KNCV TUBERCULOSIS FOUNDATION
To eliminate TB

aighd

Guidance document for the evaluation of TB prediction tests to inform WHO endorsement

An update

Frank Cobelens & Sandra Kik
f.cobelens@aighd.org

KNCV Tuberculosis Foundation
The Hague, Netherlands

Amsterdam Institute for Global Health and Development
Amsterdam, Netherlands

Liverpool, 26 October 2016

Conflict of interest disclosure

- I have the following, real or perceived direct or indirect conflicts of interest that relate to this presentation:

Affiliation / financial interest	Nature of conflict / commercial company name
Tobacco-industry and tobacco corporate affiliate related conflict of interest	
Grants/research support (to myself, my institution or department):	Qiagen is donating Quantiferon Plus testkits for the WHIP3TB trial on which i am a co-investigator
Honoraria or consultation fees:	
Participation in a company sponsored bureau:	
Stock shareholder:	
Spouse/partner:	
Other support or other potential conflict of interest:	

This event is accredited for CME credits by EBAP and speakers are required to disclose their potential conflict of interest going back 3 years prior to this presentation. The intent of this disclosure is not to prevent a speaker with a conflict of interest (any significant financial relationship a speaker has with manufacturers or providers of any commercial products or services relevant to the talk) from making a presentation, but rather to provide listeners with information on which they can make their own judgment. It remains for audience members to determine whether the speaker's interests or relationships may influence the presentation. Drug or device advertisement is strictly forbidden.

Why need an evaluation framework?

1. To set a standard for *admissible evidence* for WHO endorsement (GRADE process)

2. To inform test manufacturers, researchers and research funders about the types of studies that are required for WHO endorsement

Which test should we concentrate on?

Test	Population					
	Unexposed	Infected, no incipient TB	LTBI treated	Incipient TB	Overt, clinical TB	TB treated
"LTBI test" (TST, current IGRA?)	-	+	+/-	+	+	+/-
Persistent infection test	-	+	-	+	+	-
Incipient TB test	-	-	-	+	+	
Active TB test	-	-	-	-	+	-

+ = test is positive; - = test is negative; +/- = test is sometimes positive, sometimes negative

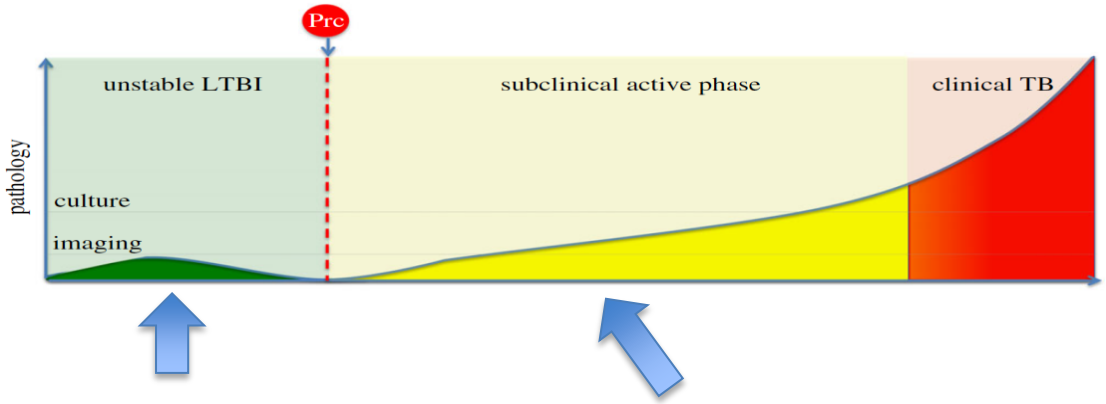
→ Concentrate on a **test for incipient TB** as this is expected to have high predictive value for incident TB disease (*rule-in test*)

Test for incipient TB

Predicts clinical TB occurring within 12-18 months

May have low sensitivity depending on when the test is done → may need to be repeated

