UCSF
University of California
San Francisco

Center for
Tuberculosis

UCSF
University of California
San Francisco

# Unfavorable outcomes, bacteriological relapse, and treatment success: Are we speaking the same language?

Patrick Phillips

# What proportion of patients are cured with the 6-month standard regimen for DS-TB?

MITT = Modified Intention-To-Treat; PP = Per Protocol

| | |
|---|---|
| 99% - SHRZ/HR | Fox, 1981 |
| 95-99% - SHRZ/HR(Z) | FDA guidance for pulmonary TB trials, 2013 |
| 96% - SHR | EA/BMRC Study R 1972 |
| 95.1% (PP) | RIFAQUIN, Jindani et al. 2014 |
| 92% | NIRT, Jawahar et al. 2013 |
| 92% (PP) | REMoxTB, Gillespie et al. 2014 |
| 88.7% (PP) | OFLOTUB Phase III, Merle et al. 2014 |
| 85.6% (MITT) | RIFAQUIN, Jindani et al. 2014 |
| 84% (MITT) | REMoxTB, Gillespie et al. 2014 |
| 82.8% (MITT) | OFLOTUB Phase III, Merle et al. 2014 |

- What do we mean by ‘cure’?
- When is it measured?
- What is the denominator (patient population)?
- How do we classify death or loss to follow-up?
- What about treatment changes for adverse events?



UCSF

FDA NEWS RELEASE

## FDA approves new drug for treatment-resistant forms of tuberculosis that affects the lungs

*"…Of the 107 patients who were evaluated six months after the end of therapy, 95 (89%) were **succ** exceeded the **his** treatment of exte*

https://www.fda.gov/news-events/ treatment-resistant-forms-tubercu

A Trial of a Shor

Andrew J. Nunn, M.Sc., Patrick P.J. Phillips, Ph.D van Deun, Ph.D., Phan-Thuong Dat, Ph.D., Ng

*"…**Favorable sta** participants in the those in the short adjustment for HI (P=0.02 for noninferiority)."*

ORIGINAL ARTICLE

# Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis

Stephen H. Gillespie, M.D., D.Sc., Angela M. Crook, Ph.D., Timothy D. McHugh, Ph.D., Carl M. Mendel, M.D., Sarah K. Meredith, M.B., B.S.

*"…in the per-protocol analysis, a **favorable outcome** was reported in fewer patients in the isoniazid group (85%) and the ethambutol group (80%) than in the control group (92%)… Results were consistent in the modified intention-to-treat analysis and all sensitivity analyses.*

## WHO updates its treatment guidelines for MDR/RR-TB

*The updated WHO treatment guidelines recommend that drug-resistant TB be treated with oral drugs only, including newer, more potent drugs with fewer side effects.*

**Table 2.2. Relative risk for (i) treatment failure or relapse and (ii) death (versus treatment success), 2018 IPD-MA for longer MDR-TB regimens and delamanid population)[22]**

| Treatment failure or relapse versus treatment success | | Death versus treatment success | |
|---|---|---|---|
| Number treated | Adjusted odds ratio (95% confidence limits) | Number treated | Adjusted odds ratio (95% confidence limits) |
| 3 143 | 0.3 (0.1–0.5) | 3 551 | 0.2 (0.1–0.3) |
| 1 391 | 0.3 (0.2–0.4) | 1 480 | 0.2 (0.2–0.3) |
| 1 216 | 0.3 (0.2–0.5) | 1 286 | 0.3 (0.2–0.3) |
| 991 | 0.3 (0.2–0.5) | 1 096 | 0.4 (0.3–0.6) |
| 5 483 | 0.6 (0.4–0.9) | 6 160 | 0.6 (0.5–0.8) |
| 1 163 | 0.4 (0.1–1.0) | 1 245 | 0.5 (0.1–1.7) |
| 289 | 1.1 (0.4–2.8)* | 290 | 1.2 (0.5–3.0)* |
| 1 248 | 2.7 (0.7–10.9) | 1 272 | 1.2 (0.1–15.7) |



UCSF

BRITISH MEDICAL JOURNAL

LONDON SATURDAY OCTOBER 30 1948

# STREPTOMYCIN TREATMENT OF PULMONARY TUBERCULOSIS

A MEDICAL RESEARCH COUNCIL INVESTIGATION

TABLE II.—*Assessment of Radiological Appearance at Six Months as Compared with Appearance on Admission*

| Radiological Assessment | Streptomycin Group | | Control Group | |
|---|---|---|---|---|
| Considerable improvement .. | 28 | *51%* | 4 | *8%* |
| Moderate or slight improvement | 10 | *18%* | 13 | *25%* |
| No material change .. .. | 2 | *4%* | 3 | *6%* |
| Moderate or slight deterioration | 5 | *9%* | 12 | *23%* |
| Considerable deterioration .. | 6 | *11%* | 6 | *11%* |
| Deaths .. .. .. .. | 4 | *7%* | 14 | *27%* |
| Total .. .. | 55 | *100%* | 52 | *100%* |

