Tuberculosis and diabetes: bidirectional association in a UK primary care dataset

Fiona Pearson1
Peijue Huangfu1
Richard McNally2
Mark Pearce2
Nigel Unwin3
Julia Critchley1

1 Population Health Research Institute, St George’s University of London, London, UK
2 Institute of Health and Society, Newcastle University, Newcastle, UK
3 MRC Epidemiology Unit, University of Cambridge, Cambridge, UK

Corresponding author:
Fiona Pearson,
Population Health Research Institute, St George’s University of London, Cranmer Terrace, London SW17 0RE
fpearson@sgul.ac.uk

Key words: Tuberculosis, Diabetes, Co-morbidity, Epidemiology, Public Health

Word Count: 2997
Abstract: 248
Reference: 41
Abstract

Background
Many studies have found an increased risk of pulmonary tuberculosis (PTB) amongst those with diabetes mellitus (DM). However, evidence on whether the association is bi-directional remains sparse. This study investigates DM rates amongst those with and without prior tuberculosis (TB) disease as well as the reverse.

Methods
Data on a UK general practice population, between 2003 and 2009, were obtained from The Health Improvement Network database. A series of retrospective cohort studies were completed. Individuals were successively classified as ‘exposed’ or ‘unexposed’ to TB, PTB, extra-pulmonary TB (EPTB), or DM. Multivariate negative binomial regression was used to calculate incidence rate ratios (IRR) amongst each exposure group for outcomes of interest (TB, PTB, EPTB, or DM in turn) adjusting for plausible confounding variables (age, sex, region, Townsend quintile and smoking status). Potential confounding due to ethnicity was adjusted for using McNamee’s external method.

Results
DM risk was substantially raised amongst individuals with a history of TB disease (IRR 5.65 (95% CI 5.19-6.16)), PTB (IRR 5.74 (95% CI 5.08-6.50)), and EPTB (IRR 4.66 (95% CI 3.94-5.51)) compared to those without; results were attenuated after external adjustment for ethnicity (IRR 2.33 (95%CI 2.14-2.53)). TB risk was raised modestly amongst individuals with DM (IRR 1.50 (95%CI 1.27-1.76)), and was attenuated slightly after adjustment for ethnicity (IRR 1.26 (95%CI 1.07-1.48)).

Conclusion
DM risk was raised among those with previous TB disease; this finding has implications for follow-up and screening of TB patients, who may be at high risk of developing DM or related complications.
Introduction

Despite consolidated control efforts, the TB burden remains high in many places in the world. In 2016 an estimated 10.4 million people were diagnosed with TB and 1.7 million died from the disease(1). Global diabetes (DM) prevalence has been steadily increasing in recent decades. Estimates produced by the International Diabetes Federation suggest that by 2045 a 55% increase in DM burden will occur worldwide, with the greatest increases projected to occur in regions where TB remains endemic(2).

An association between TB and DM has long been known although its importance has been under recognised(3-5). Two systematic reviews published nearly 10 years ago revitalised awareness of this link showing that DM increases TB risk by three-fold(3, 6). This association has since been further confirmed by more contemporaneous studies conducted in both TB endemic regions and developed countries(7-11) but not always by those completed in low TB burden countries(12). To date, most studies of the association have been cross-sectional or cohort in design with DM as the “exposure” and TB as the “outcome”.

A bi-directional association between TB and DM is biologically plausible(13, 14), but there is limited evidence to support or refute whether people with a history of TB are at increased risk of developing DM(8). It remains important to identify epidemiologically whether the increased risks of TB seen among DM patients are truly bidirectional, and clinically whether patients with TB might benefit from future screening for DM. TB patients often develop hyperglycaemia during TB treatment, though this may be transient(10). Whether or not future risk of DM is raised for those with “transient hyperglycaemia” is unknown, though logical(15).

It is generally thought that DM increases the risk of pulmonary TB (PTB) rather than extra-pulmonary TB (EPTB)(3, 16). However, there is limited available evidence at present on associations between different TB manifestations (PTB, EPTB) and DM subtypes (T1DM, T2DM)(8). Most studies are underpowered to explore these potential associations as EPTB accounts for less than 20% of global TB and only around 10% of DM is T1DM, thus it is difficult to confirm cases with diagnosis of both in many low and middle income countries (LMIC) where TB burden is high. The Health Improvement Network (THIN) – a very large, widely validated and used administrative dataset of UK primary care health data (17)—was used in this study to facilitate the investigation in this research area.
Methods

Ethical Approval

Data collection for THIN was approved by the South East Multicentre Research Ethics Committee (MREC) in 2003. This individual study did not require separate ethical approval as only anonymised THIN data is used, however it was reviewed by THIN independent Scientific Review Committee.

