Endpoints that may correlate with cure & validation of biomarkers

Gerry Davies

Reader in Infection Pharmacology
Institutes of Infection & Global Health and Translational Medicine

UNIVERSITY OF LIVERPOOL

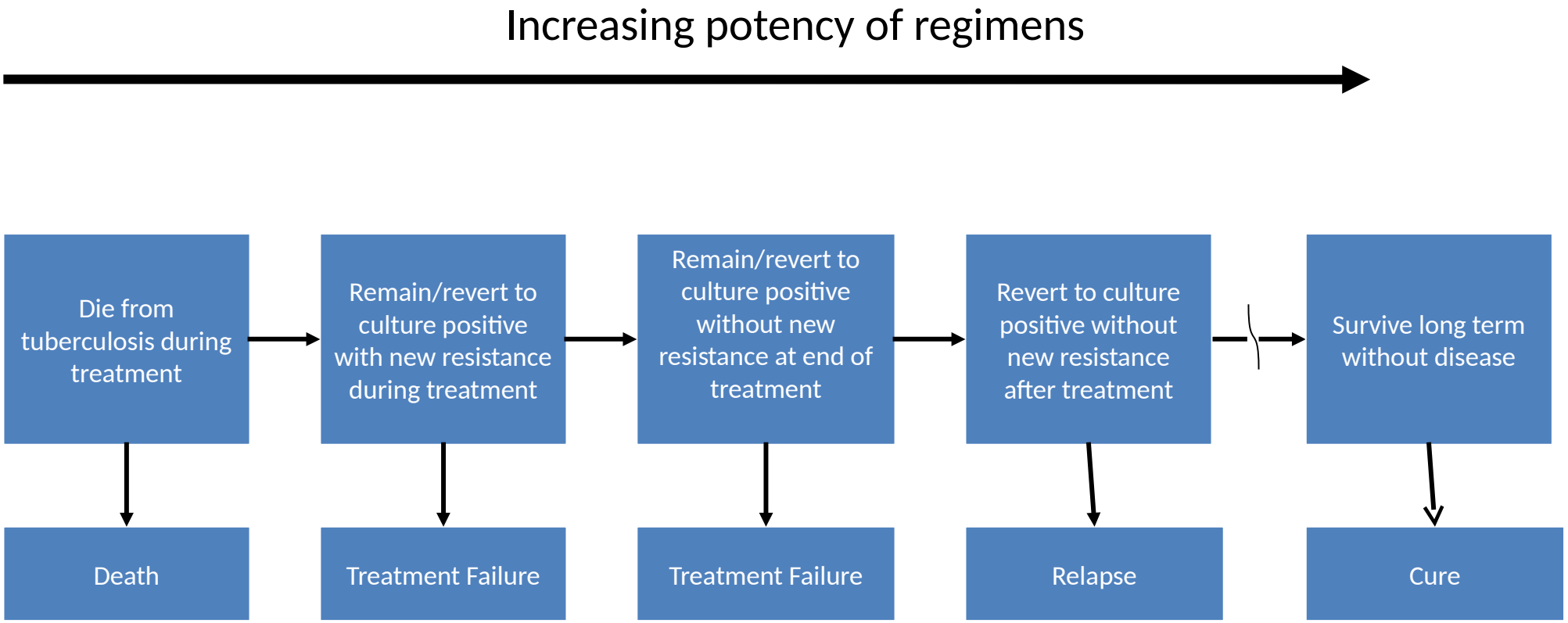
EMA Workshop on Update of TB Guidelines London 25th November 2016

Overview

UNIVERSITY OF LIVERPOOL

- How are endpoints be related to each other ?
- What data are available to us ?
- What do we mean by surrogacy ?
- What does the current evidence show ?
- What do we mean by validation ?
- Which endpoints should we use in the future ?

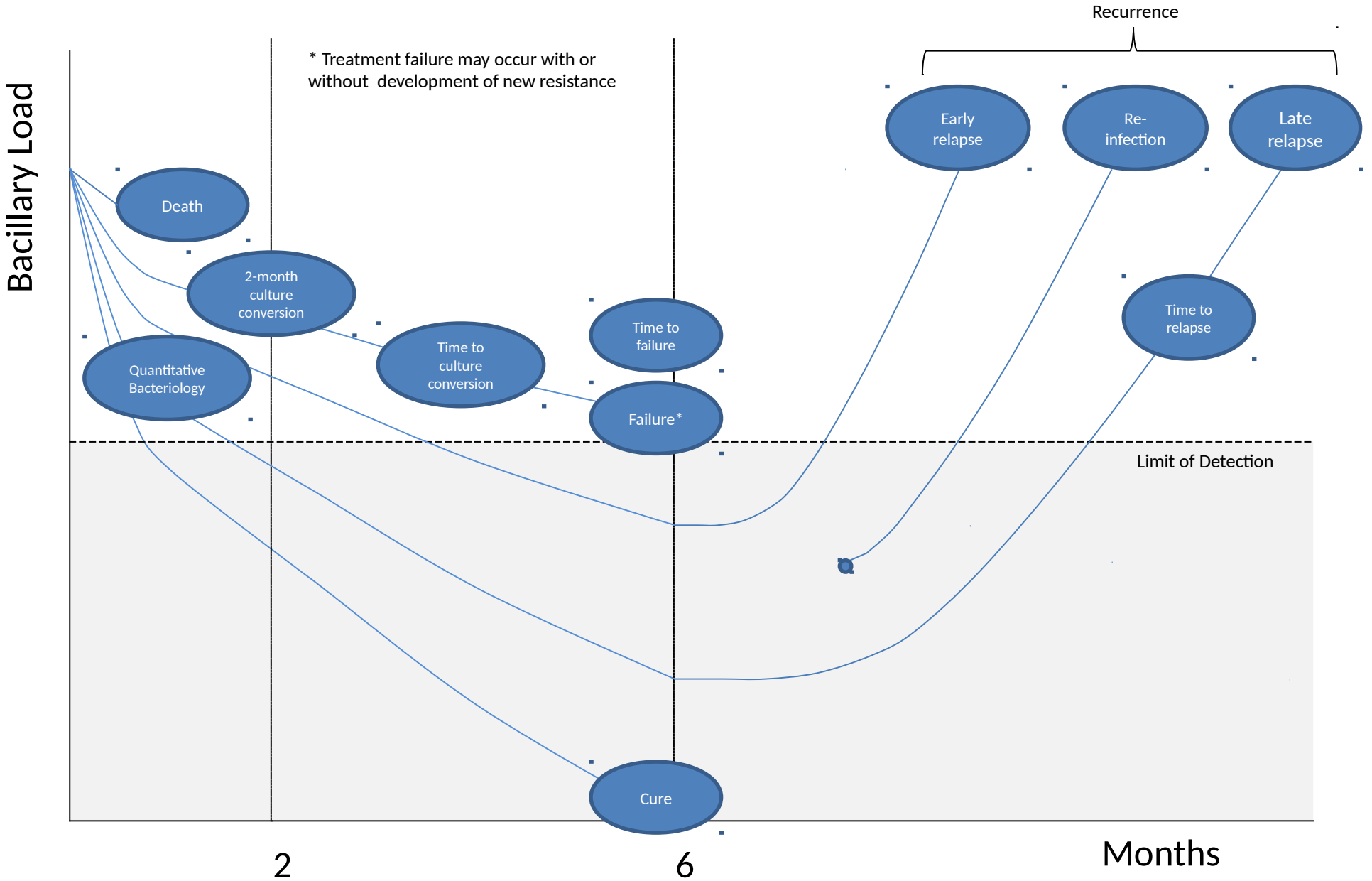
Causal linkage of efficacy endpoints



"Bactericidal activity"