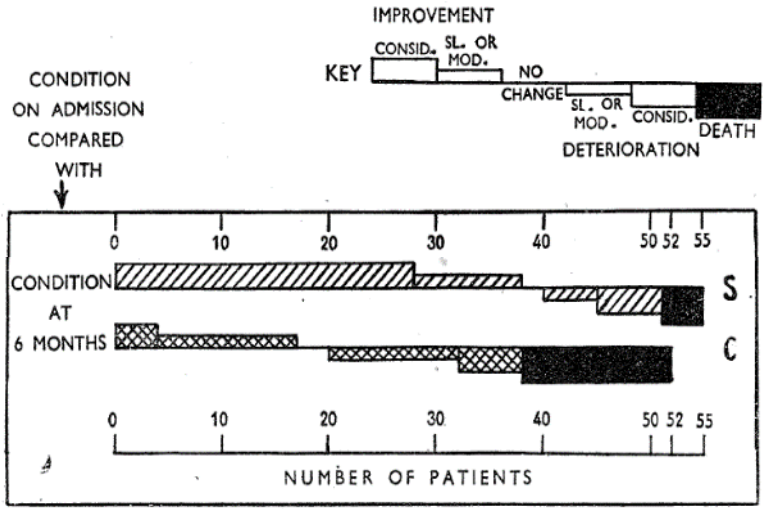



CHART II.—Condition on admission compared with condition at two, four, and six months (radiological assessment).



UCSF

BRITISH MEDICAL JOURNAL

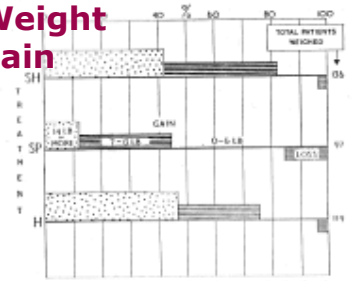
LONDON SATURDAY MARCH 7 1953

# ISONIAZID IN THE TREATMENT OF PULMONARY TUBERCULOSIS

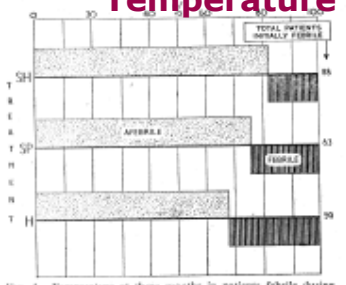
SECOND REPORT TO THE MEDICAL RESEARCH COUNCIL BY THEIR TUBERCULOSIS CHEMOTHERAPY TRIALS COMMITTEE*

- Trial comparing various combinations of Isoniazid, Streptomycin and PAS.
- No clear ‘primary endpoint’
- *“…It is concluded, judging solely from the results [on 10 endpoints] at three months, that streptomycin + isoniazid… is clinically the most effective of the treatments studied.”*

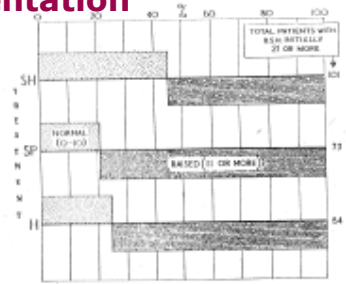
Weight gain

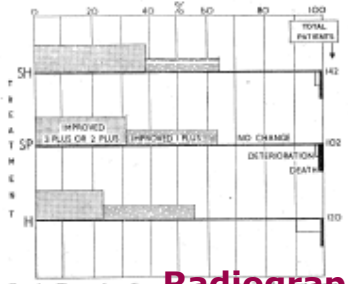


Temperature



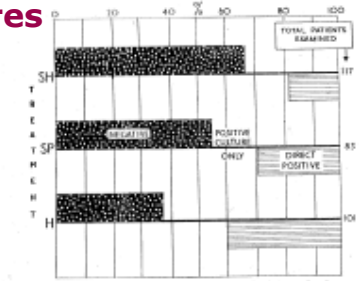
Sedimentation rate

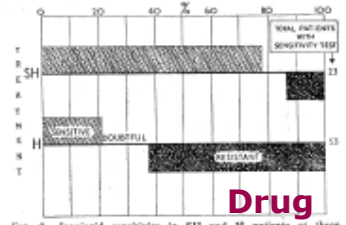




Radiographic changes

Cultures





Drug resistance

# Rifapentine FDA NDA

## Details from Statistical Review of 2-year follow-up, 2000

# Rifapentine FDA NDA

Details from Statistical Review of 2-year follow-up, 2000

- The FDA approved rifapentine, but the statistical reviewer noted the following caution:

  - *"It might be in the patients' best interest to add a statement to the proposed label cautioning that relapse rates could actually be much higher than they appear due to the fact that* ***we don't know what happened to almost a third of the patients who converted****."*

# Sensitivity analyses

## RIFAQUIN

5 Supplementary Tables

| | Analysis | Control N (%) unfavourab / assessable |
|---|---|---|
| | *Per Protocol Analysis (PP)* | |
| 1. | Primary PP analysis as in Table 2 | 8 (4.9%) / 163 |
| 2. | Primary unadjusted PP analysis | |
| 3. | Classifying all reinfections as unfavourable | 11 (6.6%) / 166 |
| 4. | Classifying all deaths as unfavourable | 13 (7.7%) / 168 |
| 5. | Subgroup of HIV negative patients | 6 (5.2%) / 115 |
| 6. | Subgroup of HIV positive patients | 2 (4.2%) /48 |
| | | p-value for |
| | *Modified Intention to Treat Analysis (mITT)* | |
| 7. | Primary mITT analysis as in Table 2 | 27 (14.4%) / 188 |
| 8. | Primary unadjusted mITT analysis | |
| 9. | Alternative mITT analysis: mITT model 2 | 23 (12.2%) / 188 |
| 10. | Strict mITT – all post-randomisation exclusions as unfavourable | 79 (32.9%) / 240 |
| 11. | Classifying all reinfections as unfavourable | 30 (15.7%) / 191 |
| 12. | Classifying all deaths as unfavourable | 31 (16.1%) / 192 |
| 13. | Subgroup of HIV negative patients | 20 (14.9%) / 134 |
| 14. | Subgroup of HIV positive patients | 7 (13.0%) / 54 |
| | | p-value for ir |

Table S1. Summary of sensitivity analyses of the primary efficacy

## REMoxTB

Table S3A Summary of sensitivity analyses

| Analysis | Control arm N unfavourable / N assessable (%) | Is N unfavourable / N assessable (%) |
|---|---|---|
| **Per Protocol Analyses (PP)** | | |
| Primary analysis: adjusted for stratification factors | 43/510 (8%) | 78/514 (15%) |
| Adjusted for additional covariates | | |
| Unadjusted | | |
| All deaths as unfavourable | 48/515 (9%) | 84/520 (16%) |
| Primary endpoint based on only LJ results | 43/501 (9%) | 78/500 (16%) |
| Primary endpoint based only on MGIT result | 65/498 (13%) | 98/498 (20%) |
| Status at end of active treatment phase (EOT) | 16/474 (3%) | 17/489 (3%) |
| 18m status in those favourable at EOT | 21/458 (5%) | 63/472 (13%) |
| 12m status in those favourable at EOT | 27/435 (6%) | 64/459 (14%) |
| **Modified Intention to Treat (MITT)** | | |
| Primary analysis: adjusted for stratification factors | 87/555 (16%) | 132/568 (23%) |
| Adjusted for additional covariates | | |
| Unadjusted | | |
| Reinfections as unfavourable | 97/565 (17%) | 145/581 (25%) |
| Pinetown and Mexico as unfavourable | 97/565 (17%) | 140/576 (24%) |
| All deaths as unfavourable | 92/560 (16%) | 135/571 (24%) |
| Secondary bacteriological endpoint | 75/555 (14%) | 127/567 (22%) |
| Primary endpoint based on only LJ results | 87/546 (16%) | 132/554 (24%) |
| Primary endpoint based only on MGIT result | 109/543 (20%) | 153/553 (28%) |
| Status at end of active treatment phase | 16/485 (3%) | 19/503 (4%) |
| 18m status in those favourable at EOT | 31/469 (7%) | 75/484 (15%) |
| 12m status in those favourable at EOT | 37/446 (8%) | 76/471 (16%) |
| **All randomised patients** | | |
| Missing outcome as unfavourable | 172/640 (27%) | 219/655 (33%) |
| Missing outcome as favourable | 87/640 (14%) | 132/655 (20%) |
| Missing outcomes as last observation carried forward | 119/640 (19%) | 165/655 (25%) |
| **All randomised patients excluding late screening failures** | | |

## STREAM

**Table S5. Summary of sensitivity analyses of the p**

Sensitivity analysis are as follows:

A. Participants on the Long regimen with dura
B. Primary outcome adjusted for randomizati
C. Primary outcome in additional analysis pop
D. Primary outcome reclassification - classifica
E. Primary outcome reclassification - classifica

| | N |
|---|---|
| **Primary analysis (mITT population)** | |
| **Primary analysis (PP population)** | |
| | |
| Sensitivity analysis A. (mITT population) | |
| Sensitivity analysis A. (PP population) | |
| | |
| Sensitivity analysis B. (mITT population) | |
| Sensitivity analysis B. (PP population) | |
| | |
| Sensitivity analysis C. (ITT population) | |
| Sensitivity analysis C. (Safety population) | |
| | |
| Sensitivity analysis D. (mITT population) | |
| Sensitivity analysis D. (PP population) | |
| | |
| Sensitivity analysis E. (mITT population) | |
| Sensitivity analysis E. (PP population) | |