May be combined with a test for persistent infection



persistent infection test

Rule-out *progression to TB disease*

Incipient test

Rule-in *progression to TB disease*

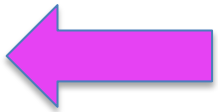
Evaluation phases

1. Analytical evaluation

evaluation of different subsets of well characterized (banked) samples

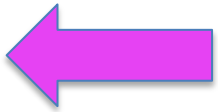
2. Clinical evaluation

evaluate the test in the intended target population
in a controlled setting with high quality standards
(compare the results of the new test against a reference standard)



3. Evaluation for (public) health impact

evaluate the test under routine conditions
for impact on patient-important or health system-important outcomes
(comparison against a reference standard not necessary)



Clinical evaluation - *admissible evidence*

For targeting preventive treatment we are not interested in latent TB infection as such, but in predicting disease

→ WHO endorsement must be ultimately based on **prediction of disease**

→ Some designs as used in evaluation of IGRA will be non-informative:

- studies comparing test results with that of IGRA or TST as 'reference' standard (beyond very early stages of test evaluation – candidate selection)
- studies that analyze test results along a *M. tuberculosis* exposure gradient
- Cross-sectional studies (= without follow-up)

Clinical evaluation – research questions

Purpose

Establish the predictive ability of the test **in the absence of preventive treatment**

Research questions:

1. What is the accuracy (sensitivity and specificity) of the test to predict incident active TB within a specified period?
2. What is the positive and negative predictive value of the test for incident active TB within a specified period, and what is the corresponding number needed to screen to find 1 positive test (NNS) and number needed to treat to prevent one incident TB case (NNT)?
3. What is the incidence rate (IR) of active TB after a positive test? What is the incidence rate after a negative test? What is the corresponding incidence rate ratio (IRR) of the test?

Clinical evaluation - designs

Key questions:

1. Is the test **positive** in persons who **develop active TB** over 12-18 months?
2. Is the test **negative** in persons who remain **without active TB** over same period?

Design:

Follow-up studies of persons with high likelihood of recent exposure or otherwise at high risk of developing TB

Options:

1. Cohort designs
2. Nested case-control designs

Clinical evaluation - cohort designs

Follow tested individuals actively over 12-18 months

Active ascertainment of incident TB, stratified by test result

Essential requirements:

Probability of being included as a TB case should be independent of test result

TB case ascertainment should be blinded with regard to test result

TB diagnosis should have high specificity (bacteriological confirmation)

Clinical evaluation – design challenges (1)

Design challenge	Low incidence country	High incidence country	Potential effect	Possible mitigation strategy
Use of preventive therapy	Present for majority of suitable study populations	Present for some study populations, but not all	Bias of accuracy estimates (if included) or limiting enrolment (if excluded)	• Choose study population in which IPT is not given (MDR-contacts, ineligible per country guidelines, declining IPT, non-adherent to IPT) • Include individuals assigned to non-intervention arm in RCT of e.g. TB preventive therapy or post-exposure vaccines trials • RCT, comparing LTBI test and treat strategy with new TB-PT test and treat strategy

Clinical evaluation – design challenges (2)

Design challenge	Low incidence country	High incidence country	Potential effect	Possible mitigation strategy
Follow-up time long	Present	Present	Long study duration, loss to follow-up (potential for new infection as discussed above)	• Use shorter follow-up time (e.g. 12 months) or analyze results for different lengths of follow-up (6, 12, 18 months) • Compare RR and IRR to determine how differential loss to follow-up may have affected study outcomes

Clinical evaluation – design challenges (3)

Design challenge	Low incidence country	High incidence country	Potential effect	Possible mitigation strategy
Progression rare	Present	Present	Large sample size needed	• Focus on highest risk groups

Clinical evaluation – design challenges (4)

Design challenge	Low incidence country	High incidence country	Potential effect	Possible mitigation strategy
Re-infection	Absent	Present	Biased estimates: ↓ sensitivity == specificity ↑ PPV ↓ NPV	• Use shorter follow-up time (e.g. 6 months) • Focus on populations with a lower risk of exposure to ongoing transmission in community (e.g. young children)