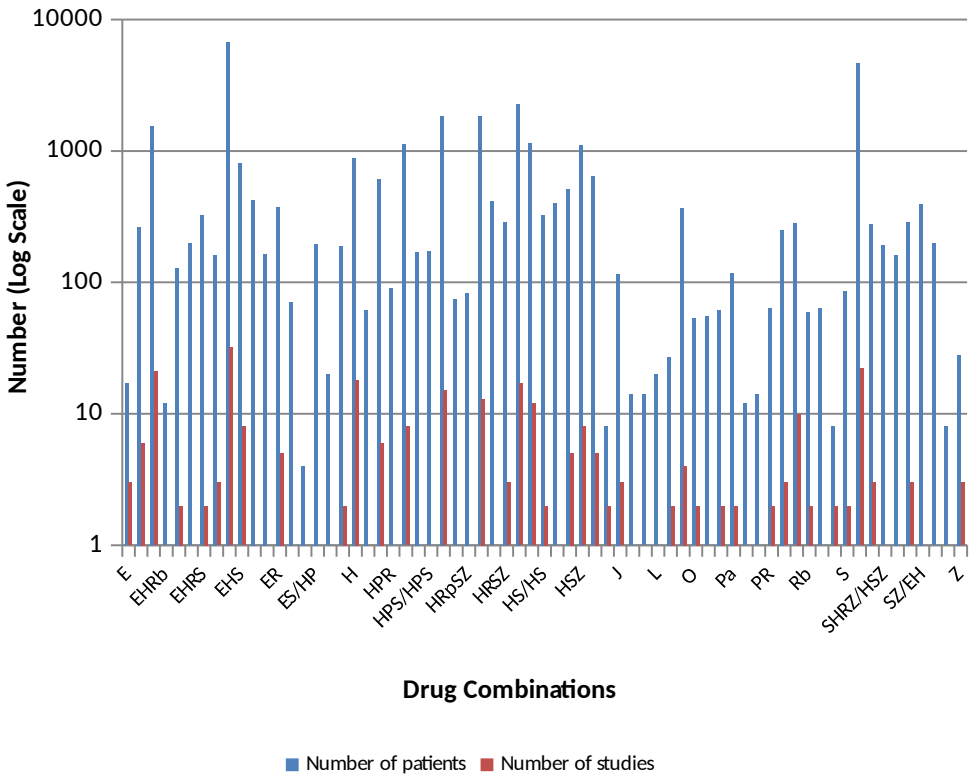
"Sterilizing activity"

Quantitative basis of efficacy endpoints



Phase II Systematic Reviews

N=37,173 67 drugs/combinations



133 trials with Phase IIA/B Outcomes

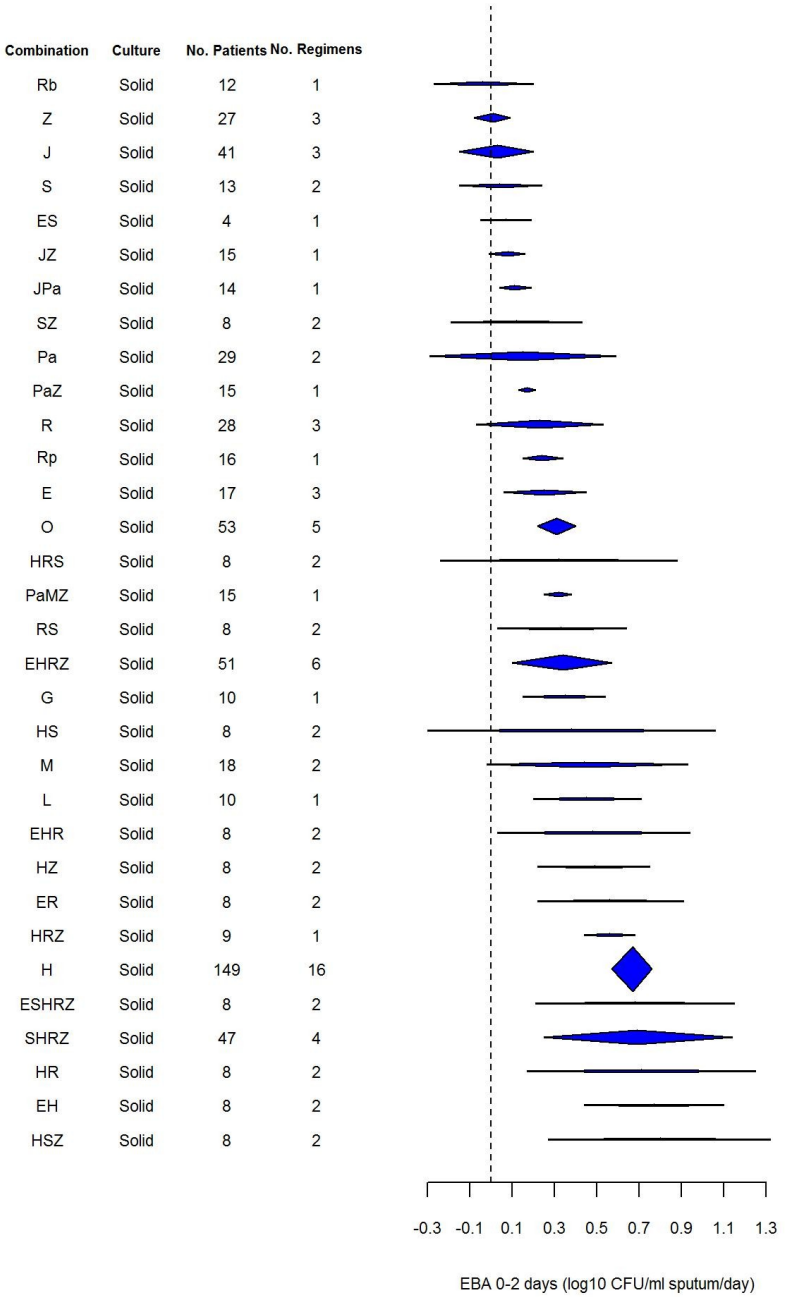
96 Phase III trials with intermediate outcomes

EBA_{0-2} and 8w CC most commonly reported endpoints

Inconsistent reporting of other EBA endpoints (EBA_{0-7}, EBA_{0-14}) and alternative approaches