Data source, definition of exposures and outcomes

Data were obtained from the THIN database which contains the electronic medical records of 6.9 million patients collated over 385 GP practices in the UK, covering approximately 5.7% of the total population in a representative manner\(^{(17)}\). The study population comprised all individuals in THIN database with active records between 2003 and 2009 that met THIN minimum data acceptability standards at both practice and patient level with no gaps in practice records. All patients were actively registered either currently or historically.

Each exposure cohort (DM, TB, PTB, and EPTB) was constructed in the same way, dynamically with individuals selected in from the date of their first record on file for said exposure between the 01/01/2003 and 31/12/2009. Each individual’s files were then searched for a subsequent record of an outcome of interest. Individuals left each cohort either on the study end date (31/12/2009), on their date of death, or GP de-registration. Unexposed cohorts were constructed from aggregated denominator data provided by CSD-Medical Research UK, the gatekeepers to THIN at the time this study was completed. Records from the THIN database undergo in-house data quality checks and only data flagged as passing these checks were used. Further data quality checks were completed to ensure no implausible values existed for any variable of interest and that date of birth, registration and death occurred sequentially.

Using a combination of hierarchical Read codes, individuals were categorised as having: TB, PTB, EPTB, or DM. TB diagnosis codes refer to TB disease, and do not include latent TB infection. When there was uncertainty on how a Read Code should be categorised, a third opinion was sought by a chest physician as appropriate (see Appendix 1—6). The majority of, but not all, individuals within the TB cohort are subcategorised in to the PTB or EPTB cohort. Data on key confounders including age, sex, socioeconomic status (SES) based on Townsend...
Quintile, a score of material deprivation linked to each individual's postcode, and smoking status were obtained from the appropriate THIN data fields.

**Statistical Methods**

Univariate analyses were conducted to identify which variables were associated with both TB and DM. Results were calculated as IRRs with 95% CI and an alpha value of $p<0.05$ was used as standard. Following this, multivariate analyses using a negative binomial model were completed to explore whether DM was associated with TB, PTB or EPTB, and the converse accounting for plausible confounding. Ethnicity could be an important confounder of the explored relationships, in particular where TB is the exposure and DM the outcome. Unfortunately, ethnicity has been poorly recorded in UK primary care datasets. Therefore, we have externally adjusted for ethnicity as a confounder of the bidirectional association between TB and DM using a method outlined by McNamee (18); the detailed method can be found in Appendix 7. Individuals with missing data for main variables of interest were excluded from datasets since the total number with missing data was small (<1%).
Results

Within the THIN dataset there were 224,508 individuals in the DM cohort, 5,470 in the TB cohort, 1,589 in the PTB cohort; and 1,006 in the EPTB cohort. The mean age of entry into the DM cohort was 60 years, TB cohort 48 years, PTB cohort 53 years, and EPTB cohort 47 years respectively. Average follow up for each cohort was 4.1 years, 3.7 years, 3.7 years, and 4 years respectively. Sex, age, SES, and smoking distribution for the DM and TB cohorts are given in Table 1.

The rate of DM was significantly increased amongst those who had previously had TB (all subtypes) (IRR 5.65 (95% CI 5.19-6.16)). This was also the case for both PTB (IRR 5.74 (95% CI 5.08-6.50)) and EPTB (IRR 4.66 (95% CI 3.94-5.51)), compared with those who had not had TB (see Table 2). We estimated that IRRs may be inflated by about 2 to 2.5 fold due to residual confounding by ethnicity (Appendix 7), suggesting that after this external adjustment for ethnicity, the true IRR could be attenuated from 5.65 to around 2.33 (95%CI 2.14-2.53) up to 2.90 (95%CI 2.66-3.16).

The IRR of TB was significantly increased amongst individuals with DM (IRR 1.50 (95%CI 1.27-1.76)) compared to those without (see Table 2). The rate of PTB amongst individuals with DM (IRR 1.24 (95%CI 0.93-1.64)) and the rate of EPTB amongst those with DM (IRR 1.43 (95% CI 0.99-2.07)) when compared to those without DM was not significantly raised. External adjustment for ethnicity attenuated the estimates of DM amongst those with TB by 1.2 to 1.4 fold, see Appendix 7.
Discussion

The key finding of this study is the increased risk of DM following TB; over a five-fold increased risk of DM in a multivariate analysis, possibly reducing to around a two-three fold increased risk after external control of confounding. As far as we are aware, no other cohort studies has been able to adequately quantify the risk of DM among people who have had TB; although many experts have speculated that this risk might be bi-directional (13, 19). Whilst studies in high TB incidence countries are ideally needed to confirm this risk and its population importance, they may be difficult to conduct in practice, due to the large sample size (and length of follow-up) required. Thus, other studies using big datasets, more routinely available in the developed world, should be analysed to confirm this association.

