PhD thesis

Thomas Stig Hermansen

Interferon Gamma Release Assays in Denmark, population-based studies from a tuberculosis low-incidence country

Academic supervisors: Pernille Ravn and Troels Lillebaek
Submitted to the University of Copenhagen April 15th 2016
This thesis was handed in April 15th 2016 for evaluation at the University of Copenhagen. The PhD defence will take place at Copenhagen University Hospital, Hillerød, June 27th 2016, 14.00 at the auditorium.

ACADEMIC SUPERVISORS:

Pernille Ravn, MD, PhD
Department of Pulmonary and Infectious Diseases, Nordsjaellands Hospital, Hillerød, University of Copenhagen, Hillerød, Denmark

Troels Lillebaek, MD, DMSc
International Reference Laboratory of Mycobacteriology, Statens Serum Institut, Copenhagen, Denmark

EVALUATION COMMITTEE

Chair
Ole Kirk, MD, DMSc
Department of Infectious Diseases, Copenhagen University Hospital Rigshospitalet, Denmark

Assessors
Christian Wejse, MD, PhD
Department of Infectious Diseases, Aarhus University Hospital, Skejby, Denmark

Frank van Leth, MD, MSc, PhD
Amsterdam Institute for Global Health and Development, Netherlands
PUBLICATIONS (referred to by their roman numerals in the text)

STUDY I (1)

STUDY II (2)

STUDY III (3)
CONTENTS

PREFACE ................................................................. 4
Acknowledgements ....................................................... 4
Abbreviations ............................................................. 5
Funding ................................................................. 5
INTRODUCTION ............................................................. 6
1. BACKGROUND ............................................................. 7
1.1. The TB spectrum .................................................... 7
1.2. Diagnosis of active TB ............................................. 8
1.3. Diagnosis of latent TB ............................................... 8
1.4. Interferon gamma release assays ................................. 9
1.5. QuantiFERON-TB Gold In-Tube Test .......................... 10
1.6. Non-tuberculous mycobacteria .................................. 10
2. OBJECTIVES ............................................................ 12
3. METHODOLOGICAL CONSIDERATIONS ......................... 13
3.1. Settings, design and populations ............................... 13
3.2. Data sources .......................................................... 13
3.3. Statistics ............................................................. 14
4. MAIN RESULTS .......................................................... 16
4.1. Study I .............................................................. 16
4.2. Study II .............................................................. 16
4.3. Study III ............................................................. 16
5. DISCUSSION ............................................................. 18
5.1. QFT Performance ................................................... 18
5.2. Non-tuberculous mycobacteria ................................. 18
5.3. QFT performance in extreme age groups ...................... 19
5.4. Predictive value of the QFT ......................................... 20
6. LIMITATIONS ............................................................ 21
6.1. Bias ................................................................. 21
6.2. QFT variability ........................................................ 21
6.3. QFT reproducibility and repeat indeterminate tests ............ 22
6.4. Screening indication ................................................. 22
7. PERSPECTIVES .......................................................... 22
7.1. The local perspective, IGRA use in Denmark ................. 22
7.2. The global perspective, IGRA use worldwide ................ 23
7.3. The future .......................................................... 23
SUMMARY ................................................................. 25
DANSK RESUME .......................................................... 26
REFERENCES ............................................................. 27
STUDY I ................................................................. 32
STUDY II ............................................................... 41
STUDY III ............................................................... 48
ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and support from a number of people.

Pernille Ravn, the principal supervisor, has been indispensable in planning the articles on which this thesis is based. I am grateful for all your encouragement, guidance, constructive criticism and introduction to the world of TB.

Troels Lillebaek, director for IRLM, my daily work place throughout the majority of this thesis, has provided valuable help always finding time for advice and fruitful discussion. Thank you for providing an inspiring environment for young researchers to learn and grow.

I am grateful to my colleagues at the PhD office, Mathias Klok Pedersen, Dorte Bek Folkvardsen, Julie Prahl and Karen Bjoern-Mortensen. We have created a fantastic working environment that I will surely miss in the future.

I am indebted to Erik Svensson and Michael Rasmussen for sparring, always keeping a door open to discuss R statistics, data extracts or something else completely.

The Department of Pulmonary and Infectious Diseases, Nordsjaellands Hospital, Hillerød, has provided me with the opportunity to initiate this PhD and enrol as a PhD student. Thanks to the entire department, in particular the research group at LIA: Dorthe, Ellen, Andreas, Pelle, Stine, Gertrud, Camilla and Kristina for sparring and exchange of ideas on research meetings and conferences around the world.

Thanks to the staff at IRLM, for putting up with me through the last years and for always being willing to assist.

The former director for IRLM, Vibeke Østergaard Thomsen, gave me the opportunity to use data from the department and perform the projects in this thesis; for that I am very grateful. Peter H. Andersen has been very valuable in providing data extracts regarding notified TB cases.

Last, but most important, I am grateful to my friends and family, Louise, Sebastian, Karl August and Selma, for putting everything into perspective and having made my life enjoyable and versatile outside the world of TB for the last three years.

Thomas Stig Hermansen, April 2016
ABBREVIATIONS

AFS – Acid Fast Staining
ATS – American Thoracic Society
BCG – Bacille Calmette-Guérin
CDC – Centers for Disease Control and Prevention
CI – Confidence Interval
CFP-10 – Culture Filtrate Protein-10
DIDE – Department of Infectious Disease Epidemiology
ELISA – Enzyme-linked Immunosorbent Assay
ESAT-6 – Early Secretory Antigenic Target-6
FDA – United States Food and Drug Administration
IDSA – Infectious Disease Society of America
IGRA – Interferon-gamma release assay
IRLM – International Reference Laboratory of Mycobacteriology
KP – Kaplan Meyer
LIMS – Laboratory Information and Management System
LTBI – Latent M. tuberculosis Infection
MAC – Mycobacterium Avium Complex

Mtb – Mycobacterium tuberculosis
NICE – National Institute for Health and Care Excellence
NNT – Number Needed to Treat
NPV – Negative Predictive Value
NTM – Non-tuberculous mycobacteria
OPUS – The name of a patient administration and reporting system
PCR – Polymerase Chain Reaction
PHA – Phytohaemagglutinin
PPV – Positive Predictive Value
PY – Person Years of follow-up
QFT – QuantiFERON-TB Gold in-Tube Test
TB – Tuberculosis
TB7.7 – TB protein 7.7kD
TBNET – Tuberculosis Network European Trials Group
TST – Tuberculin Skin Test
TU – Tuberculin Unit
TNF – Tumour Necrosis Factor
WHO – World Health Organization

FUNDING

The PhD fellowship was funded by the IRLM, Statens Serum Institut, The Research Council at Nordsjaelland Hospital, Hillerød and Olga Bryde Nielsens Fond.
INTRODUCTION

The aim of this thesis is to evaluate the performance of the QuantiFERON-TB Gold In-Tube Test (QFT) in tuberculosis (TB) and latent *Mycobacterium tuberculosis* infection (LTBI) – often referred to as active TB and latent TB – as well as the ability of the QFT to differentiate between TB and disease caused by non-tuberculous mycobacteria (NTM).