Only 3 regimens with EBA over >2 days and 8w CC data

EBA0-2



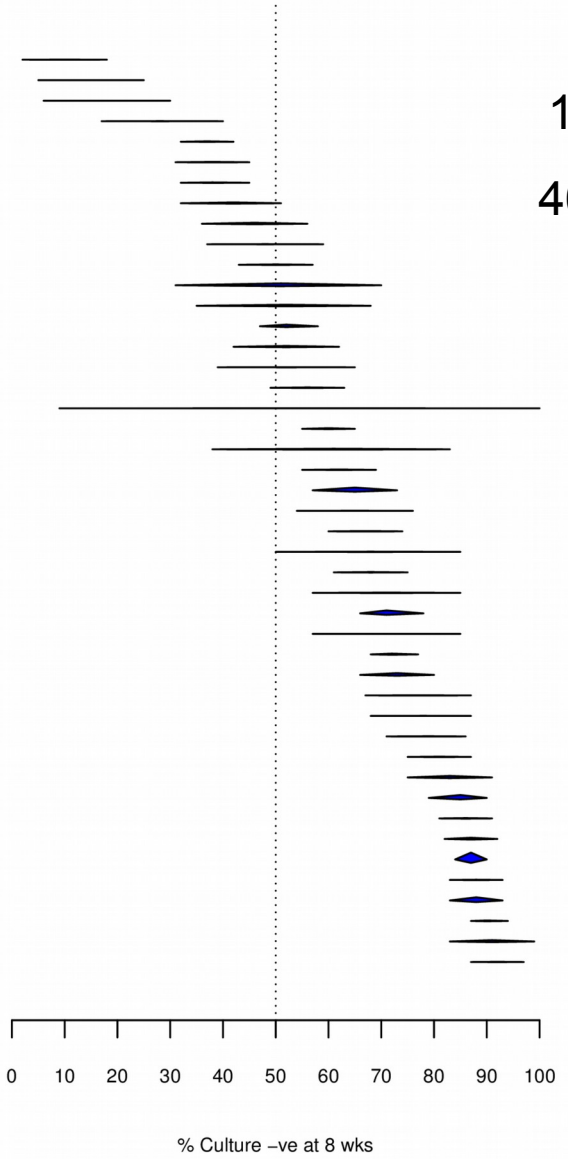
N=681

24 studies

141 regimens

8w culture conversion

Combination	Culture	No. Patients	No. Regimens
P	Solid	58	1
S	Solid	54	1
PS	Solid	37	1
HRpSZ	Solid	83	2
SZ/EH	Solid	214	2
ES/HP	Solid	195	1
SZ/HZ	Solid	198	1
HP	Solid	424	5
HT	Solid	522	5
SHP/HP	Solid	83	1
SZ	Solid	206	3
SHP	Solid	969	7
H	Solid	533	6
SH	Solid	849	8
HST	Solid	442	5
OHRZ	Solid	53	1
EHS/EH	Solid	214	2
EH	Solid	38	2
HS/HS	Solid	327	2
HSZ	Solid	319	3
HPS/HPS	Solid	171	1
HR	Solid	1415	10
R	Solid	77	2
SHRZ/HT	Solid	162	1
ER	Solid	135	2
SHRZ/HSZ	Solid	191	1
EHS	Solid	157	3
HRS	Solid	1500	13
PR	Solid	42	1
HRS/HR	Solid	413	4
EHR	Solid	919	13
HOR	Solid	62	1
ERZ	Solid	71	2
HPR	Solid	120	2
EHRS	Solid	171	1
GHRZ	Solid	793	2
HRZ	Solid	1492	15
EHRpZ	Solid	198	1
MHRZ	Solid	625	3
SHRZ	Solid	3479	25
EHRSZ	Solid	153	1
HRZE	Solid	1618	8
SHRZ/HR	Solid	278	4
EMRZ	Solid	701	2
EHRbZ	Solid	107	1