# What are the implications of this language disconnect?

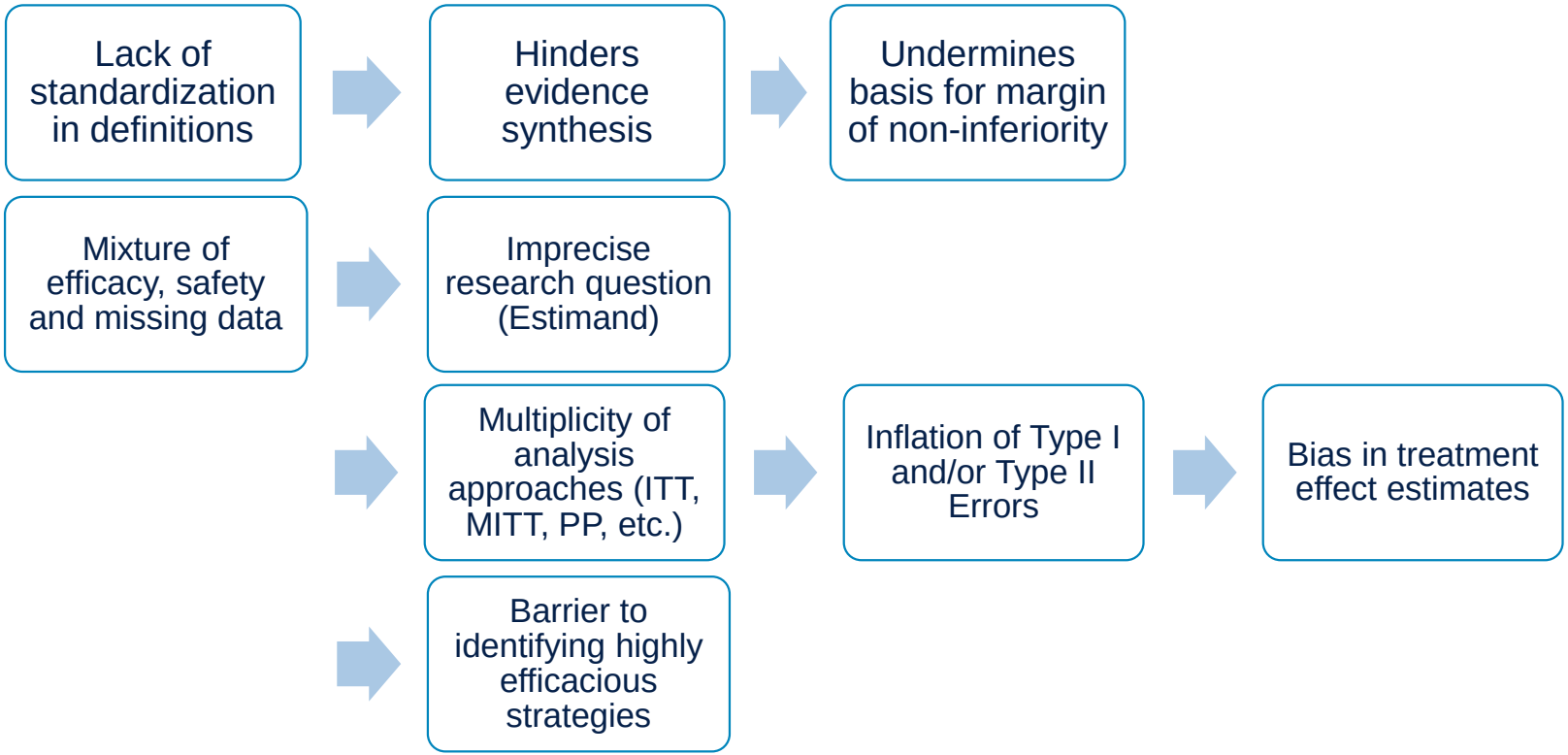
Lack of standardization in definitions

Mixture of efficacy, safety and missing data

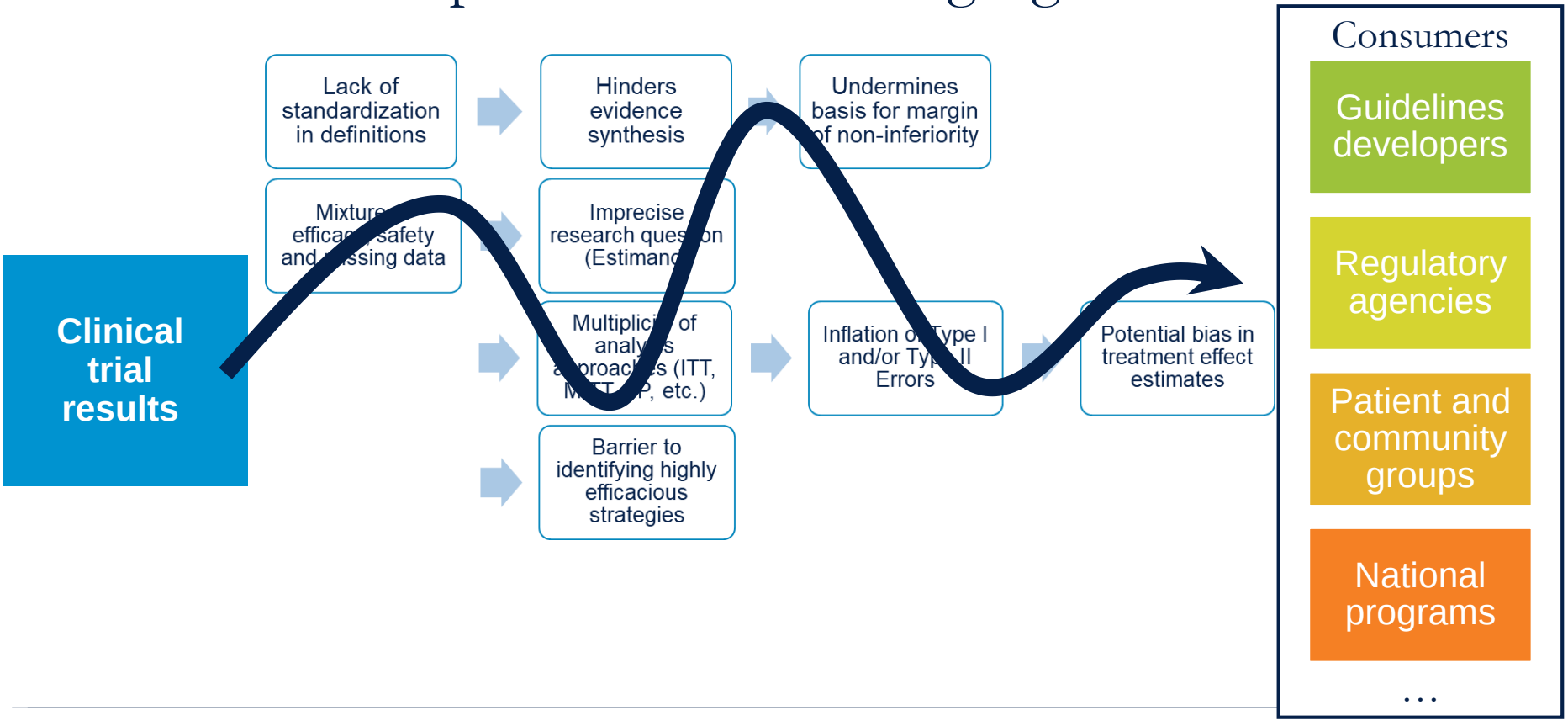
→ Multiplicity of analysis approaches (ITT, MITT, PP, etc.)

→ Barrier to identifying highly efficacious strategies

# What are the implications of this language disconnect?



# What are the implications of this language disconnect?



# REMoxTB

| | | |
|---|---|---|
| Were considered not able to be assessed — no. | | |
| Had reinfection with a different strain | 1 | 7 |
| Had a negative culture at 76 weeks but lost to follow-up thereafter | 5 | 1 |
| Were included in primary outcome analysis — no. | 124 | 24 |
| **Outcome** | | |
| Attained favorable status — no. (%)† | 99 (79.8) | 193 ( |
| Had an unfavorable outcome — no. (%) | 25 (20.2) | 52 (2 |
| Determined on the basis of bacteriologic findings‡ | | |
| Had no negative cultures§ | 1 | 5 |
| Had bacteriologic reversion during treatment period¶ | 4 | 1 |
| Had bacteriologic relapse after treatment period and started ≥2 additional drug therapies‖ | 0 | 7 |
| Had positive culture at last assessment** | 2 | 1 |
| Determined on the basis of criteria other than bacteriologic findings | | |
| Had negative culture at last assessment but died during the treatment or follow-up period | 5 | 9 |
| Had treatment extended or changed after adverse event | 3 | 4 |
| Started ≥2 additional drug therapies owing to decision by the investigator†† | 3 | 2 |
| Withdrew consent for treatment, was given a different regimen, or was lost to follow-up before 76 weeks | 4 | 8 |
| Had treatment extended or changed after poor adherence or loss to follow-up | 0 | 2 |
| Had negative culture at last assessment but was lost to follow-up before 76 weeks | 3 | 1 |