The QFT is the most widely used immunological test to diagnose LTBI and belongs to the group of Interferon Gamma Release Assays (IGRA). By evaluating the current use and limitations of these assays, we can potentially advance in the seemingly never-ending task to control and fight TB. IGRA is considered an important asset in the diagnosis of infection with *Mycobacterium tuberculosis* (Mtb), and in many TB low-incidence, high-income countries, they have replaced the tuberculin skin test (TST) as the preferred screening tool.

The development of the IGRA was made possible by studies performed at Statens Serum Institut (SSI) by Professor Peter L. Andersen and his group (4). Almost two decades ago, they identified the mycobacterial antigens used in these tests. The International Reference Laboratory of Mycobacteriology (IRLM) at SSI was at the frontier when it came to introducing the first commercial test using the new antigens. The third generation of this test, the QFT, was first used in Denmark late 2005, and in the following five years, the QFT was almost exclusively performed at IRLM. Consequently, the results of the nationwide coverage of TB cases and performed QFT could be obtained from a single laboratory. This provided a unique opportunity to combine data on all notified TB cases and all cultured NTM cases with QFT data, allowing us to create a cohort of persons screened with a QFT that was useful for large-scale evaluation of test performance.

This cohort serves as the basis for all studies in this thesis.
1. BACKGROUND

1.1. THE TB SPECTRUM

TB has unfortunately maintained a status as one of the world’s most important infectious diseases. Alongside HIV, it is a leading cause of mortality and morbidity worldwide (5). Although a reported decrease in TB cases and TB related deaths has taken place over the last decades, the disease seems in no way to be controlled or controllable with available measures. An estimated 1.9 billion people are harbouring Mtb (6) and serve as a reservoir for future TB disease. In 2014, the WHO estimated 9.6 million new TB cases and 1.5 million deaths due to TB, with the vast majority of both TB cases and fatalities occurring in the developing world (5). In Denmark, over the last decade, the occurrence of TB has remained at a constant level of 300-400 new cases per year, with approximately two thirds of these cases found among the immigrant population (7).

Infection with Mtb begins with the inhalation of infectious droplet nuclei containing the bacteria. The bacteria reaches the lungs and are ingested by alveolar macrophages (8) where they are able to replicate, and through an immune reaction involving macrophages, T lymphocytes, B lymphocytes, and fibroblasts create the primary site of infection, referred to as the Ghon focus. Within this granuloma, the host’s immune response is suppressed allowing Mtb to survive and replicate (9). The majority of infected persons will then remain infected without clinical symptoms and are said to have LTBI, in general referred to as “latent TB”. The bacteria are viable but contained by the immune system without evidence of clinical disease, and the majority of patients will never progress to overt disease. A proportion will however develop clinical manifest TB with ongoing bacterial replication, in general referred to as “active TB”.

The general perception that infection with Mtb exists solely as active TB or latent TB has been challenged over the last decade (10,11), and the different ways in which the causative agent, Mtb, can impact the host are being recognized as a continuum, rather than a binary process (12). This continuum spans from (i) early elimination of Mtb by the innate immune system without any immunological priming to (ii) the latent state that evokes an acquired immune response, but does not surface as a clinical significant disease to (iii) active TB, responsible for 1.5 million deaths every year (5). These observations have led to speculations on how Mtb infection perhaps should be considered more a spectrum rather than separate entities (13) (Figure 1).

The lifetime risk of developing active TB, once latently infected, is suggested to be 5-10% (14), with increased risk the first two years after infection, where half of the cases occur (15,16). This “reactivation” happens when the bacterial load increases (Figure 1) and the infection can no longer be contained by the host immune system. Essential in controlling TB, especially in low-incidence countries

![Figure 1. Depiction of a proposed framework for considering tuberculosis infection as a spectrum (Adapted from (13)).](image-url)
where the majority of TB cases are due to reactivation (17), is the prevention of disease. Thus, it is important to identify individuals latently infected, who could benefit from preventive treatment. Administration of preventive treatment in a population with LTBI prevents 60-90% from reactivation (18,19). Determinants of progression to active TB are numerous, and are often a mixture of host and environmental factors (12), that increase the risk of progression from latent TB. Such known host factors are suppressed cellular immunity by HIV infection (20,21), tumour necrosis factor alpha (TNF-α) inhibitors (22), glucocorticoids (23), transplantation (24,25), advanced age (26) and diabetes (27). Reported environmental determinants are prison stay (28), homeless status (29), recent immigration from a TB high-burden country (30), and recent TST conversion (31).

Depending on symptoms, bacterial load and recognition by the immune system, different treatment strategies may be used. Active TB disease with a drug susceptible strain is traditionally treated during 6 months using the standard four-drug regime consisting of isoniazid and rifampicin in combination with ethambutol and pyrazinamide for the first 2 months (intensive phase) and isoniazid and rifampicin alone for the subsequent 4 months (continuation phase). LTBI may be treated by a variety of different regimes depending on availability of drugs, compliance and adverse effects. The WHO 2015 guidelines on LTBI recommend five options for the treatment of LTBI: isoniazid monotherapy daily for 6 or 9 months, the combination of rifampicin and isoniazid daily for 3-4 months, the combination of rifapentine and isoniazid once a week for 12 weeks, and rifampicin as monotherapy daily for 3-4 months (32). In Denmark, official guidelines suggest a 6 month course of isoniazid therapy for the treatment of LTBI (33).

1.2. DIAGNOSIS OF ACTIVE TB

TB is the symptomatic disease caused by bacteria from the 
*M. tuberculosis* complex (*M. tuberculosis*, *M. africanum*, *M. bovis*, *M. canetti*, *M. microti* (34-36)). Classic symptoms of pulmonary TB, although not necessarily present, include chronic cough (sometimes with blood-tinged sputum), weight loss, fever, and night sweats (37). The reference standard for microbiological diagnosis of TB is culture, but this requires specialized laboratory facilities, and bacterial growth can take weeks or longer. Therefore, direct microscopy of bacteria using acid fast staining (AFS) from sputum samples remains the most common diagnostic method worldwide. Various staining methods exist, but the Ziehl-Neelsen stain, by which bacteria is dyed and viewed under oil immersion or the newer auramine-rhodamine stain (38) using fluorescence microscopy, are commonly used. AFS is inexpensive and simple to perform, but lacks sensitivity and specificity, as it cannot distinguish mycobacterial species. In recent years, molecular methods to detect TB have improved, and they are used increasingly in both low- and high-incidence countries. PCR-based methods such as the Xpert MTB/RIF (Cepheid) and the Geno-Type line probe assays (HAIN Lifescience) offer advantages due to the fast detection of *Mtb* and possible resistance to first line anti-TB drugs. However, these methods detect bacterial DNA solely and cannot separate live from dead bacteria - consequently they are not useful for treatment monitoring.

In Denmark, the microbiological detection of TB is based on direct smear microscopy using auramin-rhodamine stain as well as culture in fluid and solid media for up to 56 days using MGIT 960 and Löwenstein-Jensen slants respectively. In addition, molecular methods are used to obtain rapid results regarding susceptibility, species identification and to monitor overall transmission dynamics (genotyping methods). All TB cases are notifiable by the treating physician, including culture-negative cases diagnosed based on clinical criteria, radiological findings, histology, treatment response, and sometimes also on immunological tests (TST and IGRA). In Denmark, the proportion of culture-negative TB cases was 23.5% from 1992 through 2011 (39) (including notified cases where no specimens were sent for culture).

