

Data Analysis and Standardization

Jamie Posey, PhD
Lead, Applied Research Team
Division of TB Elimination/ Laboratory Branch
Centers for Disease Control and Prevention

February 4, 2014

National Center for HIV/AIDS, Viral Hepatitis, STD & TB Prevention
Division of TB Elimination

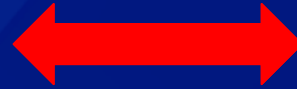


Data Sets

Genetic

Phenotypic

Clinical



Genetic Data

- ❑ DNA sequence is DNA sequence
- ❑ Targeted sequencing
 - ❑ Sanger
 - ❑ Next generation sequencers
- ❑ Whole genome sequencing
 - ❑ Analysis pipeline
 - ❑ What loci to include for each drug
- ❑ Heteroresistance
- ❑ Standardizing data entry
 - ❑ Nucleotide and protein sequence
 - ❑ One or three letter for nucleotide and amino acid

Phenotypic Data

- ❑ Not a single method used universally
 - ❑ Agar proportion
 - ❑ Absolute concentration
 - ❑ Liquid systems

- ❑ Various types of media for each assay
 - ❑ Solid – 7H10, 7H11, LJ
 - ❑ Liquid

- ❑ MIC data
 - ❑ Convert to R or S

- ❑ Enzymatic data

Analytic Capacity

- ❑ Keep it simple

- ❑ Who are the end users?
 - ❑ Academia
 - ❑ Industry
 - ❑ Laboratorians
 - ❑ Clinicians

- ❑ How will the data be used?
 - ❑ Probability a mutation confers resistance
 - ❑ What level of resistance?
 - ❑ Sensitivity / Specificity
 - ❑ Identify gaps

- ❑ Statistical Platform
 - ❑ Downloadable

Access

- ❑ Public versus tiered access

- ❑ Who is allowed to upload data
 - ❑ How is this regulated?
 - ❑ Standards
 - ❑ EQA program
 - ❑ Accredited
 - ❑ Minimal set of data

- ❑ Previously collected data versus new data