Culture results during treatment for TB as surrogate endpoints for long-term treatment outcome: The promise and the pitfalls

Patrick Phillips
patrick.phillips@ctu.mrc.ac.uk
Medical Statistician, MRC Clinical Trials Unit

WHO HIV/TB Research meeting, Rome
17th July 2011
Outline

• Terminology

• The Promise of surrogate endpoints

• The Pitfalls of surrogate endpoints

• Culture results as surrogate endpoints in TB trials

• Conclusions
Terminology

• **Biomarker**
  • “A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.”

• **[Trial-level] Surrogate Endpoint**
  • “A biomarker that is intended to substitute for a clinical endpoint.”

(Biomarker Definitions Working Group, 2001)
Terminology

- **Prognostic Marker / Individual-level Surrogate Marker**
  - A marker used for the estimation, in an individual, of the relative probabilities of the various possible outcomes of a disease.

- **Intermediate Endpoint / Auxiliary Endpoint**
  - A marker measured prior to the clinical appearance of the disease that bears some relationship to the development of that disease.
Terminology: Prentice Criteria
(Prentice, 1989)

- The statistical hypothesis test on the putative surrogate $S$ should be equivalent to the hypothesis test on the true endpoint $T$.

1. There is some association between $S$ and $T$ at the individual level.

2. $S$ should **fully capture the treatment effect** on $T$ at the trial level.

- The **treatment effect** on surrogate endpoint should reflect the **treatment effect** on true endpoint.
The Promise of Surrogate Endpoints
Phase III trials for new TB therapy

- Recognised endpoint in Phase III trials:
  - Failure or relapse within 18-24 months of starting treatment

- Endpoint rare (5%-10%) leading to trials involving a large number of participants

- Upwards of 4-5 years for a phase III trial from conception to presentation of results.
## Current TB drug pipeline

### Wide Gap

<table>
<thead>
<tr>
<th>Stage</th>
<th>Drug Name</th>
<th>Company</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preclinical Development</td>
<td>CPZEN.45</td>
<td>Caprazine nucleoside</td>
<td>New Chemical Entity (NCE)</td>
</tr>
<tr>
<td></td>
<td>Quinolone DC. 159a</td>
<td>Fluoroquinolone Antibiotics</td>
<td>NCE</td>
</tr>
<tr>
<td></td>
<td>SO008</td>
<td>Dipeptidyline</td>
<td>Sequoia NH</td>
</tr>
<tr>
<td></td>
<td>SO041</td>
<td>Capramycine</td>
<td>Sequoia</td>
</tr>
<tr>
<td></td>
<td>Benbaziloxime</td>
<td>Benzazoloximes</td>
<td>New Medicines For Tuberculosis (NM418)</td>
</tr>
<tr>
<td></td>
<td>Q201 Novel anti-TB agent</td>
<td></td>
<td>Quin Science Inc.</td>
</tr>
<tr>
<td>Phase I</td>
<td>PNU.100480</td>
<td>Oxazolidinone</td>
<td>Pfizer</td>
</tr>
<tr>
<td></td>
<td>SO109</td>
<td>Ethylisoxalamine</td>
<td>Siquoia NH</td>
</tr>
<tr>
<td></td>
<td>AZD5847</td>
<td>Oxazolidinone</td>
<td>AstaRזence</td>
</tr>
<tr>
<td>Phase II*</td>
<td>PA.824</td>
<td>Nitromidazole-oxazine</td>
<td>TB Alliance</td>
</tr>
<tr>
<td></td>
<td>Novel Combinations/Regimens</td>
<td>NCE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nitromidazole-oxazine, diaphyranone, fluoroquinone, Nitric acid derivative</td>
<td>TB Alliance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TMC207 for MDR TB</td>
<td>Diallylpyridine</td>
<td>TB Alliance, Jansen Pharmaceuticals</td>
</tr>
<tr>
<td></td>
<td>TMC207 for Drug-Sensitive TB</td>
<td>Diallylpyridine</td>
<td>TB Alliance, Jansen Pharmaceuticals</td>
</tr>
<tr>
<td></td>
<td>OPC 67663</td>
<td>Thioamide-4-ylidene-benzimidazolone, Thioamide-4-ylidene-thiamidone,</td>
<td>TB Alliance, Jansen Pharmaceuticals</td>
</tr>
<tr>
<td></td>
<td>Linezolid for MDR TB</td>
<td>Oxazolidinone</td>
<td>TBTC, Pfizer</td>
</tr>
<tr>
<td></td>
<td>Rifapentine (TBTC Study 26)</td>
<td>Rifapentine</td>
<td>CDC, Sanofi-Aventis</td>
</tr>
</tbody>
</table>

[WHO Stop TB partnership Working Group for New Drugs](http://www.newtbdrugs.org/pipeline.php)
From Compounds to Combinations

Ginsberg AM, 2010.
MRC | Medical Research Council
The Promise

- A **well validated surrogate endpoint** could be used in a phase III trial leading to:
  - shorter, smaller, cheaper trials.