Clinical evaluation - nested case-control design

Follow tested individuals passively over defined period (passive cohort)
Passive ascertainment of incident TB
Test status among incident TB cases compared to that of random subset of non-TB cases
Allows for larger sample sizes

Requirements and design challenges:
As for cohort studies

Additional challenges:
Incomplete TB case ascertainment: no bias, but sample size trade-off

Clinical evaluation – subgroups

Of interest for stratified/subgroup analysis:

- history previous TB disease
- children
- gender
- BCG vaccination status
- comorbidities: e.g. HIV, diabetes, malnutrition

Evaluation of (public) health impact - *admissible evidence*

The new test may identify the same absolute number of persons who develop TB disease as TST or IGRA but with much higher PPV (= **lower number-needed-to-treat**)

→ Comparative studies cannot just have effectiveness endpoints but must also have **cost-benefit endpoints**

Cost-benefit should entail:

- Individual patient benefits
- Public health benefits
- Health system monetary costs
- Patient monetary costs
- Additional costs, e.g. adverse events

Evaluation of health impact – research questions

Purpose

Assess the impact of the assay on patient important outcomes and its public health impact **when used to guide preventive treatment decisions** under routine conditions

Research questions:

1. What is the effectiveness of the test for reducing incident TB when combined with a strategy to offer preventive treatment upon a positive test?
2. Is the test combined with a preventive treatment a cost-effective strategy to reduce incident TB in individuals for whom testing and preventive treatment is currently not recommended?
3. Is the test combined with preventive treatment a more effective and cost-effective strategy compared to alternative LTBI test and treat strategies using TST and/or IGRA?
4. What is the effect of the test combined with preventive treatment on the occurrence of adverse effects (e.g. hepatotoxicity), when compared to alternative LTBI test and treatment strategies (e.g. based on TST and/or IGRA)?
5. What is the effect of the test combined with preventive treatment on the uptake and acceptance of preventive treatment?
6. Which treatment regimen (monodrug or multidrug preventive treatment) is most effective when used for individuals with a positive test?

Health impact evaluation - designs

Key questions:

1. Does the test when used in routine settings **improve health outcomes**?
2. Does the test when used in routine settings **improve cost-effectiveness**?