1.3. DIAGNOSIS OF LATENT TB

Currently, no gold standard for diagnosing latent TB exists, and it is not possible to find viable bacteria directly. Instead, indirect methods are used either in vivo by the widely used TST or in vitro using the newer, blood based IGRA. Both methods depend on cell-mediated immunity and recognition of mycobacterial antigens, as shown by the illustration from Andersen et al (4) (Figure 2).
The tuberculin skin test

The TST contains a purified protein derivative tuberculin unit (TU) which is a protein fraction from a Mtb strain. The standard is injecting five TU intradermal using the Mantoux technique. If cell-mediated immunity is present, the person will develop a delayed-type hypersensitivity reaction causing local inflammation and skin induration. After 48-72 hours, if an induration appears, the result is given by the transverse diameter of the induration. However, pre-test probability of infection, induration size and risk of TB disease should be taken into consideration when interpreting TST reaction (40). Also, the TST has important limitations. The need for a return visit in 48-72 hours to read the skin reaction is a logistic drawback, and because of manual reading, the reproducibility of the test may be a problem. Furthermore, boosting caused by a previous TST and resulting hypersensitivity can make serial or repeat testing difficult to interpret (41). More important, false positive results can occur due to infection with the group of NTM and if the person has been vaccinated with the *M. bovis* Bacille Calmette-Guérin (BCG) strain. These false positive results are caused by a cross-reaction between the non-specific Mtb protein mixture contained in the PPD with the majority of the NTMs (42) and the BCG strain (43). False negative results of the TST is known to occur in immunocompromised persons (due to both medical conditions and immunosuppressive therapy), where test sensitivity is poor (44). However, the TST remains the most widely used test worldwide, requiring less laboratory setup than the newer IGRAs.

1.4. INTERFERON GAMMA RELEASE ASSAYS

IGRAs are in vitro, blood based assays, which have their main use in diagnosing Mtb infection. They cannot discriminate between active disease and latent infection, and are not intended as a stand-alone diagnostic test for TB, but in certain situations they can however aid in the TB diagnosis.

IGRAs are based on a cell-mediated release of interferon gamma (IFN-γ) from T-lymphocytes, after stimulation with antigens specific for Mtb. When Mtb enters the human host (and is not eliminated by the innate immune response, see Figure 1), its antigens will evoke a cellular response with sensitization of T-lymphocytes by antigen presenting cells (Figure 2). The sensitized T-lymphocytes are stimulated by Mtb specific peptides, and recognition of these peptides in the presence of major histocompatibility complex class II will result in the production of different biomarkers, one of these being IFN-γ. Infected individuals will have an IFN-γ level above the threshold and the test will be positive. The basis of these assays is a quantification of the IFN-γ production.

The peptide antigens in the IGRAs are the ESAT-6 (Early Secretory Antigenic Target-6) and CFP-10 (Culture Filtrate Protein-10) and the genes encoding these proteins are located in the region of difference 1 locus in the Mtb genome, whereas TB7.7 (used in the QFT) is encoded in region of difference 11 (45). IGRAs have superior specificity compared to the TST because the antigenic regions are not encoded in the majority of NTM (except *M. gastri*, *M. kansasii*, *M. marinum*, *M.riyadhense* and *M. szulgai* (42)) or the vaccine strain *M. bovis* BCG (46).

Two commercially available IGRAs exist, the T-Spot.TB (T-Spot, Oxford Immunotec, Abingdon, UK) and the QuantiFERON-TB Gold In-Tube Test (QFT, Cellestes, Quiagen, Chadstone, Australia). The T-Spot utilizes an enzyme-linked immunosorbent pot assay and report the number of IFN-γ producing T-lymphocytes that respond to ESAT-6 and CFP-10. If the spot count in the Mtb antigen well reaches a predefined threshold after subtraction of the count in the negative control well, the test is positive. The
QFT provide a measurement of the interferon-γ release from T-lymphocytes stimulated by the Mtb specific antigens ESAT-6, CFP-10 and TB7.7, using the ELISA technique.

1.5. QUANTIFERON-TB GOLD IN-TUBE TEST

Third generation of the Quantiferon assay uses three pre-coated tubes to collect heparinized whole blood from the individual tested. One tube is coated with the mycobacterial antigens (TBAG tube), one is a positive control containing phytohaemagglutinin (PHA, an unspecific mitogen) used for validation and control of the assay, and one is a negative control (Nil tube) containing saline solution. The positive control serves both as control of correct handling and as control of the person’s immune status and ability to respond to the test antigens. Within 16 hours from blood drawing, the tubes are incubated for 16-24 hours before centrifugation and harvesting of plasma. At this point the samples can be stored between 2-8 °C for up to 8 weeks. The samples are then analysed by an ELISA method in which the amount of IFN-γ in IU/mL is measured.

For the test to be valid the Nil value has to be below 8.0 IU/mL and the positive control above 0.5 IU/mL. Indeterminate tests are reported to be associated with an array of different host factors, as well as environmental factors and multiple sources of variability can influence test results (described in the limitations section). The test is positive if the IFN-γ readout in IU/mL is above 0.35 after the Nil value has been subtracted, whereas a test is negative if the IFN-γ level in the TBAG tube is below 0.35 IU/mL.

The clinical presentation of NTM disease is often very diverse, e.g. in the form of pulmonary lesions, cervical lymph node enlargement, skin or soft tissue lesions or disseminated disease (49). In children, the most common clinical

<table>
<thead>
<tr>
<th>Interpretation</th>
<th>Nil*</th>
<th>TB response*</th>
<th>Mitogen Response*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive**</td>
<td>≤8.0</td>
<td>≥0.35 IU/mL and ≤25% of Nil</td>
<td>Any</td>
</tr>
<tr>
<td>Negative***</td>
<td>≤8.0</td>
<td>&lt;0.35 IU/mL or &lt;25% of Nil</td>
<td>≥0.5</td>
</tr>
<tr>
<td>Indeterminate**</td>
<td>≤8.0</td>
<td>&lt;0.35 IU/mL or &lt;25% of Nil</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td></td>
<td>&gt;8.0</td>
<td>Any</td>
<td>Any</td>
</tr>
</tbody>
</table>

* The interferon gamma (IFN-γ) concentration in plasma from blood incubated without antigen.
* The IFN-γ concentration in plasma from blood stimulated with ESAT-6, CFP-10 and TB7.7
* Interpretation indicating that Mtb infection is likely
* Interpretation indicating that Mtb infection is not likely
** Interpretation indicating an uncertain likelihood of Mtb infection.

Table 1. Interpretation Criteria for QuantiFERON-TB Gold in-Tube Test based on package insert.
presentation of NTM disease is cervical adenitis, most often due to mycobacteria from the *M. avium* complex (MAC) (55). This localized lymph node enlargement and chronic infection is regarded an independent disease entity different from the pulmonary disease observed among adults, and it is occurring in immunocompetent children aged 1-5 years without systemic symptoms (55). As NTM infection can mimic both pulmonary and glandular TB, a diagnostic tool that could facilitate differentiation would be valuable.
2. OBJECTIVES

1. To assess the performance of the QFT among patients with NTM disease and to review available literature on the subject.