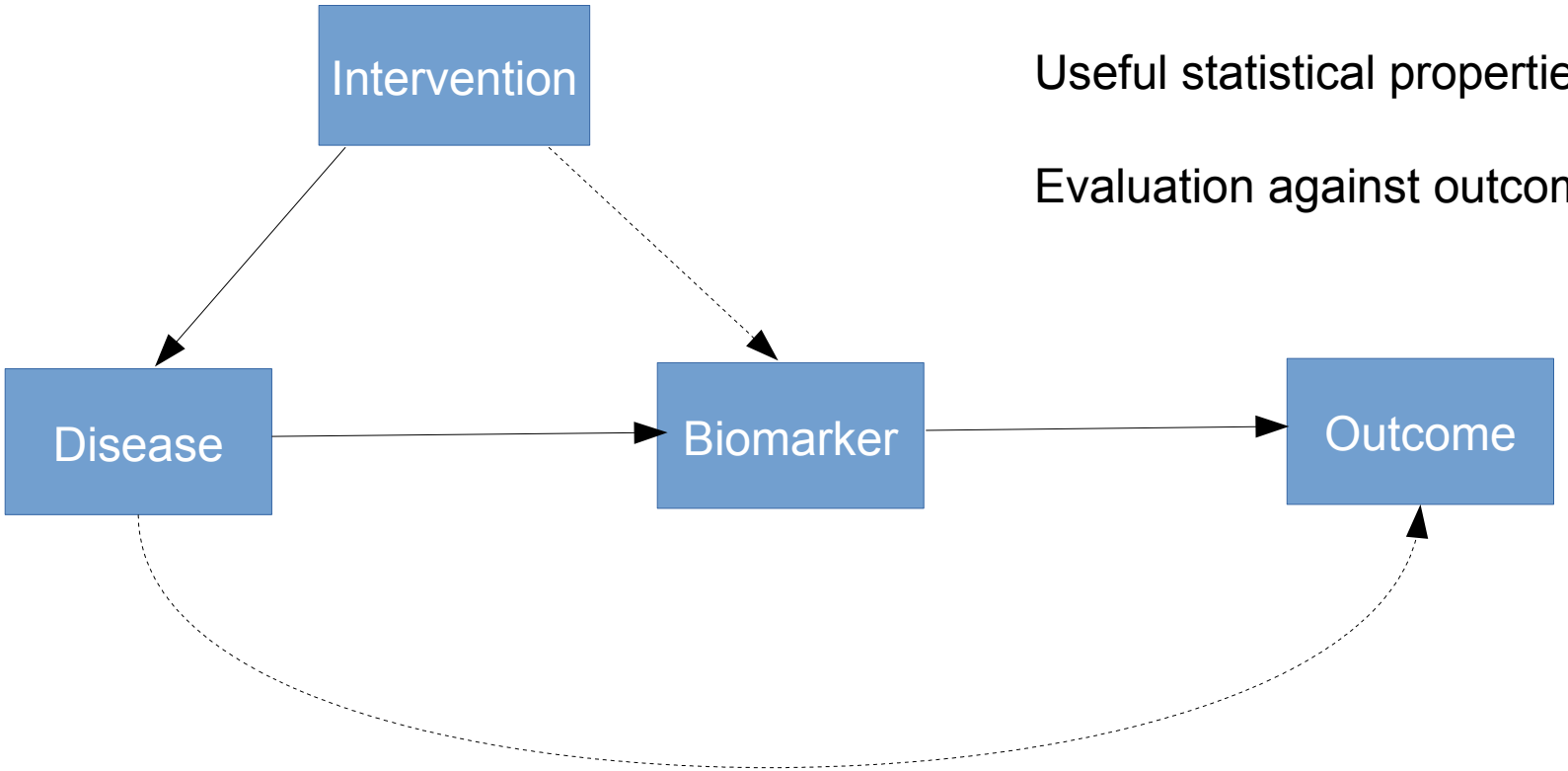


N=18276

107 studies

46 regimens

Concept of surrogacy



Biologically plausible

Useful statistical properties

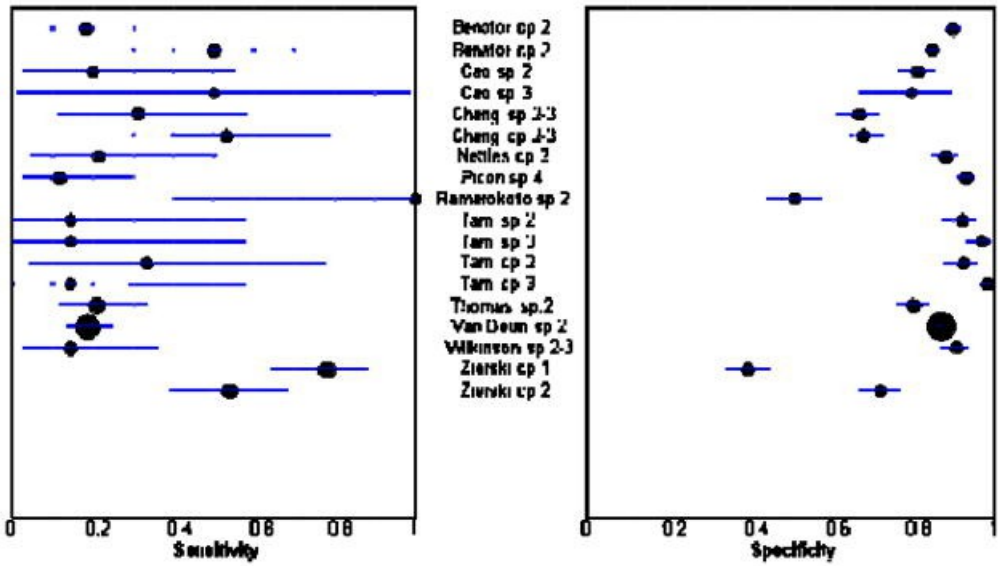
Evaluation against outcome

Surrogate endpoint

UNIVERSITY OF LIVERPOOL

- A biomarker that can replace the reference endpoint
- Trial level : The ability to capture treatment effect on the definitive endpoint (ρ_z ,RE,PTE,R^2_{trial})
- Individual level : The ability to predict an individual's definitive outcome (PPV, NPV, ROC, $R^2_{individual}$)
- These levels are theoretically independent (Simpson's paradox) though in practice often go together

8w CC : individual level



	Studies (n)	Sample size (N)	Hierarchical regression model		Odds ratio (95% CI)	PPV* (95% CI)	NPV* (95% CI)
			Sensitivity (95% CI)	Specificity (95% CI)			
Relapse							
Culture	4	1298	40% (25–56%)	85% (77–91%)	3·8 (2·2–6·8)	18% (14–21%)	95% (95–96%)
Smear	6	9848	24% (12–42%)	83% (72–90%)	1·5 (1·1–2·2)	10% (8–12%)	93% (93–94%)
Failure							
Smear	7	20 062	57% (41–73%)	81% (72–87%)	5·8 (4·3–7·8)	9% (9–10%)	98% (98–98%)

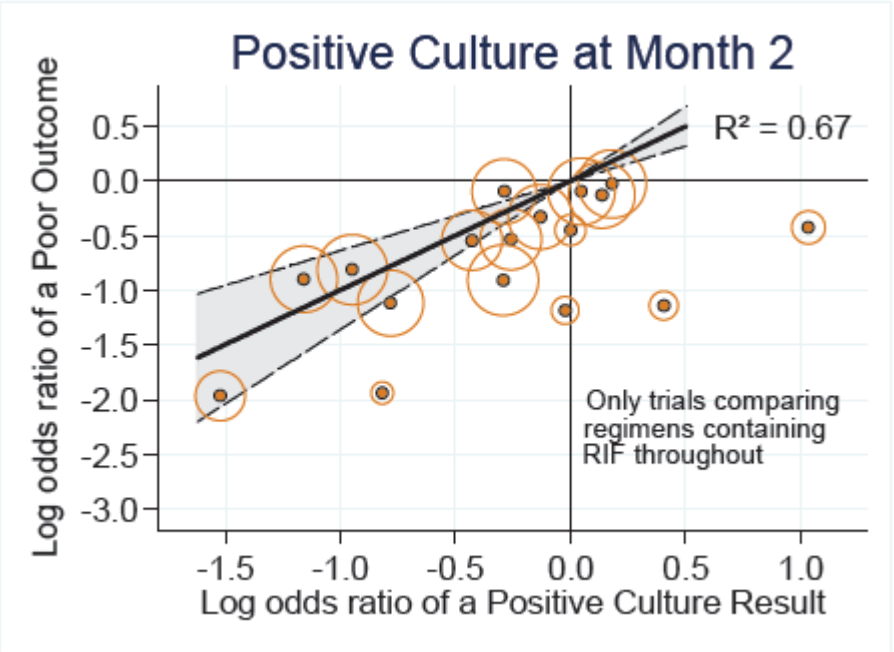
*Ability of smear to predict poor outcomes, assuming 7% risk of relapse and 3% risk of failure. NPV=negative predictive value; PPV=positive predictive value.

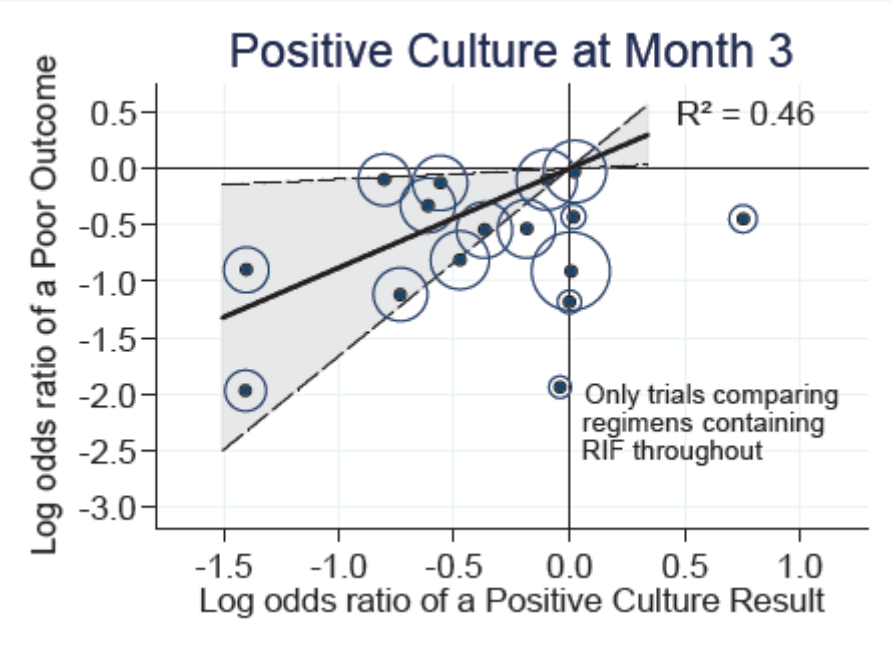
Table 5: Pooled summary estimates for relapse or failure for patients with a positive sputum culture or smear at 2 months