# STREAM

| | | | |
|---|---|---|---|
| **Favorable outcome — no. (%)** | | | |
| Patients with outcome | 467 (92) | 436 (85) | 419 ( |
| Culture-negative status at 18 mo | 409 (80) | 389 (76) | 367 ( |
| Unable to produce sputum | 0 | 2 (<1) | 0 |
| Unable to produce sputum at 18 mo but culture-negative status earlier | 49 (10) | 31 (6) | 35 ( |
| Missing data on L–J culture at 18 mo and MGIT negative | 9 (2) | 14 (3) | 17 ( |
| **Unfavorable outcome — no. (%)†** | | | |
| Patients with outcome | 43 (8) | 78 (15) | 105 ( |
| 6-Mo treatment phase | | | |
| Nonviolent death | 5 (1) | 6 (1) | 7 ( |
| Treatment failure‡ | | | |
| Culture-confirmed | 3 (1) | 4 (1) | 1 ( |
| Not culture-confirmed | 4 (1) | 1 (<1) | 4 ( |
| Adverse reaction | NA | NA | N |
| Withdrawal of consent | NA | NA | N |
| Relocation | NA | NA | N |
| Other investigator decision | NA | NA | N |
| No completion of treatment | NA | NA | N |
| Follow-up | | | |
| Relapse after culture-negative status | 12 (2) | 46 (9) | 64 ( |
| Retreated for tuberculosis | 14 (3) | 17 (3) | 27 ( |
| Death from tuberculosis or respiratory distress | 2 (<1) | 0 | 0 |
| No culture-negative status | | | |
| Ever | 1 (<1) | 1 (<1) | 0 |
| At last visit | 2 (<1) | 3 (1) | 2 ( |

# Delamanid C213 trial

### 5.2 Treatment Success or Failure at Month 30

**Table S7 Treatment Success or Failure at Month 30 (MITT),**

| Endpoint – no. (%) | Delamanid + optim background regi (N=226) |
|---|---|
| Treatment Success* | 173 (76·5) |
| Treatment Failure | 53 (23·5) |
| Achieved 6-month SCC, then died | 6 (2·7) |
| Achieved 6-month SCC, discontinued and alive/unknown at Month 30 | 10 (4·4) |
| Achieved 6-month SCC, discontinued, then died | 0 (0,0) |
| Achieved 6-month SCC, then had positive culture | 9 (4·0) |
| Died before 6 months | 1 (0·4) |
| Failed to achieve 6-month SCC and died after 6 months | 3 (1·3) |
| Discontinued before 6 months and alive/unknown at Month 30 | 9 (4·0) |
| Discontinued before 6 months then died | 1 (0·4) |
| Failed to achieve 6-month SCC, discontinued, and alive/unknown at 30 months | 2 (0·9) |
| Failed to achieve 6-month SCC, discontinued, then died | 1 (0·4) |
| Failed to achieve 6-month SCC and completed 30 months | 11 (4·9) |

***SCC by 6 months, completed trial to 30 months with sustained**
**MITT=modified intent-to-treat, SCC=sputum culture conversi**

### 5.3 End of Treatment Outcomes

**Table S8 Treatment Outcome at End of Treatment with OBR (MITT), MGIT**

| Endpoint – no. (%) | Delamanid + optim background regi (N=224) |
|---|---|
| Favourable Outcome | 182 (81·3) |
| Cured | 173 (77·2) |
| Completed | 9 (4·0) |
| Unfavourable Outcome | 42 (18·8) |
| Failed | 11 (4·9) |
| Defaulted | 22 (9·8) |
| Died | 9 (4·0) |

Lack of standardization → Challenges for evidence synthesis

**Completed treatment, culture converted and culture negative at end of follow-up**

| | REMoxTB (MITT) | REMoxTB (PP) | STREAM (MITT) | Delamanid C213 30m outcomes | WHO DR-TB Guidelines* |
|---|---|---|---|---|---|
| Favorable / Success | | | | | |
| Unfavorable / Failure | | | | | |
| Excluded | | | | | |

* ‘Treatment failure or relapse versus treatment success’ analysis in guidelines



UCSF

Lack of standardization → Challenges for evidence synthesis

## Lost to follow-up after treatment completion

| | REMoxTB (MITT) | REMoxTB (PP) | STREAM (MITT) | Delamanid C213 30m outcomes | WHO DR-TB Guidelines* |
|---|---|---|---|---|---|
| Favorable / Success | | | | | |
| Unfavorable / Failure | | | If before 15 months | | |
| Excluded | | | If after 15 months | | |