Although data is limited it could be biologically plausible that the risk of developing DM is higher after TB disease. Some recent animal models of Mtb have seen unexpected development of DM (e.g. among guinea pigs), particularly those treated with anti-glycaemic therapy as host-directed therapy for TB (20). Studies have shown induced hyperglycaemia and/or impaired glucose tolerance occur during the early phases of active TB disease (21, 22), and these metabolic states themselves are linked with progression to overt DM amongst 20-50% of individuals after three to five years (23). In a recent Indian cohort study, TB patients with newly diagnosed DM had markedly lower HbA1c values (although still abnormal) compared to those with known DM, suggesting that TB might, at least transiently, stimulate progression from intermediate hyperglycaemia to frank DM or identify those individuals who may be more prone to metabolic alterations or DM in the future (24). Thus, TB disease may identify individuals at higher risk of progression to DM, in much the same was as gestational DM identifies groups at high risk of progression to overt DM. Another possible mechanism by which TB may increase DM risk is through changes in body composition during and following the illness. TB patients frequently lose a substantial amount of weight before and in early stages of treatment; limited evidence from cohort studies suggests that weight re-gain during treatment could increase the proportion of body fat in recovered TB patients (25, 26); hence increasing their future risk of DM.

This study found modest associations between DM (as exposure) and TB (as outcome) after controlling for key confounders (15). However, the novelty of this analysis is the large sample
size and hence power to show an increased risk of EPTB as well as PTB (of similar magnitude; 1.4 to 1.5 fold increases). Some studies from LMIC have suggested that only PTB risk is increased among DM patients, which may be explained by a cell-mediated immune response in DM patients inhibiting dissemination outside the lung(27). In the UK, most TB is likely to be due to re-activation compared with LMIC, so it is possible that in the UK individuals may have seeds of Latent Mtb at extra-pulmonary sites that occurred prior to the onset of DM, and suggest that immune containment of Mtb is generally deficient among people with DM.

The magnitude of the increased risk of TB in people with DM (about 1.3 fold increased risk over both sub-types) is lower than that first reported in systematic reviews(28, 29), but in line with more recent and larger studies, particularly those in developed countries including one recently published from the Clinical Practice Research Database (CPRD) in the UK(11, 30). The slightly weaker association in these recent studies may reflect a strong primary care health system with relatively good glycaemic control and management of co-morbidities among patients with DM(11). This hypothesis is supported by the stronger association (about a 2.5-fold association) found in the early NHANES II study (1978-80), even after control for key confounders(7), and several recent studies suggesting stronger associations between DM and TB risk where glycaemic control is poor(31).

Our finding of an increased DM risk after TB needs to be interpreted cautiously. Some studies have noted that the hyperglycaemia observed amongst populations of individuals with TB is transient, reversing after the early, acute phase of TB infection(22, 32), although others have suggested it is permanent(33). It could be that some of these observations of hyperglycaemia are due to short-term side effects of treatment with rifampicin and isoniazid(34), or, to stress hyperglycaemia rather than being signs of true metabolic dysfunction and DM. Our findings by TB sub-type also need to be carefully assessed as many patients did not have TB type classified; though most likely these are PTB cases.

Diagnosis with T2DM is often delayed, even in developed countries with strong primary health care systems such as the UK(35). With the majority of DM cases in the UK being T2DM we therefore cannot be certain if the TB disease occurred prior to DM onset, even among those individuals without a previous diagnosis of DM. Longer follow-up time periods are ideally required to fully tease out whether TB disease is itself increasing the risk of DM or whether it
is a more a presenting feature of DM. However, T1DM is usually much more rapid in onset and should be quickly diagnosed in the UK (36), so possible reverse causality is less likely when T1DM is diagnosed after having TB disease. Given the lack of any screening programmes for TB and DM, the effect of misclassification may be non-differential; biasing estimates towards the null. However, even if uncertainties concerning the mechanism and direction of the association remain, our finding of an increased risk of DM among former TB patients could have important clinical and public health implications. Lifetime risk of developing DM is already very high in many TB endemic countries (2); assuming the increased risk among TB patients is real, it could turn out to be highly cost-effective to screen former TB patients at regular intervals for DM; thus reducing the risks of expensive complications and improving long-term health outcomes from DM. TB patients with DM also appear to be more infectious (having both a higher bacterial load, and remaining smear positive for longer) than patients without DM, and also experience more recurrent TB (3, 37). These factors also suggest TB patients at risk of, or already with, overt DM may be driving continued TB transmission disproportionate to their population size and may require different treatment guidelines (38).