Horne DJ Lancet ID 2010 10:387-94

8w CC : trial level

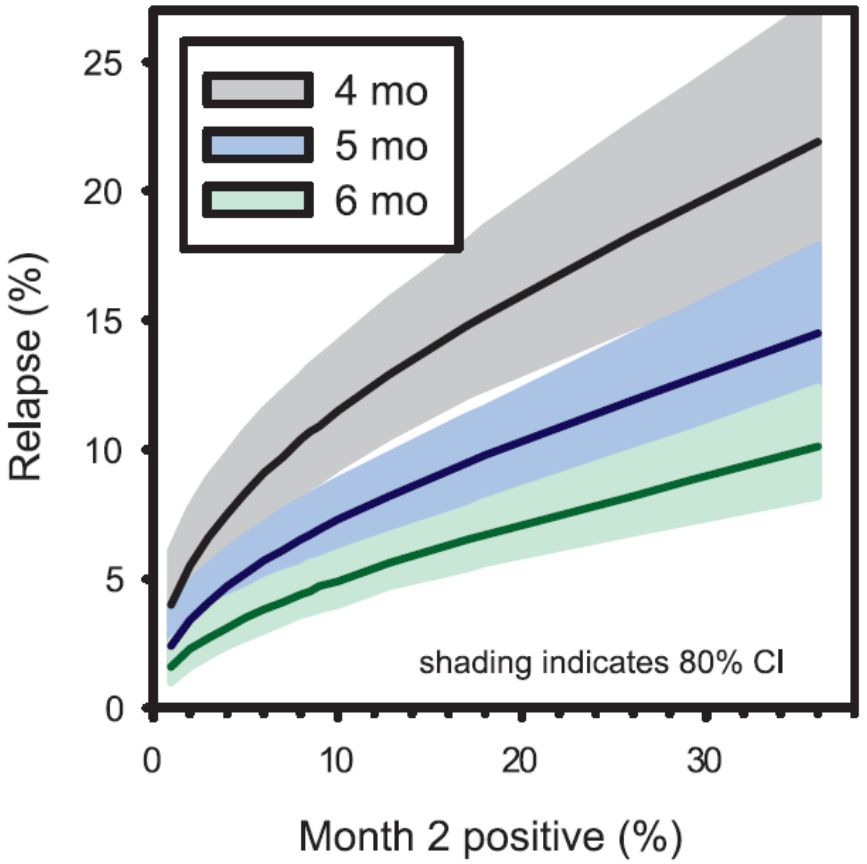
15 BMRC trials 6974 participants 37 treatment comparisons

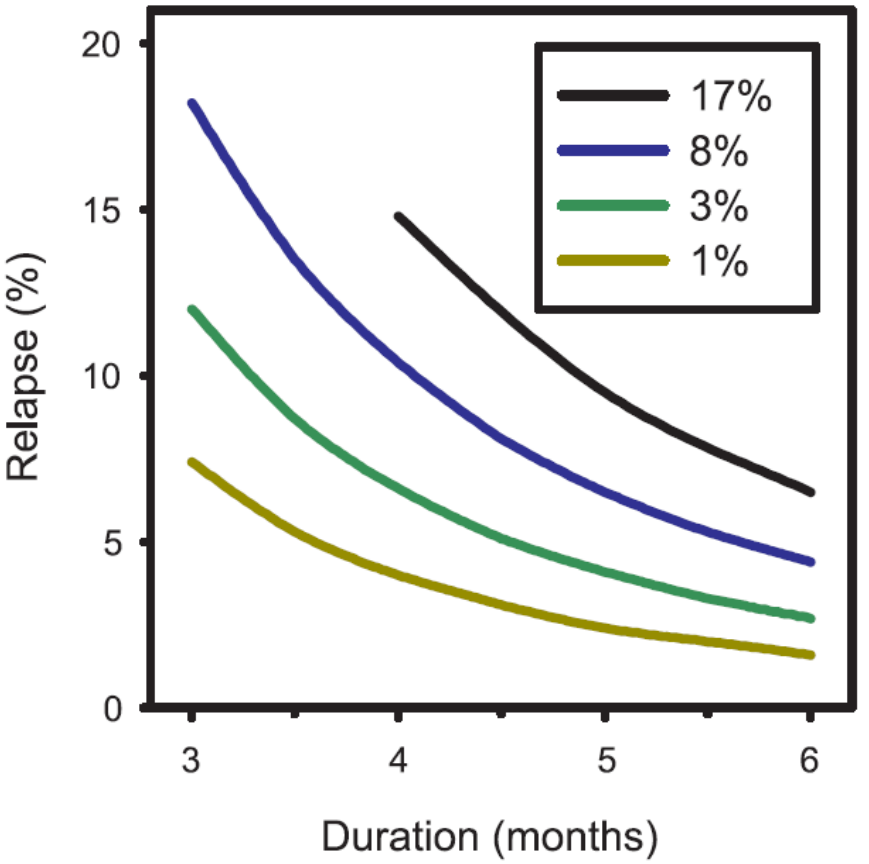




Phillips P and Fielding K 2008 IUATLD Conference Paris

8w CC : Predicting duration





Wallis RS PloS ONE 8(8) : e71116

Evaluation, Validation, Qualification

UNIVERSITY OF LIVERPOOL

- Prentice criterion
- R^2_{trial} “sufficiently close to” 1
- Reasonably likely to predict clinical benefit
- Widespread agreement about the significance of the test results
- Can be relied upon to have a specific use and interpretable meaning

Chau CH 2008 Clin Cancer Res 14: 5967-76

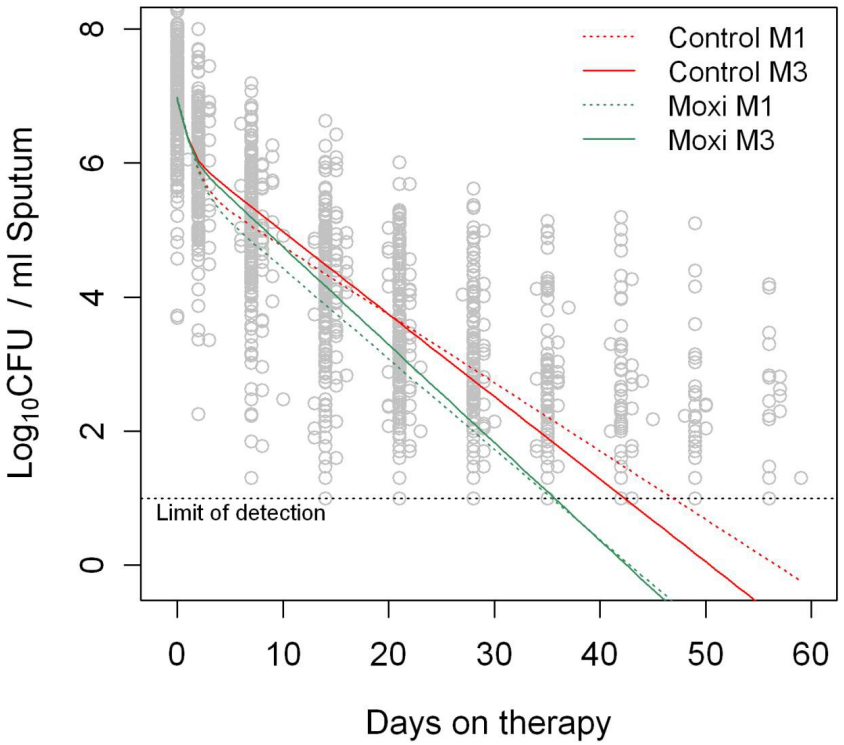
Longitudinal or time-to-event endpoints

UNIVERSITY OF LIVERPOOL

- Independent of sampling timepoints
- No need for future ad hoc re-evaluation
- Unrestricted scale of measurement
- Greater statistical power
- Well-adapted to cumulative meta-analysis
- Little trial level evaluation due to design and reporting
- Model choice, LOD, missing data

