
**3 - 4 February 2014**

Workgroup 3  
Infrastructure, access and sustainability

Chair: Lara Bethke, Wellcome Trust  
Rapporteur: Ivo Van Walle, ECDC

# Objectives

- Determine compatibility of existing databases and consider possible solutions for a datashare /
- Consider existing infrastructure that could be mirrored in other systems
- Clarify private/public data needs and governance (decisions regarding linkage to resistance and regulatory collaboration)
- Clarify the policy regarding intended access to the data and the difference between open and public access
- Identify challenges and define process for curation, maintenance and sustainability
- Address funding needs
- Discuss potential and possibility for expanding into a bio-repository or specimen bank

# Workgroup 3 – Infrastructure, access and sustainability

- Centralised vs decentralised databases
  - More or less 5/12 of the current questions can be answered with the current dbs, so would not need a centralised storage.
  - Should have a front end for existing dbs, plus own storage for TB specific data.
- Laboratory vs clinical data
  - Must have
    - Define mutations that influence drug resistance, requiring both geno- and phenotype data for the same sample.
    - Clinical outcome data.
  - Nice to have
- Storage vs analysis
  - Design extensibility from the start, e.g. future drugs.
  - Data representativity required?
- Quality assurance and curation
  - Prevent low quality data from getting in: possibly through certification scheme for data providers.
  - Promote self curation.
  - Automated curation: scores for data quality, lower scores only accepted after manual curation.
  - Sequencing is easiest, phenotyping but more difficult, clinical data and especial long term follow up most difficult.

# Workgroup 3 – Infrastructure, access and sustainability

- Data access policy
  - Publicly accessible but not editable (i.e. no open source).
  - Embargo must be possible.
  - Authorship must be fairly distributed.
  - May drive the schedule, rather than technical aspects.
- Data protection
  - Data must be anonymised, already standard practice.
- Funding
  - Different funders should be aligned on what the db should look like.
  - Much easier to fund a db for a particular goal rather than just storage.
  - Requirement to upload to public repository, in order to receive funding. Possibly include minimum quality level for submission.
  - Unclear on who would lead application to funders.
- Specimen bank
  - Should have a mechanism to request the strain or find the owner of the strain through the db.
  - Not further discussed.