2. To assess the performance of the QFT among individuals with and without active TB.

3. To determine the predictive value of the QFT for development of active TB in a Danish cohort and identify factors potentially associated with increased risk of progression.
3. METHODOLOGICAL CONSIDERATIONS

3.1. SETTINGS, DESIGN AND POPULATIONS

All three studies were retrospective and based on a cohort of individuals screened with a QFT from January 2005 until the end of December 2010. During this period, more than 95% of all QFT done in Denmark were analysed at IRLM, SSI. From 2011 onwards, the test has been increasingly performed at several local biochemistry departments instead of SSI, making similar nationwide studies difficult to perform in the future. The present study is thus based on a unique data material covering the majority of the Danish population screened for LTBI with a QFT in a six-year period.

Study I combined the QFT cohort with nationwide data on persons with a positive NTM culture retrieved from the IRLM. Study II and III both combined the QFT cohort with an extraction from the Department of Infectious Disease Epidemiology (DIDE) on notified cases of TB. In study III additional data were retrieved from hospital case records. Figure 3 depicts the participant flow from available data sources. All databases were combined using the Personal Identification Number (CPR) included in the Danish Civil Registration System (56). Further description of study design, study settings and populations can be found in article I-III (1-3).

3.2. DATA SOURCES

LIMS

Data on QFT performed at IRLM from 2005 through 2010 were extracted from the Laboratory Information and Management System (LIMS). LIMS is used to register and report all diagnostic examinations performed at the Division of Diagnostics & Infection Control at SSI. All entries in LIMS are based on the CPR, and this number is used to connect LIMS with other registries enabling reporting of diagnostic examinations to the requestor. Data stored in LIMS is incorruptible and cannot be altered once entered in the system. This way data is stored in its original format over time, ensuring perfect reproducibility and making the LIMS a comprehensive data collection well suited for research purposes. The drawback of having a database where data cannot be modified or updated is that extracts from the LIMS are very diverse if the reporting of a

![Figure 3. Participant flow from available data sources study I-III](image-url)

**Abbreviations:**

LIMS - Laboratory Information and Management System;
IRLM - International Reference Laboratory of Mycobacteriology;
DIDE - Department of Infectious Disease Epidemiology
diagnostic test changes. In the present thesis, this adversely affected the extract because the description of a QFT result in LIMS had been changed numerous times during the initial years of testing. Hence, a lot of time-consuming data cleansing was necessary before the analysis could be performed. To eliminate some of the diversity in the QFT reporting, we excluded the second generation of the test, the QuantiFERON-TB Gold and included only the third generation of the test, QuantiFERON-TB Gold In-Tube Test (QFT).

**IRLM Database**

In Denmark, the culturing of Mycobacteria is centralized at IRLM. The database contains information on all samples received at the laboratory that turn out positive by either microscopy, molecular methods or culture. Positive specimens are manually entered into the registry by patient ID. Sample type, citizenship, country of birth and address are included in the registry, whereas negative samples are not. The original and current purpose of the database at IRLM is fast access to information on all patients with a positive mycobacterial specimen and the opportunity to perform population based research. As opposed to the LIMS, the data in the IRLM database is updated and modified in continuity, making data extracts readily usable without cleansing and the need for interpretation of data output. In this thesis, an extract of positive NTM cultures was used in study I.

**DIDE Database**

Since 1990, the Department of Infectious Disease Epidemiology (DIDE) at SSI has received nationwide data on notified TB cases (57). In Denmark, the physician responsible for treating the TB patient are required to notify DIDE using a standardized form (58). This form contains demographical, clinical and epidemiological data regarding the TB patient, such as country of exposure, way of transmission, type of infection (pulmonary or extra-pulmonary), outbreak conditions and diagnostic method. The database at DIDE is maintained and updated regularly, and data concerning TB surveillance and disease control in Denmark is published through the SSI website as short reports and graphical displays available to the public (57).

### 3.3. STATISTICS

**Descriptive statistics**

The quantitative (continuous) variables were described as mean and standard deviation of normally distributed variables and as median and interquartile range of non-normally distributed variables. The qualitative (categorical) variables were summarized using frequency distributions and described as numbers and percentages out of the total.

**Statistical analysis**

Wilcoxon rank sum/Mann-Whitney U was the non-parametric test used to evaluate continuous variables and Pearson’s chi-squared or Fisher’s exact test (frequency counts <5) was used to compare categorical variables. Cochran-Armitage test for trend is a modification of the Pearson chi-squared test where a suspected ordering in the proportions of the categorical variables is incorporated. The null hypothesis is that the linear trend parameter is zero, so a significant p-value is consistent with the presence of a trend. It was used to evaluate an association between binomial variables and variables with more than two categories. Kaplan Meyer (KM) survival estimates were used. To investigate significant differences between the KM curves and the two derived survival distributions, the nonparametric log-rank test was applied.

<table>
<thead>
<tr>
<th>Database Name</th>
<th>Data source</th>
<th>Used in Study</th>
<th>Purpose of database</th>
<th>Description of extracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIMS</td>
<td>LDUS enquiry in CSF format</td>
<td>I,II and III</td>
<td>Register and report diagnostic investigations at SSI</td>
<td>Data on all QFT performed from 2005-2010</td>
</tr>
<tr>
<td>IRLM</td>
<td>Extract from &quot;Mycobakterieregister&quot; in Excel format</td>
<td>I</td>
<td>Data collection on patients with a positive Mycobacterial specimen</td>
<td>All positive NTM cultures from 2005-2010</td>
</tr>
<tr>
<td>DIDE</td>
<td>Extract from DIDE Database in Excel format</td>
<td>II and III</td>
<td>Central reporting of all notified TB disease cases</td>
<td>Notified TB cases from 2005-2011</td>
</tr>
<tr>
<td>OPUS</td>
<td>OPUS data on incident TB patients, manual data entry into spreadsheet</td>
<td>III</td>
<td>Records of all Hospital admissions in Denmark</td>
<td>Patients with Incident TB from 2005-2011</td>
</tr>
</tbody>
</table>

**Table 2.** Data sources in Study I-III.
Statistics regarding diagnostic tests

Sensitivity was calculated as \( \frac{\text{#true positive tests}}{\text{#true positive tests} + \text{#false negative tests}} \). Positive predictive value of a test was defined as the proportion of patients with an initial positive test that developed disease, and negative predictive value defined as the proportion that scored negative in the test and did not develop disease. Overall, a two-sided p-value <0.05 was considered significant. For the statistical analysis, SAS® 9.3 (SAS Institute Inc., Cary, NC, USA) and R version 3.1.1 (59) were used.
4. MAIN RESULTS

4.1. STUDY I

Background
The incidence of NTM infections have been reported to be on the rise and clinical manifestation of NTM disease can mimic TB. If a blood test can differentiate between TB and NTM disease, this could provide a minimally invasive tool for the clinician compared to traditional microscopy, culture and PCR. The reason why the QFT could potentially be useful is that the majority of the NTM do not share the M. tuberculosis specific antigenic regions used in the QFT.

Aim
The study objective was to investigate test performance among patients infected with NTM and to review available literature on IGRA among patients with NTM disease.

Results
The majority (50/53) of patients had NTM disease due to mycobacteria not sharing the antigenic region with Mtb, and we found 2 positive, 43 negative and 5 indeterminate QFT results, resulting in a specificity of 96% (43/45). Three patients had disease due to Mycobacteria known to share the antigenic region with Mtb, and in this group, 67% (2/3) were positive.