Longitudinal endpoints

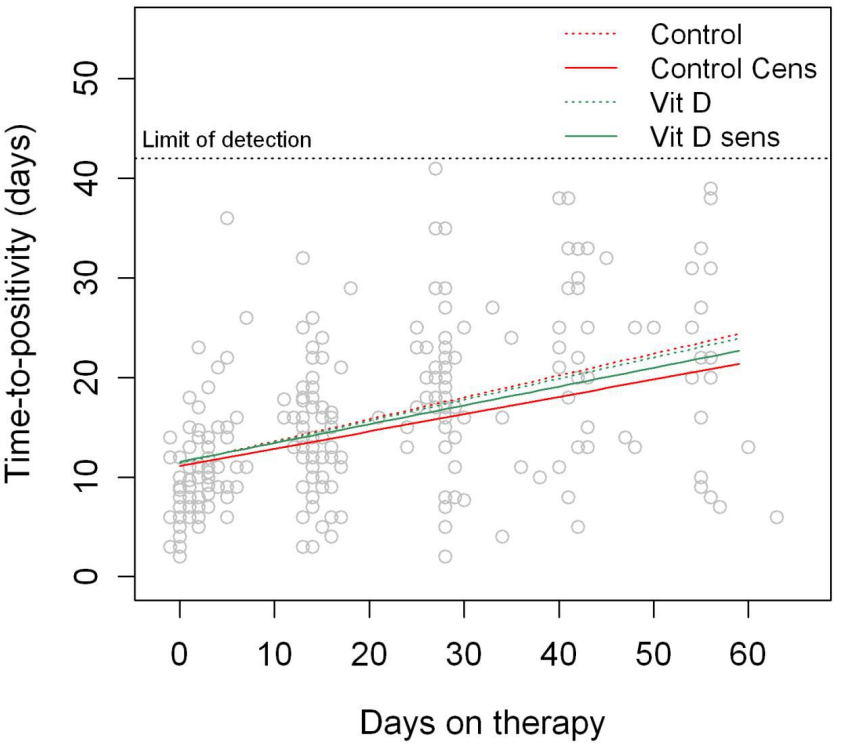
Oflotub



M3 method in NONMEM

Rustomjee R IJTLD. 2008 12:128-38

AdjuVit

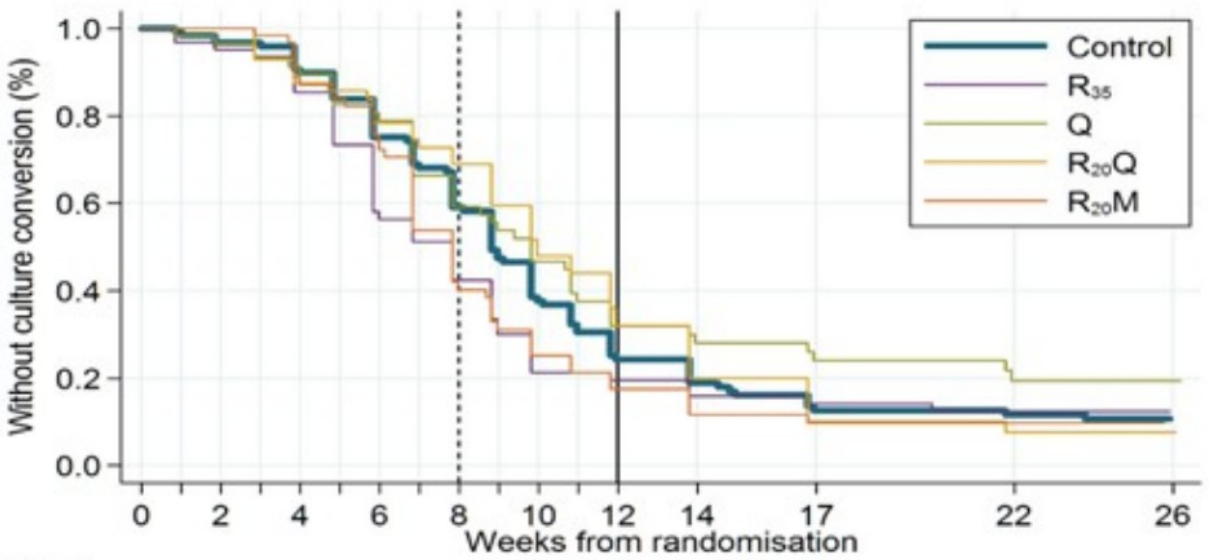


`I()` in WinBUGS

Martineau Lancet. 2012 377:242-50

Davies G Gordon conference on TB drug development 2011

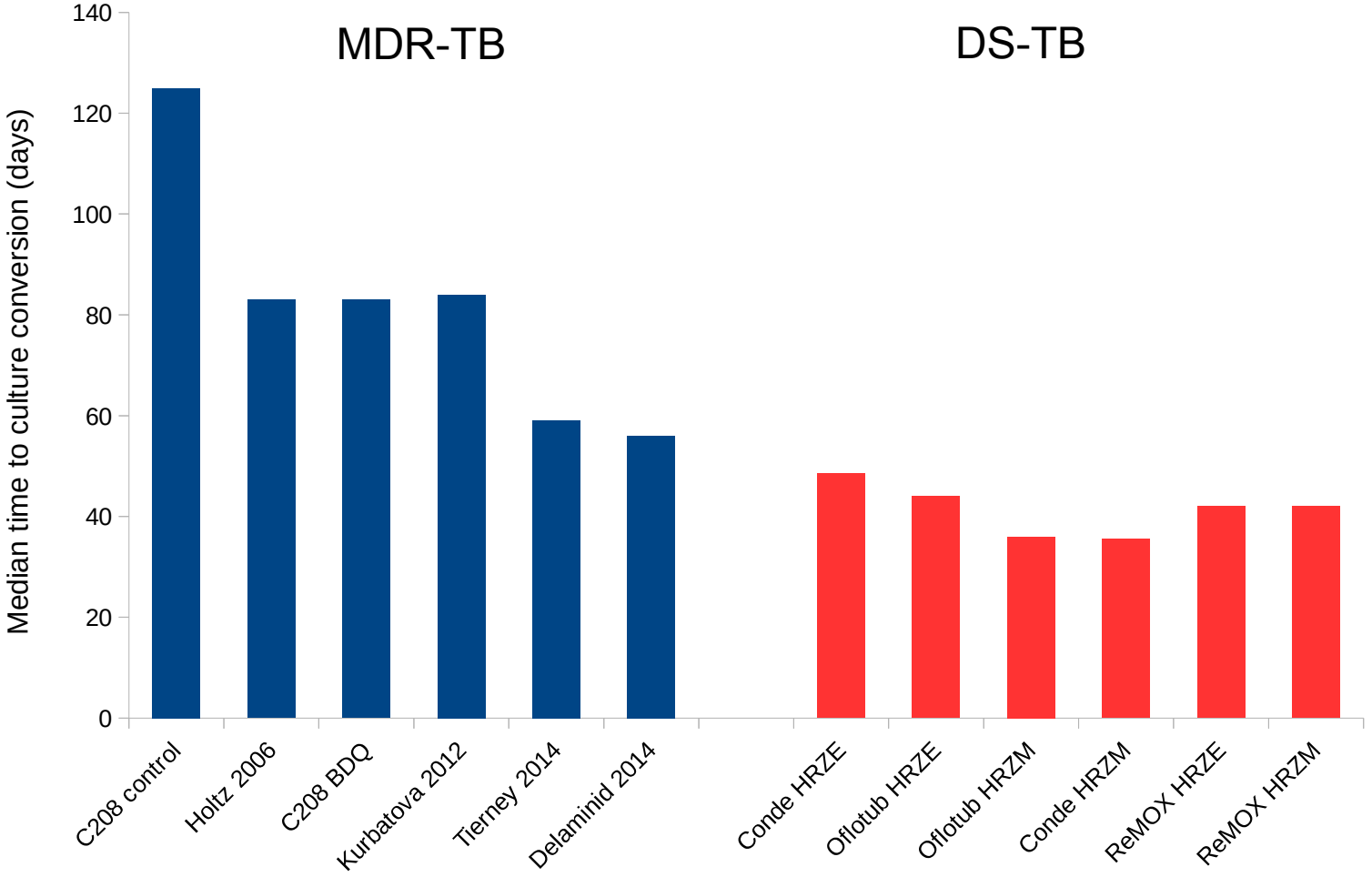
Time-to-event endpoints



	Control	Q	20RQ	20RM	35R
Included in analysis	123	58	56	63	63
Median time	62 days	63 days	66 days	55 days	48 days
Adj. HR[1] (95% CI)		0.85 (0.57 - 1.27)	0.76 (0.50 - 1.17)	1.42 (0.98 - 2.05)	**1.78 (1.22 - 2.58)**
		p=0.42	p=0.21	p=0.07	**p=0.003**