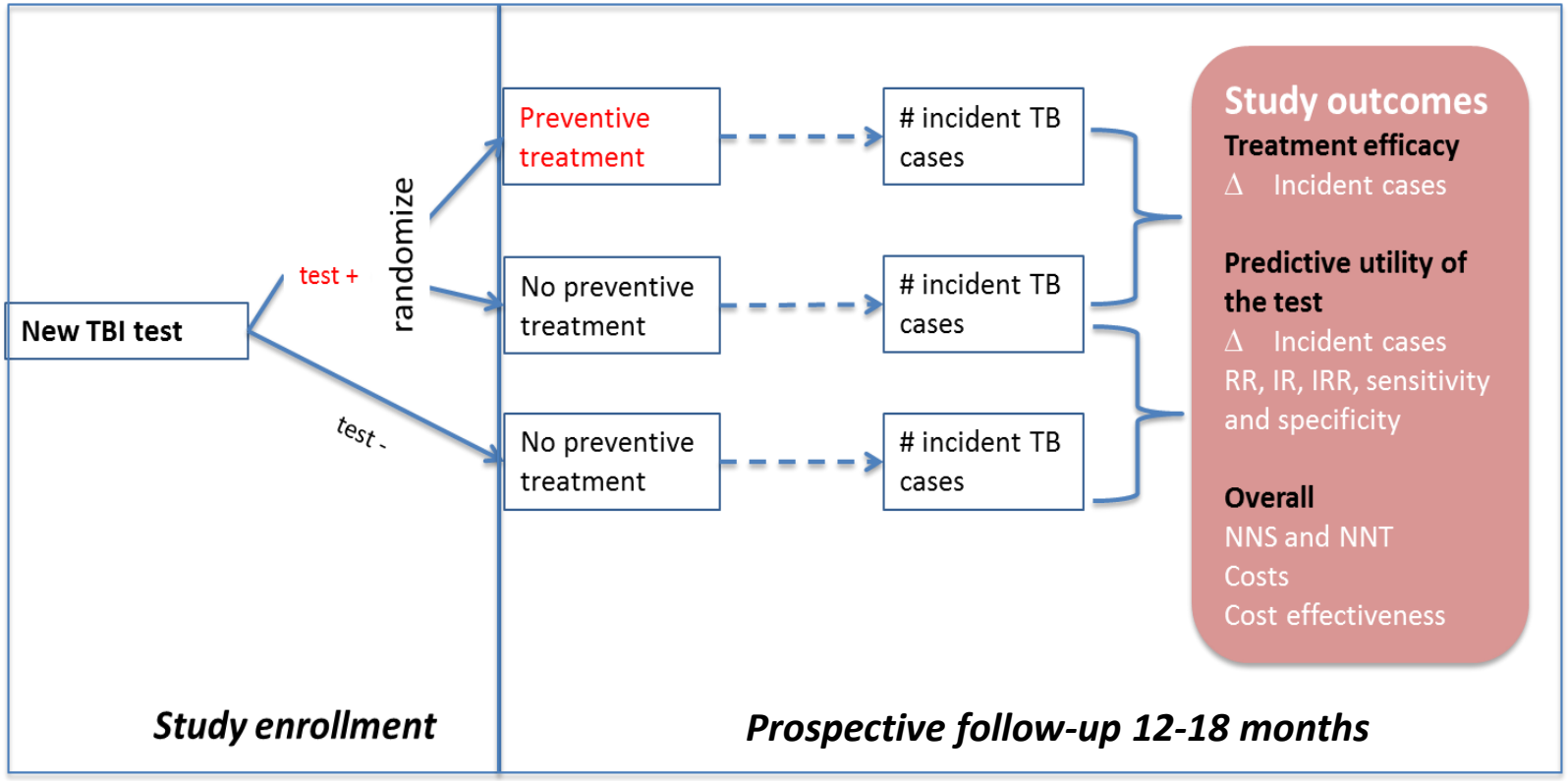
Design:

Comparative designs, ideally randomized trial (individual/group):

- Randomize individuals with a positive test for treatment vs no treatment
- Randomize individuals for old test & treat strategy vs new test & treat strategy