We found 9 published studies with a total of 478 patients with NTM disease and an IGRA. In patients infected with an NTM without the antigenic region, the QFT was positive in 12% (ranging from 0-50% in the different studies) compared to our rate of 4.4% (2/45). Patients infected with NTM sharing the region with Mtb had an overall positivity rate of 57%, which is in line with our finding of 67% (2/3) were positive.

Conclusion
The QFT might hold potential in discriminating between NTM and Mtb infections, in particular in children, where lymphadenitis due to mycobacteria from the MAC was predominant and because this clinical presentation may mimic TB lymphadenitis.

4.2. STUDY II

Background
Even though IGRA are used increasingly for the diagnosis of LTBI worldwide, few studies have investigated the test performance in extreme age groups. Current guidelines warrants caution when using IGRA in children less than five years of age due to lack of studies performed. Similarly, only a limited number of studies evaluate test performance among elderly, and no guidelines specifically addresses the ageing population.

Aim
To evaluate QFT performance among patients of different age groups with and without TB.

Results
Among 383 patients with active TB, we found a low indeterminate rate of 3.9% and a high sensitivity of 86.1%. The latter was unaffected by localization of TB (pulmonary/extra-pulmonary), sex and diagnostic group (culture and PCR confirmed TB versus clinical TB diagnosis only). Sensitivity declined with increasing age and was thus found to be 71% among elderly persons aged more than 65 years, but was not impaired in young children with active TB (100%). Among 15,709 persons without TB, the indeterminate rate was overall 5.1%, and higher in infants aged <1 year (15.6%) and elderly aged >65 years (8.1%). From age one and above the indeterminate rate was not significantly increased compared to adults aged 15-64 years.

Conclusion
QFT sensitivity among TB patients was overall 86% but declined with increasing age (>65 years). In contrast to what current guidelines on IGRA use in small children state, our data suggest that the QFT performs well in children from age 1, but that caution should be taken when using the test among the elderly and infants.

4.3. STUDY III

Background
Latent TB is prevalent in high numbers throughout the world, but only a small proportion of persons infected with...
**Mtb** will eventually develop TB disease. A valuable diagnostic test detecting the presence of latent infection should be evaluated by its ability to predict how many will develop disease and ultimately be able to identify those in highest risk of progression.

**Aim**
To investigate the predictive value of the QFT.

**Results**
We included patients with a QFT performed and combined these with notified TB cases. We defined “co-prevalent TB” as cases 0-90 days after testing and “incident TB” as cases developing after more than 90 days. In total, 231 patients had co-prevalent TB and 40 patients developed incident TB within the follow-up period. The cohort was subject to a follow-up for 52,807 person-years, and median follow-up time was 3.36 years. For incident TB, we confirmed a very high negative predictive value of 99.85% and found a positive predictive value of 1.32%. Expressed as risk of developing incident TB, this corresponds to an incidence rate (IR) of 383 per $10^5$ PY for QFT positive patients and an IR of 45 per $10^5$ PY for QFT negative patients. The risk of incident TB was highest the first two years after a positive test, and only two patients out of 20 with a positive test completed preventive therapy.

**Conclusion**
We found a high NPV and a PPV of 1.32% for development of incident TB. Development of incident TB was associated with time interval after the QFT, but not with age.
5. DISCUSSION

5.1. QFT PERFORMANCE

Sensitivity among patients with active TB
Sensitivity of the QFT among patients with TB in Denmark was comparable to what has been found in available literature (86%). In line with previous studies, risk factors such as gender (60), localization of disease (pulmonary vs. extra-pulmonary TB) (61) and culture result (presence or absence of a positive Mtb culture) (62) did not affect test sensitivity (2). Sensitivity has been evaluated extensively during the last decade. In general, the TST, QFT and T-Spot have comparable sensitivities of 80-90% in most studies of immunocompetent individuals (61,63,64), but because the tests are based on the cellular immune response, reduced sensitivity are linked to immunological function and immunocompromised patients. HIV infection and very young age are shown to reduce sensitivity for all diagnostic tests (65,66), highlighting a paradoxical problem for all immune based tests for TB; they suffer reduced performance among those patient groups that are most susceptible to both active and latent TB.

The QFT cannot separate latent from active TB and should not be used alone in patients suspected of active TB. It can however, be used as a tool in the diagnosis of active TB in situations including; patients with extra-pulmonary TB, patients who are persistently negative by microscopy or culture, TB diagnosis in children, or as a supplement in differential diagnosis between infections with NTM and TB (67). Only few studies have assessed whether adding an IGRA test to the existing diagnostic work-up for active TB can successfully identify a greater proportion of TB cases in TB high-incidence settings (68,69), TB low-incidence settings (70) and in child populations (71). Discouragingly, in most clinical situations an IGRA or TST was not of additional value when combined with standard methods to diagnose active TB (67).

Distinguishing active from latent TB using IFN-γ level
IGRAs (as well as the TST) cannot distinguish between latent and active TB. If an algorithm based on the commercially available IGRAs could make this distinction, it would be immensely valuable. Numerous studies have indeed investigated whether the level of IFN-γ readout is higher in patients with active TB compared to those with latent TB (70,72-76). In study III, we describe significantly higher IFN-γ levels in the group that develops incident TB compared to the group with a positive QFT that remain healthy. However, the confidence limits are very wide and the test cannot discriminate between the two groups, so this observation has no practical use. A new generation of the QuantiFERON test has recently been introduced (77). It contains two tubes with mycobacterial antigens directed at different subpopulations of T-lymphocytes and might have potential in discriminating active TB from LTBI. The first independent evaluation of this test (78) shows similar specificity and sensitivity but suggest that accuracy might have improved in patients with low CD4+ T-cell counts (78), whereas discriminatory ability is not evaluated. So far, in one decade of reporting and studying the IGRAs, no human studies have conclusively shown that they can distinguish between active and latent TB, and currently available IFN-γ based assays will most likely not be able to (79). Numerous studies are focusing on discriminating TB stages (80,81) and the search for alternative biomarkers and dormancy antigens that could be used in the development of next generation assays is ongoing (82).

5.2. NON-TUBERCULOUS MYCOBACTERIA

The role of IGRAs in the diagnosis of NTM disease is as yet insufficiently explored. European Centre for Disease Control and Prevention guidelines state that an IGRA can contribute with supplementary information in the differential diagnosis between NTM and TB disease (67), but this recommendation is based on very limited evidence. Study I (1) as well as other studies find a low proportion of positive tests among patients infected with NTM without the RD1 antigenic region and a significantly higher proportion among patients infected with NTM sharing the RD1 (83-93). Based on the high specificity of the QFT in patients with NTM disease in study I and on available literature, we find that the QFT holds potential to discriminate between NTM and TB (1).
NTM disease is a rare event in Denmark (approximately 100 new patients per year), and the need to separate NTM from TB using a QFT is not widely applicable in a Danish setting. However, in children presenting with cervical lymphadenitis, our study found all 15 children aged 0-9 years with MAC lymphadenitis to be QFT negative, resulting in a specificity of 100%. Lymphadenitis due to NTM is common in young children and the clinical presentation can mimic extra-pulmonary TB. The differentiation relies on a suitable biopsy specimen from which the mycobacteria is identified by culture or molecular methods. If a less invasive test such as an IGRA could aid in this separation it could be useful, but the IGRA's suboptimal sensitivity as reported in study II (86%) calls for caution when using such tests to distinguish between TB and NTM.