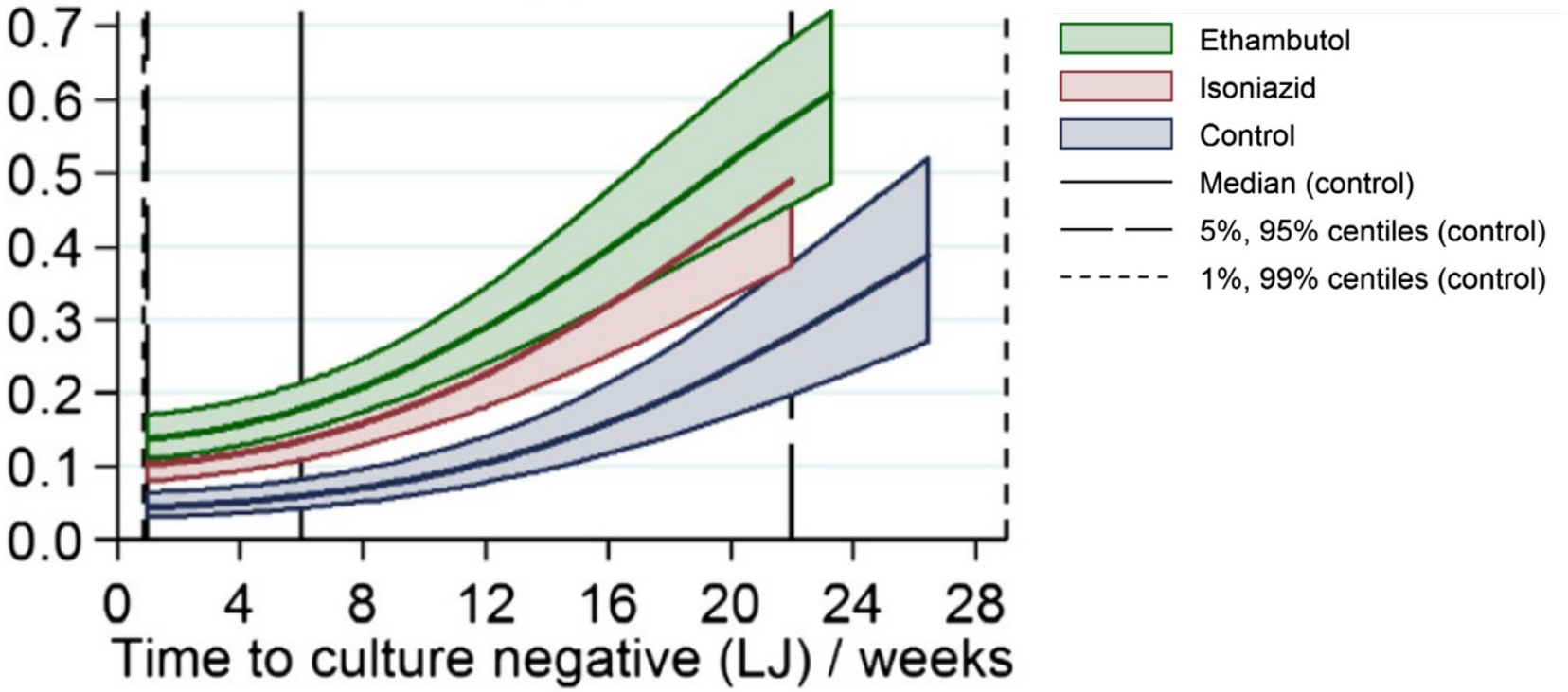
Boeree MJ, Hoelscher M CROI 2015 Abstract 95 LB