- Smoother transition from **compounds to combinations**

- **Faster overall drug development process**

- But, requires **rigorous validation**

- There must be
  - Biological plausibility
  - Sufficient data/evidence from clinical trials
The Pitfalls of Surrogate Endpoints

Correlation $\neq$ Surrogacy

(Baker, 2003)
Example: ventricular arrhythmia as a surrogate for cardiac death

- **Ventricular arrhythmia** is associated with an almost 4-fold increase in the risk of death related to cardiac complications
  - Clear **individual-level prognostic marker**
    - Patients with ventricular arrhythmia are more likely to die with cardiac complications

- Encainide, flecainide and moricizine were shown to effectively suppress arrhythmias and were approved by the FDA

- More than 200,000 persons per year took these drugs in the USA
Example: ventricular arrhythmia as a surrogate for cardiac death

- The post-licensing Cardiac Arrhythmia Suppression Trial (CAST) was evaluated these three drugs in patients who had had a myocardial infarction.

- The trial was stopped early finding an increased risk of death in all three treatment arms.

- Ventricular arrhythmia is a poor trial-level surrogate endpoint

  - A reduction in ventricular arrhythmia does not necessarily correspond to a reduction in risk of death.
Causal pathways: failed surrogates

Fleming and DeMets (1996)
Causal pathways: a true surrogate

Fleming and DeMets (1996)
Culture results during treatment as surrogate endpoints for long-term outcome of TB treatment
The data

- British MRC treatment trial conducted in 1970s and 1980s across East Africa and East Asia.

- Included in the analysis:
  - 6974 trial participants on 49 different treatment arms in 12 trials (37 comparisons).

- All regimens:
  - 6 months duration,
  - 6 months of INH, at least 2 months of RIF
  - Addition of STR, PZA, EMB, THI (Thiacetazone)
  - Combination of daily and intermittent treatments.
The data

- **Highly standardized** clinical and bacteriological protocols.

- **Monthly** cultures during treatment and follow-up

- All cultures on **LJ solid media**.

- Likely only a handful of cases of HIV

- See (Fox, Ellard & Mitchison, 1999) for more details
The analysis

- Individual patient data was available from original paper treatment cards.

- **Poor outcome** is a composite endpoint defined as:
  1. Failure at the end of treatment, or
  2. Recurrence after the end of successful treatment, or
  3. TB death during treatment or follow-up.

- Two-stage analysis based on methods developed in statistical literature.
Measuring the utility of a surrogate: $R^2_{\text{trial}}$

- $R^2_{\text{trial}}$ is the established metric for evaluating surrogate endpoints (Burzykowski, 2005).

- $R^2_{\text{trial}}$ is the proportion of variation, at the comparison-level, in the treatment effect on the true clinical endpoint that is explained by the treatment effect on the putative surrogate.

- Broad guidelines:
  - $R^2_{\text{trial}} \geq 0.80$ considered ‘good’.
  - $R^2_{\text{trial}} \geq 0.95$ considered ‘very good’.
Overall results

- Each plotted point corresponds to a single within-trial treatment comparison.

- Radius of circle represents precision of estimates.

- Fitted line is weighted by precision.

- Endpoint association strongly significant (p<0.005).
Results by geographical region

- **Month 2**
  - East Africa
    - $R^2 = 0.19$
  - Hong Kong
    - $R^2 = 0.86$

- **Month 3**
  - East Africa
    - $R^2 = 0.81$
  - Hong Kong
    - $R^2 = 0.62$
Why do the results differ between Hong Kong and East Africa?

- **Hong Kong trials:**
  - more often evaluated **RIF-containing regimens**
  - saw earlier sputum culture conversion

- **Other differences:**
  - More extensive cavitation and advanced disease in EA
  - HK trials mostly urban population
  - EA trials generally earlier in time than the HK trials
  - Higher proportion of older patients in HK
Interpretation

- Geographical differences weaken the utility of endpoint

- Difference possibly due to:
  - More advanced disease leading to delayed culture conversion in East Africa
  - More regimens containing RIF in Hong Kong
  - Other unknown confounders?

- Utility of culture results as surrogate endpoints appears to be time-dependent.

- Better alternatives are likely to be:
  - Time to sputum conversion
  - Longitudinal modelling over 2-3 months (e.g. SSCC)
Other Bacteriological Markers

- **Serial Sputum Colony Counting (SSCC)**
  - Non-linear modelling of decline in CFU count over 56 days (Davies, 2006; Rustomjee, 2008)
  - Yet to be evaluated in the presence of long-term endpoint in a phase III trial

- **Time to culture conversion on liquid culture**
  - Generally recognised as best phase IIB endpoint (e.g. Diacon, 2009)
  - Yet to be evaluated in the presence of long-term endpoint in a phase III trial
Conclusions

- Caution required in selecting endpoints for Phase III trials

- Time-dependent marker preferable to fixed-time point for Phase II trials
  - Particularly in presence of HIV, with or without ART

- Important to measure evaluate putative surrogates in context of Phase III trials
  - OFLOTUB – end 2011
  - RIFAQUIN – early 2012
  - REMoxTB – end 2013
  - STREAM (MDR-TB) – early 2015
Conclusions

• If measured **before the end of treatment**, how can a marker **fully capture** the treatment effect?
  • Putative surrogate must at least be after end of intensive phase

• In the context of Phase II, what are the relative ‘costs’ associated with:
  • **False negative?** – Effective regimen rejected
  • **False positive?** – Ineffective regimen accepted
References


Burzykowski, T; Molenberghs, G & Buyse, ME. The Evaluation of Surrogate Endpoints *Springer,* 2005