5.3. QFT PERFORMANCE IN EXTREME AGE GROUPS

Children

Young children are facing an increased risk of developing active TB if they become infected, and they often have disease with few bacilli (paucibacillary TB). Also the TB disease manifestations are different from adults in the form of more miliary TB and TB meningitis (94). This makes children a particularly vulnerable group, difficult to diagnose and suffering more severe disease than adults do. As a result, a positive IGRA can be useful to support a suspicion of TB along traditional diagnostic measures (67).

Study II suggests that already from age one, the performance in children with active TB is good, although few were included. Studies of IGRA performance in children are very heterogeneous, and few investigate performance in children aged less than 5. Some studies demonstrated a lowered sensitivity in children aged <5 years (94,95), whereas others report IGRA sensitivity comparable to adults from 2 years of age (85,96,97). In TB high-incidence countries both IGRA's and TST perform poorly among children, most likely due to a combination of severe comorbidity causing immunosuppression, misclassification and high rates of reinfection (98). In the study by Rose et al in a Tanzanian population of 211 children, of whom 77 were aged <2 years, the QFT performed badly with sensitivity of 19% and indeterminate rate of 27% (98).

We evaluated the indeterminate rate in children without active TB, and found no effect on indeterminate rates in children from age 1, but among infants aged <1 the indeterminate rate was high (15.6%) (2). This suggest that an immature immune system in children seem to affect test performance only the first year of life (2), and is most likely caused by a T-cell population dominated by naïve cells (99). Studies on indeterminate rate and children are numerous, and some report that young age is associated with more indeterminate results (100-103) whereas others do not reach such a conclusion (95,97,104). Overall, IGRA and TST could both be used for latent TB diagnosis from age 1 in otherwise healthy children. None of the tests should be used for a definitive diagnosis of active TB.

Elderly

In the elderly, the T-lymphocyte response is decreased (105). This can affect the QFT indeterminate rate and sensitivity,
and is probably caused by an exhausted population of memory T-cells. In study II, we found that among those aged 65 years and above, the QFT had more indeterminate results as well as decreased sensitivity. Consequently, the likelihood of both false negative and indeterminate tests is higher amongst the elderly. Whether the decreased sensitivity is a direct effect of old age or caused by age related immunosuppression is not possible to determine (2). Our findings are supported by a recent TBNET study (106) where extreme age was the only significant factor associated with higher rate of falsely negative IGRA.

We have assessed the influence of age on the quantitative PHA response and confirmed that the PHA response is affected in the very young and the very old (Figure 4, adapted from Table 4 in Study II). Figure 4 illustrates the PHA response from persons screened for LTBI in different age groups. In children 0-9 years and elderly >65 years the response to PHA is lowered, but the indeterminate rate (PHA<0.5) is only higher in infants <1 and elderly >65, suggesting that IFN-γ release is lowered throughout early childhood but does not affect QFT result in children aged >1 (2).

5.4. PREDICTIVE VALUE OF THE QFT

Negative predictive value

In a cohort of 13,463 individuals we found a very high NPV of 99.85% (study III). We have hereby established that the NPV of the QFT in Denmark is extremely high, corresponding to the findings in other studies from TB-low-incidence settings (72,76,107). The individuals in study III constituted a heterogeneous group tested with QFT due to recent exposure, in the diagnostic work-up of active TB, or before the initiation of anti TNF-α treatment. Although we cannot separate the groups by test indication, the NPV in general was very high, and the QFT seemed a useful tool for identifying those that would not progress to TB disease. However, of the 40 patients who developed incident TB, 20 actually had a baseline negative test. This could be explained by different factors such as an impaired immune system or re-infection with a new strain. In environments with active transmission, the NPV is reduced because of new infections, and this could be an explanation among the socially marginalized persons in Denmark, where the prevalence of active TB are reported to be more than 2000 per 10^5 (108). Of the 20 patients with incident TB and a baseline negative test, 25% (5/20) were immunosuppressed at the time of testing, and these test results could potentially be false negative.

Positive predictive value

In a mixed population of patients tested with the QFT, 1.32% developed TB during follow-up (study III). This is equivalent to an 8 times increased risk of developing TB when a person has a baseline positive test compared to a baseline negative test. The PPV vary substantially in available studies (63,76,107,109) depending on the TB prevalence in general and the population in particular: In a vulnerable population of recently exposed children, Diel et al found a PPV of 28.6% (107), whereas the PPV was well below 1% (110) amongst health care workers. A PPV of 1.3% is intuitively low for a diagnostic test but the PPV varied and was higher among children and adolescents aged less than 35 years (2.22%). The risk of progression to incident TB was significantly higher within the first two years after the QFT, where 90% of incident cases developed. IR of progression was almost 800 per 10^5 PY the first year, 400 per 10^5 PY the second year and 90 per 10^5 PY after more than two years (3), in line with historical data describing an increased risk during the first years after testing (111). A general issue to consider regarding the low PPV is that only 5-10% of persons with LTBI develop disease (14), so even a flawless test with perfect accuracy will have a low predictive value.

Number needed to treat (NNT)

Generally, the QFT should be performed with a clear indication, and the reasons for testing as well as the consequences of a positive or negative result have to be considered before deciding to test. In 2011 the catchphrase “Intention to test is intention to treat” was introduced (112). The point of that concept is that unnecessary tests for LTBI should be avoided. The purpose of screening to detect LTBI is to prevent future TB cases by administration of preventive therapy. The PPV reflects the proportion of the persons screened with a positive test result, who ultimately develop TB, but the clinical usefulness of the test can also be evaluated using another measure: The Number Needed to Treat (NNT) expresses the number of persons with LTBI that have to be treated with preventive therapy to avoid one
case of active TB. NNT is central in cost-benefit analysis of the use of IGRA s in TB low-incidence countries. Sester et al (113) have reviewed the NNT to prevent a TB case among close contacts (30-37 (76,114,115)), immunocompromised (50-80 (116)), HIV infected (14-26 (116)) and health care workers (no TB cases developed among >15,000 persons (117-119)). Thus, NNT differs according to the risk group, and in all studies, confidence limits were wide because of few incident TB cases. In Denmark and comparable TB low-incidence countries mass screening is not cost effective (63) because both the number needed to screen to identify a person with LTBI and the NNT is very high. All studies point toward targeted screening in high risk groups (32).

Another relevant measure when addressing TB prevention is the Number Needed to Harm (NNH). It indicates the number of patients that can be treated with preventive therapy for one patient to suffer adverse effects. With the increasing likelihood of isoniazid-induced hepatotoxicity with increasing age (120,121) the NNH is age related and screening elderly have to take this into account.

6. LIMITATIONS

6.1. BIAS

In general, patients screened with IGRA have a high a priori likelihood for TB and LTBI. Thus, the studied population is not representative for the Danish population and our studies will overestimate the prevalence. Selection bias is also present when we investigate age and risk of latent and active TB, because the indication to screen a child or a young person is often different from screening elderly. Among children, recent exposure or suspicion of active TB is a common indication, and the risk of active TB is high if infected (122). Reversely, among elderly aged 35 and above, screening before starting immunosuppressive treatment is increasingly common, and in this group the prevalence of LTBI is low and consequently very few will develop TB. Therefore, the risk of TB could differ because the age groups have different a priori risks of TB and not because age is an independent risk factor.