Scaling of time-to-event endpoints



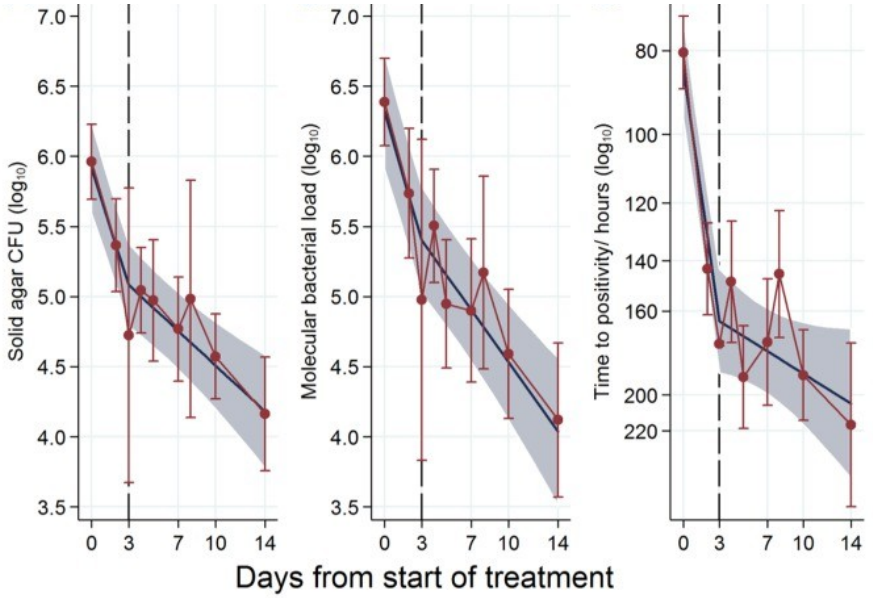
Davies GR Wallis RS 2016 IJTLD in press

Time-to-event : individual level

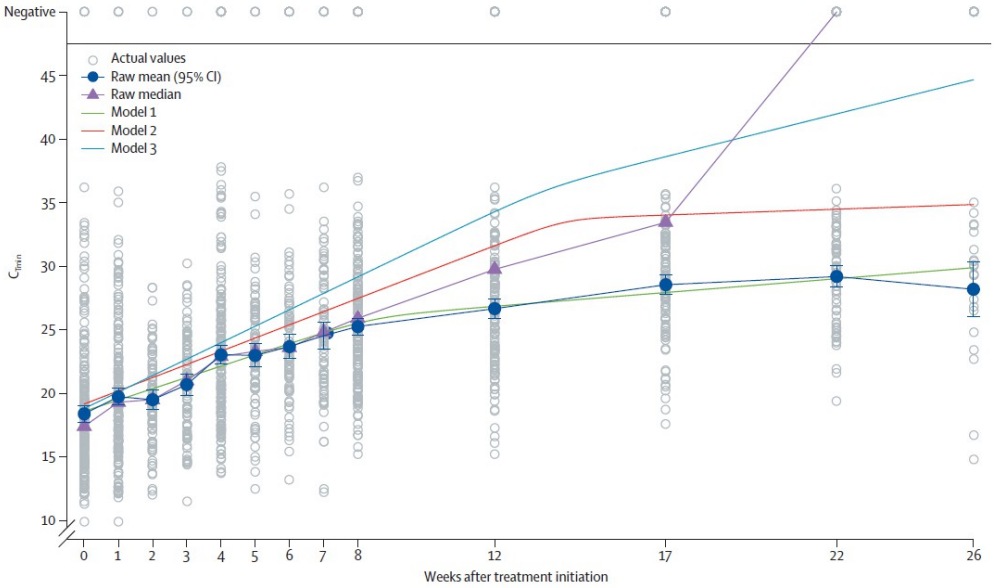


Phillips P BMC Med 2016 14 : 19

Culture-independent methods



Honeyborne I 2014 J Clin Micro 52 : 3064-7

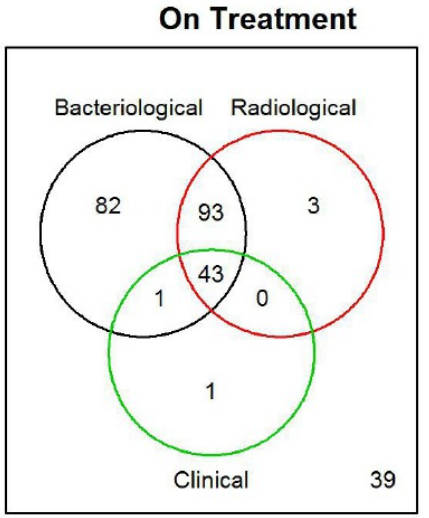


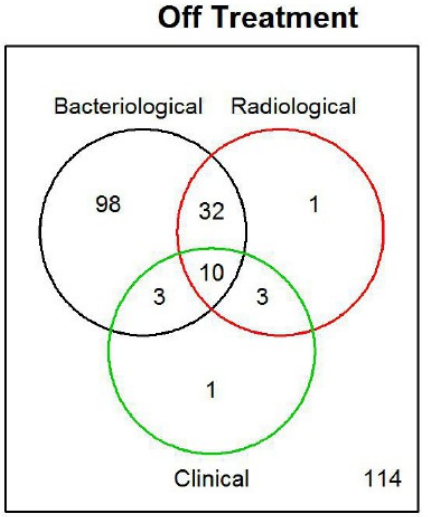
Friedrich S 2013 Lancet Resp Med 1:462

Xavier A 2013 J Clin Micro 51:1894

Alternatives to microbiological endpoints

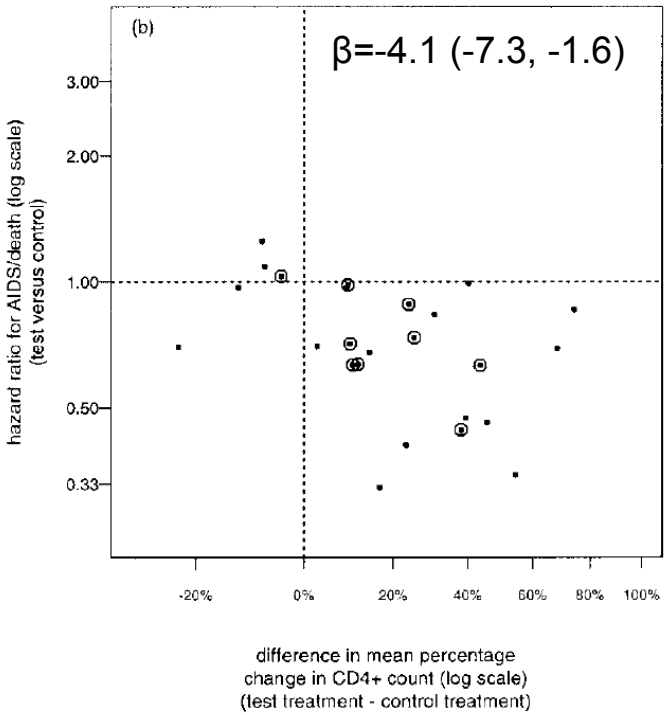
- Clinical prediction scoring
- Host response (blood/sputum IFN-γ)
- Whole blood bactericidal assay
- Host trans/proteo/metabolomic signatures
- Functional imaging (FDG)
- Composite endpoints





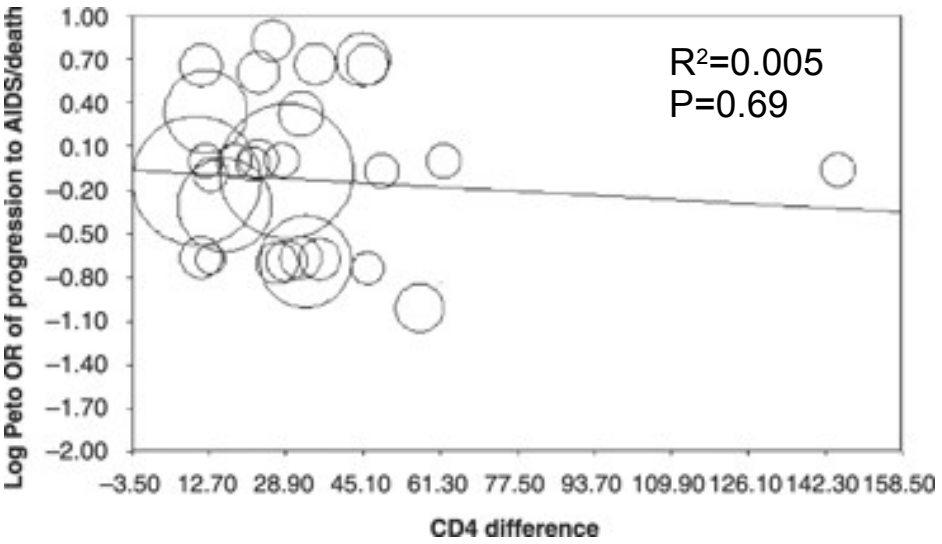
Evolution of surrogate status

16 trials 13,045 participants
1987-1997



HSMCG 2000 Aids Hum Retrovir 16 : 1123-33

22 trials 8,363 participants
1994-2006



Mills EJ 2008 HIV Med 9:849

Summary

UNIVERSITY OF LIVERPOOL

- Biological and causal plausibility of bacteriological endpoints is strong
- Extensive evidence suggests the best-reported (8w CC) is a useful surrogate endpoint and predictive of duration of regimens in DS-TB
- Lack of consensus on outcomes or analytical approaches in Phase II hampers evaluation
- Longitudinal or time-to-event approaches offer many potential advantages and have some individual-level support
- Evaluation is a process not an event and meta-analysis would ideally be curated and cumulative
- A core outcome set would be desirable

Points for discussion

UNIVERSITY OF LIVERPOOL

- What are the best endpoints and approaches to bridge the gap from Phase IIA to Phase IIB ?
- What formal statistical approaches should be favoured for evaluation of early phase endpoints ?
- How will evaluation of longitudinal or time-to-event biomarkers be achieved without a core outcome set or definitions ?
- What are the implications of adaptive approaches for evaluation of novel endpoints and biomarkers ?
- How should the TB trials community support data collection and meta-analytic approaches to address these issues ?