* ‘Treatment failure or relapse versus treatment success’ analysis in guidelines



UCSF

Lack of standardization → Challenges for evidence synthesis

## Lost to follow-up during treatment

| | REMoxTB (MITT) | REMoxTB (PP) | STREAM (MITT) | Delamanid C213 30m outcomes | WHO DR-TB Guidelines* |
|---|---|---|---|---|---|
| Favorable / Success | | | | | |
| Unfavorable / Failure | | | | | |
| Excluded | | | | | |

* ‘Treatment failure or relapse versus treatment success’ analysis in guidelines



UCSF

Lack of standardization → Challenges for evidence synthesis

## Change of two drugs in background regimen for adverse drug reaction

| | REMoxTB (MITT) | REMoxTB (PP) | STREAM (MITT) | Delamanid C213 30m outcomes | WHO DR-TB Guidelines* |
|---|---|---|---|---|---|
| Favorable / Success | | | | | |
| Unfavorable / Failure | | | | | |
| Excluded | | | | | |

* ‘Treatment failure or relapse versus treatment success’ analysis in guidelines



UCSF

Lack of standardization → Challenges for evidence synthesis

## Death after treatment completion

| | REMoxTB (MITT) | REMoxTB (PP) | STREAM (MITT) | Delamanid C213 30m outcomes | WHO DR-TB Guidelines* |
|---|---|---|---|---|---|
| Favorable / Success | | | | | ? |
| Unfavorable / Failure | TB-related | | | | |
| Excluded | Not TB-related | | | | ? |

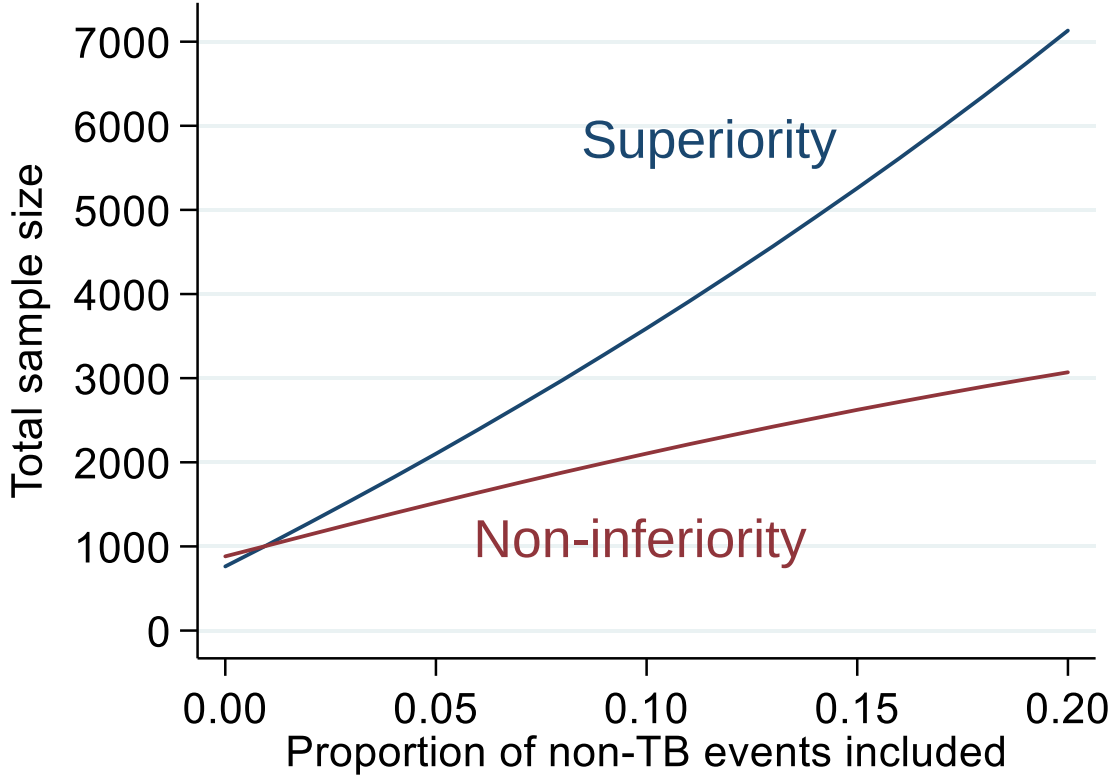
* ‘Treatment failure or relapse versus treatment success’ analysis in guidelines



UCSF

Mixture of efficacy, safety and missing data

Barrier to identifying highly efficacious regimens




**Assumptions**

- True relapse rate in control: 5%
- Power: 90%
- Superiority: Power to show reduction to 1%
- Non-inferiority: Margin of 5%
- Non-TB events independent of treatment and TB events