Misclassification can occur when cases and controls are not correctly recognized as such. False negative test results among patients with TB disease occurred in 14% of the active TB population in study II, and this misclassification could result in unrecognized TB cases. In study II, we evaluate how age, sex, localization of disease and culture/PCR versus clinical confirmation of active TB influences the rate of false negative tests. Confounding is present when two evaluated factors do not act independently in their contribution to the outcome under investigation.

Confounders that should always be considered are sex and age. In our studies these variables were controlled for in the data analysis in study II and III.

6.2. QFT VARIABILITY

Challenges with the QFT assay in terms of variability can be assigned to a number of overall sources (63); Manufacturing issues, preanalytical sources of variability, analytical sources of variability and immunological sources of variability as depicted in Figure 5.

Figure 5. Sources of variability in the QuantiFERON-TB Gold In-Tube assay. Reprint from (63)
In our cohort of persons screened with a QFT, all of the factors shown above may influence test results in both a positive and negative way. Because of the retrospective nature of our studies, the influence of the preanalytical and immunological sources of variability on our cohort is difficult to address. No information regarding the preanalytical filling of blood tubes or handling and transport of the samples before they arrive in the laboratory are available. Further, we lack information on host factors for the individual patients because no indication for screening or information on comorbidity is available. Regarding analytical variability, quality control is performed in each run of the ELISA test in accordance with manufacturer’s instructions. Manufacturing defects could be defects such as errors in the production of the kits or reagents used. It can be difficult to detect for individual laboratories, but monitoring of sudden increases in indeterminate or positive test results is an important control measure.

6.3. QFT REPRODUCIBILITY AND REPEAT INDETERMINATE TESTS

The systematic sources of variability of the QFT have been accounted for (Figure 5), but random variation due to immunological factors can also affect test results. Problems with QFT reproducibility in the same person undergoing serial testing is well known, and reports of conversions (a negative test becomes subsequently positive) and reversions (a positive test becomes negative) are numerous (118). The introduction of a borderline zone to better understand test dynamics has been proposed in order to counter this (123). The grey zone is the level of IFN-γ release in IU/mL, where most reversions and conversions appear, and it can be used to alert the clinician to interpret results with caution. In study III, we describe a larger proportion of patients who develops incident TB are in the grey-zone from 0.2-1.0 (3,117,123) compared to the group without TB. This finding is interesting albeit not surprising, because it suggests impaired immunity plays a role in the susceptibility towards developing TB disease.

In study II, the indeterminate rate of all persons tested was 5.1%. Of these, 40% were repeated and 67% changed to either positive or negative in the subsequent test. The recommendation when encountering an indeterminate QFT is to repeat the test (33) and a majority of these repeats turned out valid. More than 99% of all the indeterminate results were caused by a low positive control, and most likely, they are attributed to an impaired T-lymphocyte response because of suppressed host immunity.

6.4. SCREENING INDICATION

A limitation in both study II and III is the lack of screening indication when performing the QFT. In study II the group consisting of persons with a QFT without TB disease was heterogeneous and comprised individuals tested for different reasons. Because we did not have access to clinical information, we could not determine if the immunosuppression seen with advanced age in study II was a genuine effect of old age or an effect of age related comorbidity (2). Similarly in study III, we did not have clinical data on the actual reason for testing, so calculation of the PPV and NPV stratified by risk group was not possible. However, in a very large cohort we did provide general confidence limits for both the PPV and the NPV of a QFT per see.
widespread use of the test in active TB diagnosis, in a recent Danish study from Aarhus, Danielsen et al. evaluated the use of the T-Spot test from 2010-2011 (124). Suspicion of active TB was the indication for screening in 32.7% (533/1,631) of persons tested. Another study evaluated test indication in 2005-2006 at Hvidovre Hospital, and Browatzki et al (125) found active TB to be the indication in 81% (74/91) of all IGRA performed.

These findings indicate that the test is used to screen for active TB. This is not necessarily misuse of the test, but according to guidelines IGRA rarely have an added value in combination with standard methods for diagnosing active TB (67) indicating at least some misuse of the test. Based on the studies in the current thesis and available Danish literature, there seem to be a need for better adherence to guidelines and improved information regarding IGRA limitations. Although this is not a big problem in Denmark with the majority of TB cases being confirmed by culture, it is increasingly a concern globally (126).

Preventive treatment

According to Danish guidelines (33), the standard treatment for latent TB infection is a 6-month course of isoniazid monotherapy. Diagnosis of latent TB and prescription of preventive treatment is not a notifiable event, and therefore no summarized data is available. The total coverage of isoniazid preventive therapy among individuals with a positive QFT was estimated to be 35% (data not shown) and among patients with a positive QFT who developed incident TB, 90% did not receive preventive treatment (3). It follows that adherence to guidelines concerning screening indication and preventive treatment in Denmark is suboptimal.

7.2. THE GLOBAL PERSPECTIVE, IGRA USE WORLDWIDE

IGRA use in TB high-incidence countries

In TB high-incidence settings, IGRA use worldwide has limitations because of their extensive laboratory requirements and a high prevalence of LTBI in the population. Consequently, WHO has issued different guidelines on the management of latent tuberculosis infection in countries with a TB incidence rate below 100 per 10^5 population (32) compared to TB-highincidence countries (127,128). In TB high-incidence countries, WHO recommends to reserve the use of IGRA for HIV infected individuals (127) and children below the age of five (128), both groups being vulnerable groups with increased risk of progression to disease.

A novel RD-1 specific skin test has a potential use in resource-limited settings because it does not require a laboratory. The C-Tb (129) contains the same mycobacterial antigens as the QFT and has similar specificity and sensitivity, but still share the IGRA restraints in settings with high prevalence of LTBI.

Use and misuse of the test among patients suspected of active TB

The off-label use of IGRA in the diagnosis of active TB could potentially result in overtreatment. In a recent article from India (130), Little et al. constructed an analytical decision model and found an estimated 315,700 additional false-positive persons would be diagnosed with TB each year at an incremental cost of US$49.3 million. This finding supports recent policies by WHO that now discourage the use of IGRA for the diagnosis of active TB in settings like India (126) and the misuse of the IGRA in high-incidence low income countries is currently receiving global attention. Another problematic tool when diagnosing TB is the commercial, serological tests for recognition of Mtb antigens by the humoral immune response (as opposed to the cellular response used by the IGRA). They are marketed in many third world countries, although none of them are recommended by official guidelines or FDA approved. They are not superior to the IGRA based on the performance reported in study II and recent meta analysis (63). Overall data quality in the studies of serological tests for active TB are regarded as very low, and official recommendations strongly advice against serological tests (131). In summary, even though this is an unjustifiable malpractice, many immunological tests are marketed for diagnosing active TB, especially in high-burden countries with weak regulatory systems.

7.3. THE FUTURE

Many questions regarding IGRA and their role in TB control and ultimately TB elimination in low-incidence settings remain unanswered. Data from this thesis suggest that a negative IGRA is valuable in ruling out TB with little...
risk of progression to disease, and that the QFT has a high sensitivity for TB and a high specificity when used among the majority of patients with NTM disease. Critics often pinpoint the low sensitivity and poor predictive value of the IGRAs, and it is indeed not a perfect test. However, currently IGRAs are the best available option in low-incidence countries like Denmark, and they are important in diagnostic algorithms as tools to identify persons with active TB and LTBI.

