Stop TB Partnership

New Diagnostics Working Group

Critical Path to TB Drug Regimens

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

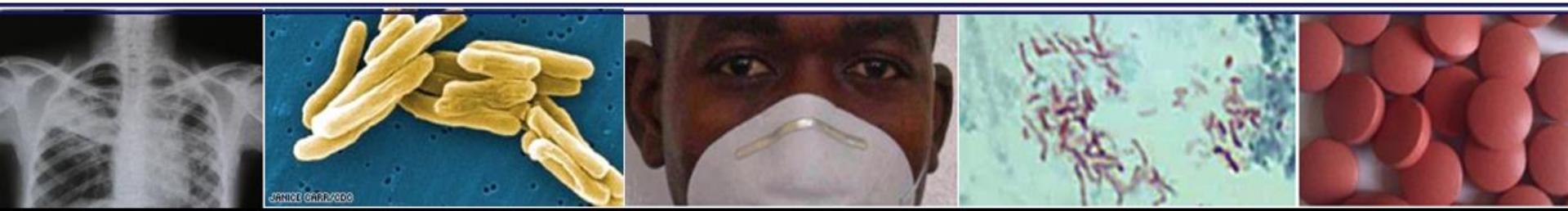
A workshop jointly organized by the New Diagnostics Working Group and CPTR

Senate House, London, UK
3 - 4 February 2014

Workgroup 2
Data Standards

Chair: Hannu Laang, European Commission

Rapporteur: Penny Wilson, TSB



Objectives

- Identify which data can be standardized or not and discuss interoperability aspects
- Discuss standards for the entire process behind deposited data from sample collections and preparation through to clinical information
- Clarify how data sets that are available have been organized
- Establish the data format required to enable integration of various databases and efficient data mining
- Define quality metrics for deposited data and a framework for validation at both microbiology and clinical levels
- Ensure correlation between genotypic and phenotypic information, as well as with clinical data
- Discuss how to handle discrepancies (i.e. between a phenotypic and a genotypic assay or within molecular assays)
- Agree on minimum requirements and on who is responsible

Some initial thinking…

- Must be TB
- Should be about Dx (other things may come from it but…)
- Need to define key questions
- Difference between prospective and retrospective

Genotypic, phenotypic, clinical

- Genotypic
  - Final agreement that should be sequence
  - WGS + Targeted
  - Need to define terminology
  - Multiple genotype numbering
  - Define Rv as reference
  - Understand the nature of the data file – size of file vs utility, Complete BAM file?
  - Minimum read and coverage?
  - Define optimal and minimum

Genotypic, phenotypic, clinical

- **Phenotypic – need to record**
  - How, where, test protocol, drug concentrations
  - Lab accreditation. EQA
  - Good to record MICs.

Genotypic, phenotypic, clinical

Clinical

- Should not be allowed to delay input
- Year of sample
- New vs relapse
- Region of the world
- Treatment regimes challenging to capture
- Define optional – accept database with evolve

- 2 out of the 3 (ie G+P, or G+C may be sufficient – do we need all 3?)

General points

- A need to sit down and define the key questions?
- Accept different levels of quality
- Periodic summary download
- Capture methodology and details of any lab accreditation and EQA
- Quality of data important but better to include than exclude
- Barriers recognised in terms of requirement to publish
- Who will carry out the work?
- Who will pay?

General points

- Don’t reinvent what already exists?
- What can we learn from existing databases?
- Can we build on what’s already out there?
- CPTR offered to update standards
- Existing data base needs more detail
- Ideally a working group to take things forward.

- Should software be distributed to labs?