UCSF

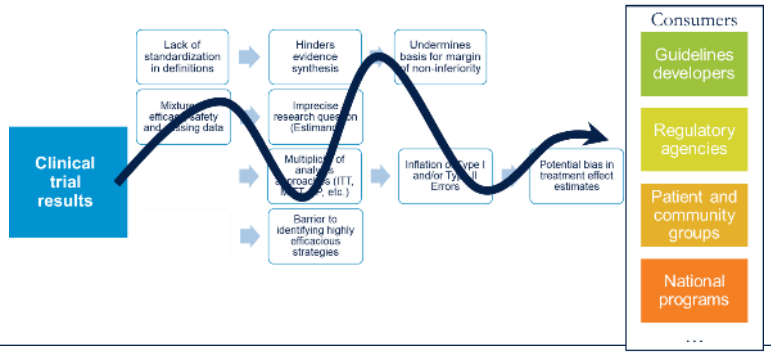
# Conclusions

- Lack of standardization of primary endpoint definitions hinders evidence synthesis
- Conflation of efficacy, safety and missing data into one dichotomous endpoint inflates Type I and Type II errors and introduces bias in treatment effect estimates
- Inclusion of non-TB events is a barrier to identification of highly efficacious treatment strategies

# Conclusions

- Catch-all endpoint definitions address <u>imprecise research questions and treatment effects</u>

  → Greater separation between clinical trial results and consumers of those results



# What should we do about it?

- Specification of appropriate estimand(s) for TB trials
  - Estimand = Precise statement of treatment effect estimate of interest
  - *Can we define estimand(s) that meet requirements for regulators, guidelines developers, patients and community groups?*
- Development of appropriate methods for analysis that yield unbiased estimates in the presence of missing data
- Greater standardization across sponsors

# A better framework for linking research questions with endpoint and analyses

## ICH E9 (R1) Addendum, draft 2014

*Mallinckrodt CH, Bell J, Liu G, Ratitch B, O'Kelly M, Lipkovich I, et al. Aligning Estimators With Estimands in Clinical Trials: Putting the ICH E9(R1) Guidelines Into Practice. Ther Innov Regul Sci. 2019*

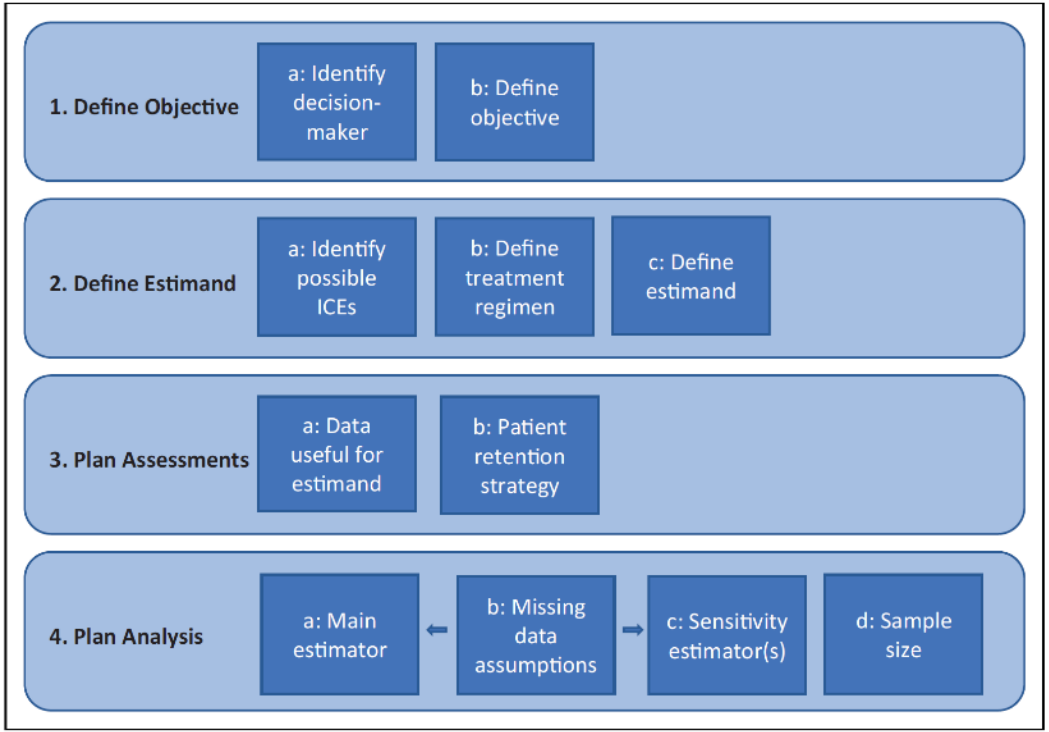



**Figure 1.** Study development process chart.

Center for
Tuberculosis

UCSF

University of California
San Francisco

tb.ucsf.edu