When the IFN-γ based assays were introduced, it was anticipated that they could provide the means to distinguish active from latent TB and identify those at the highest risk of progression. After more than a decade of studying and reporting on IGRA performance, it has become clear that none of the commercial IGRAs have met these expectations. One could speculate whether IGRAs have reached their full potential, and if the search for a better biomarker should be prioritized rather than trying to improve existing assays based on release of IFN-γ (77).

The discovery of a novel biomarker that is both highly specific, sensitive and with good predictive value would be a game changer in TB control in low-incidence countries. In recent years many new biomarkers have emerged (79-81). One example is the IP-10 marker and the recent development of a dried blood spot that enables easy storage and prolonged transport (132). IP-10 has similar properties in separating NTM and TB, and might hold potential in making immunological diagnosis of TB and differentiation between TB and NTM accessible in low-income, TB high-incidence settings (133). However, the IP-10 marker is not commercially available and has not proven superior compared to existing tests. Thus, until better biomarkers are introduced, it is important to use the current available tests, while being aware of the limitations. This is the focus of an ongoing debate on how screening strategies using IGRAs in high-income TB low-incidence countries should be organized to achieve the most optimal results towards TB control and elimination (113,134,135).

The burden of TB remains in third world countries, and immunological assays, although important in a Danish setting, is not useful for combating the disease measured in terms of global morbidity and mortality. Rather, the same basic but essential countermeasures are as important today as they were 100 years ago.

"The disease is dying a natural death with improved conditions of the working classes, and it is by further developments on such lines, and not otherwise, that its extermination will be attained."

Dr. T. D. Lister (Mount Vernon Hospital for Consumption, Hampstead, early 20th century).
SUMMARY

In the past decade, the Interferon Gamma Release Assays (IGRAs) have been increasingly used in tuberculosis (TB) low-incidence countries for the diagnosis of both active and latent TB. We evaluated test performance of the most widely used IGRA, the QuantiFERON-TB Gold In-Tube test (QFT), in a large cohort through 6 consecutive years from 2005 through 2010, and included all Danish patients with notified TB and cultured non-tuberculous mycobacteria (NTM).

It would be of great clinical value if the QFT could serve as a minimally invasive tool and aid in discriminating TB disease from NTM disease. Disease caused by the majority of NTM will not evoke a positive IGRA because they do not cross-react with the mycobacterial antigens in the QFT. We conducted an evaluation of IGRA performance in patients with NTM disease, and in light of these result and available literature, we suggest that the IGRAs could be useful in discriminating between TB and NTM disease. Among NTM patients, we found a high specificity of 96%, a finding which is especially relevant in children, where lymphadenitis due to mycobacteria from the M. avium complex are seen to predominate and can mimic TB lymphadenitis.

The ability of the QFT to identify infection with Mtb among patients with active TB was in line with other studies and the overall sensitivity was 86%. Performance was poorer in the extremes of age with lowered sensitivity among elderly (aged >65 years) and increased indeterminate rates in infants <1 year and elderly >65 years. We conclude that test performance is impaired only in children aged less than 1 and elderly, and suggest revision of current guidelines advising caution when using IGRAs in children below the age of 5.

A pivotal measure for any diagnostic test is the prediction of disease. We conducted the largest study to date in a TB low-incidence setting and determined the negative predictive value (NPV) and positive predictive value (PPV) of the test, identifying incident TB cases during more than 50,000 person-years of follow-up. PPV was 1.3% and NPV 99.85%. We conclude that a negative QFT in our cohort of persons that were screened due to various indications predicts an extremely low probability of developing TB disease. The PPV was expectedly low, but a positive test still inferred a more than 8 times increased risk of developing TB, with an additional increase following the first 2 years after QFT testing.

With the current studies, we have established and used a comprehensive, nationwide cohort to provide novel insights into the performance of the QFT. In summary, the test is used extensively in Denmark both for active and latent TB. Overall test performance is acceptable, but caution is warranted in infants and elderly and our results suggest that administration of preventive therapy for latent TB could be improved. Screening indication, host susceptibility, as well as a priori risk of TB remain vital factors to consider both when deciding to test and when interpreting test result.
DANSK RESUME

I lande med en lav forekomst af tuberkulose (TB) påvises infektion med tuberkulosebakterien i stigende grad ved hjælp af interferon-γ release assays (IGRAs). I denne ph.d. evaluerede vi den mest udbredte af disse tests, QuantiFERON-TB Gold In-Tube testen (QFT) i en kohorte bestående af alle personer med en udført test i perioden 2005 til og med 2010 i Danmark. Vi kombinerede QFT data med alle anmeldte tilfælde af TB samt alle tilfælde af dyrkningsverificeret sygdom forårsaget af gruppen af non tuberkuløse mykobakterier (NTM) og undersøgte værdien af QFT til at påvise disse sygdomme.

Det ville være en fordel hvis man kunne bruge QFT til at skelne aktiv tuberkulose fra sygdom grundet NTM, da en QFT er mindre invasiv end traditionel diagnostik ved biopsitagning. Størstedelen af NTM vil give et negativt QFT resultat, da de ikke indeholder de specifikke antigener som testen er baseret på. Vi har evalueret QFT blandt patienter med NTM sygdom og fandt en meget høj specificitet. Vi har ydermere gennemgået tilgængelige studier vedrørende IGRA og anvendelse hos patienter med NTM sygdom og samlet set fandt vi, at testen har potentielle til at skelne mellem TB og NTM. NTM sygdom er sjældent forekommende, men hos især børn kan testen være relevant, da en typisk præsentation af NTM sygdom er en hævet lymfeknude på halsen, som klinisk kan have lighed med glandel TB.

QFT evne til at påvise infektion med tuberkulosebakterien blandt patienter med aktiv TB i vores kohorte var sammenlignelig med andre studier, og samlet set var sensitiviteten 86%. QFT var påvirket i begge ender af aldersspektret, idet testen havde flere inkonklusive resultater blandt både spædbørn yngre end 1 år og ældre over 65 år. Endvidere havde testen nedsat følsomhed blandt ældre >65 år. Vi konkluderer, at testen synes at virke godt fra 1 år og opefter, men at man bør udvide forsigtighed når testen anvendes hos spædbørn og ældre over 65 år, da præstationen er nedsat blandt disse grupper. Dette er interessant i et internationalt perspektiv, idet flere retningslinjer fraråder testen hos børn under 5 år, hvorfor en revision af disse bør overvejes.

En central egenskab ved en diagnostisk test er dens evne til at forudsige udvikling af sygdom. I denne ph.d. har vi lavet den hidtil største undersøgelse af QFT’s evne til at forudsige udvikling af TB, og vi har bestemt den prædiktive værdi af en positiv og negativ test i løbet af mere end 50,000 personårs opfølgning. Vi konkluderer, at den positive prædiktive værdi er 1.3% og den negative prædiktive værdi er 99.85%, hvilket betyder at en person der screenes, og har en negativ test, iht. vores data har en meget lille risiko for at udvikle TB. Hvis man har en positiv test, har man mere end 8 gange større risiko for udvikling af TB.

REFERENCES


(13) Modlin RL, Bloom BR. TB or not TB: that is no longer the question. Sci Transl Med 2013 Nov 27;5(213):21


(89) Park SY, Jeon K, Um SW, Kwon OJ, Kang ES, Koh WJ. Clinical utility of the QuantIFERON-TB Gold In-Tube test for the


(126) World Health Organization Country Office for India. Standards for TB Care in India . 2014.