Health impact evaluation

Trial randomizing individuals with positive test



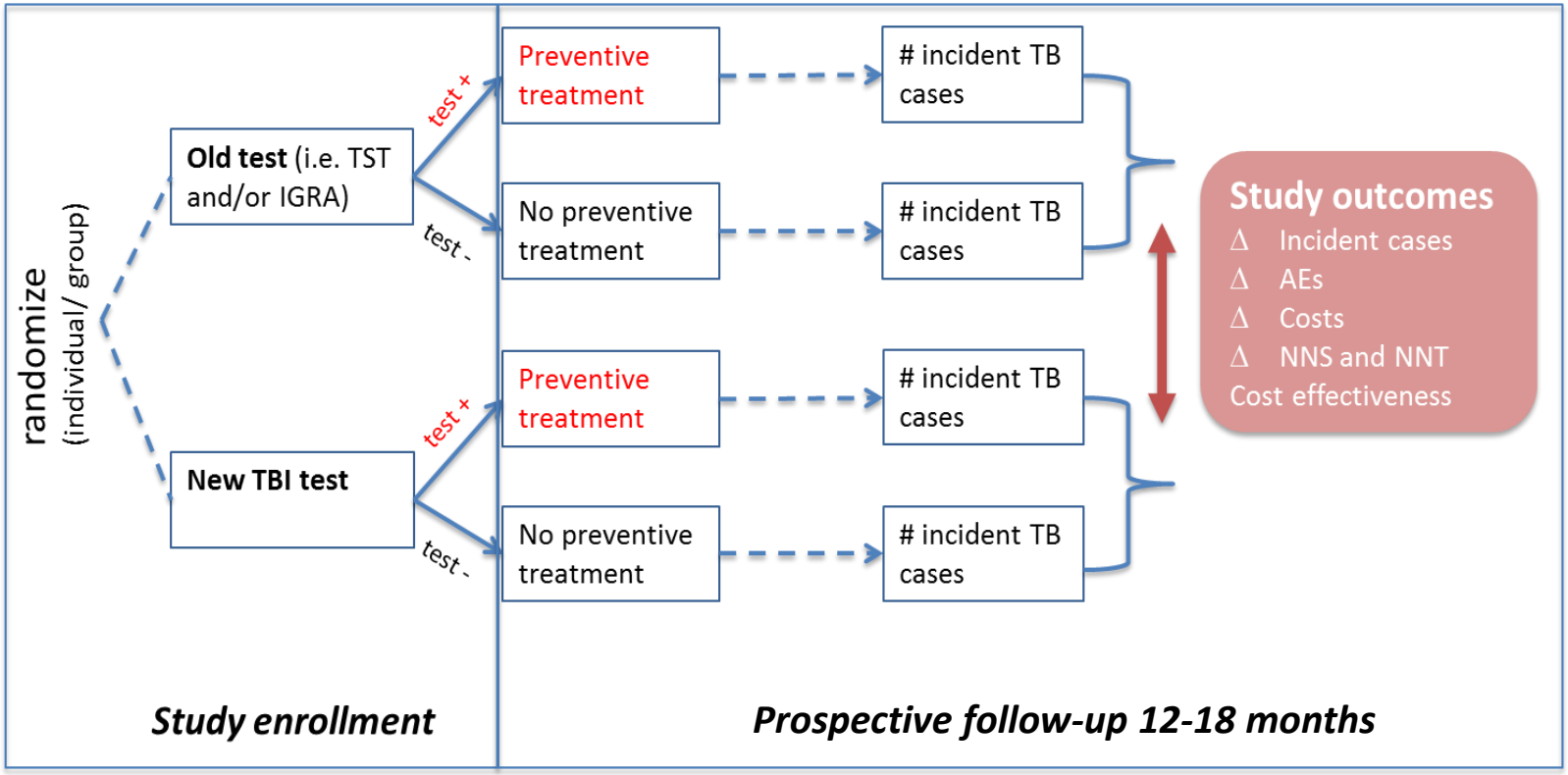
Δ=difference, IR=incidence rate, IRR=incidence rate ratio, NNS=number of individuals needed to screen to find a positive test, NNT=number of individuals needed to treat to prevent one incident TB case, RR=risk ratio, TBI=tuberculosis infection.

Based on the CORTIS study

Only in target groups that currently not eligible for preventive treatment

Health impact evaluation

Trial randomizing by test & treat strategy



Δ=difference, IR=incidence rate, IRR=incidence rate ratio, NNS=number of individuals needed to screen to find a positive test, NNT=number of individuals needed to treat to prevent one incident TB case, RR=risk ratio, TBI=tuberculosis infection.

In target groups for which preventive treatment is currently indicated

Conclusions

What we're looking for is a test for incipient TB

This requires a different evaluation approach than used for IGRA thus far

Endorsement should ultimately be based on predictive power (of incident TB)

→ follow-up studies

Cohort studies with relatively short follow-up are needed for clinical evaluation

Nested case-control studies may be useful alternative

Randomized trials are ideally done to show impact on patient/health system-important outcomes

For such trails, number-needed-to-treat, adverse events and cost-effectiveness are important endpoints

Acknowledgements

- Gavin Churchyard
- Daniela Cirillo
- Claudia Denkinger
- Chris Gilpin
- Delia Goletti
- Amita Gupta
- Alberto Matteelli
- Dick Menzies
- Molebogeng Rangaka
- Samuel Schumacher
- Alessandra Varga