The key strengths of our analyses are the true cohort design and size of the database. The very large size, several orders of magnitude larger than most of studies on this topic, has allowed us to assess the associations with disease subgroups (PTB and EPTB) and also the risk of TB disease leading to future DM; topics barely covered (sub-types) or assessed at all (TB leading to DM) in previous analyses. Our study was based on routine health care records collated in primary care. DM is routinely managed in primary care, and the UK GP Quality and Outcomes Framework (implemented after 2004) incentivised GPs to diagnose and treat DM patients (39), thus DM status was highly likely being recorded accurately (40). Our dataset may slightly underestimate the true incidence of TB as care will mainly be instigated and completed by secondary providers. However, we believe that in most cases TB relevant information will be fed back to primary care, as TB is a serious, notifiable infection. Supporting this assumption is that the rates of TB identified in THIN were comparable to UK rates (TB incidence of 14.6 per 100,000 in 2009) meaning it is unlikely that misclassification is occurring to a great extent.

Although we adjusted for key confounders (SES, age, sex, smoking status) (7, 11), our estimates could still be affected by residual confounding. We were unable to adjust directly
for other confounders that could be of interest (such as ethnicity or BMI) with this dataset.

Our previous analyses using the NHANES dataset from the US did not find substantial
attenuation of associations after adjusting for a more comprehensive range of potential
confounders(7), and another recent analysis of data from primary care in the UK using the
CPRD database also found no effect modification from ethnicity, age, or duration of DM(11).
BMI is associated with the development of DM and TB; besides, TB patients also lose weight
which is rapidly regained on treatment. Hence even single measurements of BMI may not be
sufficient to adjust for potential confounding effects. More detailed analyses, potentially
treating BMI as a time-varying exposure, are thus required, but are not currently available.

We acknowledge that the potential for residual confounding by ethnicity is clearly important
since DM risk is substantially increased in Black Afro-Caribbean and South Asian populations,
compared with White British people, and the majority of TB cases reported in the UK also
occur in these minority groups. For this reason we performed an external adjustment for
ethnicity, which suggested that the true RR for DM among those with prior TB could be
reduced from 5.65 to around 2.33-2.90, but still remained statistically significant and clinically
important.

Our analyses are the first globally to cautiously add epidemiological support to a hypothesis
that the association between TB and DM could be bi-directional. This may have implications
for future health care for people with a history of TB disease, who are not routinely assessed
for DM or informed that they may be at higher risk of DM in the future. WHO guidelines
recommend screening all TB patients for DM globally, but these are not yet routinely
implemented; moreover, it is unknown what the long term health supports are needed for
former TB patients subsequently developed hyperglycaemia. Whilst this may be of most
importance for population health in high TB-DM burden countries, it is also important among
ethnic minority groups in low burden countries such as the UK, who may be at higher risk of
both TB and DM(41).
Contributions of authors: NU and JAC originally conceived the ideas for the manuscript. FP performed the analyses, JAC and PH helped with analyses of confounding by ethnicity. FP wrote the first draft of the manuscript with input from JAC and PH. MP, RM, PH and NU edited and critically appraised the manuscript.
Acknowledgements: We thank Dr Chris Stenton (Chest Physician) for his help and expert opinion in identifying the TB Read codes. We thank the database provider THIN for making data freely available to us under a student license.

Competing Interest: None declared.

Licence for Publication: The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in JECH and any other BMJPGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence (http://group.bmj.com/products/journals/instructions-for-authors/licence-forms).


Sources of Support:

This publication was made possible by NPRP grant # 7-627-3-167 from the Qatar National Research Fund (a member of Qatar Foundation). The findings achieved herein are solely the responsibility of the authors. Julia Critchley is also funded by the Higher Education Funding Council for England. Newcastle Arch-Epi group for their support in aggregating data. The funders had no role in the design, conduct, or analysis of the study.
What is already known on this subject?

It is known that having diabetes increases your chance of developing active TB. However, although plausible, it was not known if having TB increase your future risk of developing DM.

What this study adds?

This study shows that having had TB individuals are at an increased risk of developing DM. It is thus important to consider follow up, in particular screening, for such long term complications.
References


388 41. Walker C, Unwin N. Estimates of the impact of diabetes on the incidence of pulmonary
390
391
**Figure 1 Flowchart of cohorts within study population**

Population at risk
Patients active between 2003-2009 that met THIN minimum data acceptability standards, flagged A or C

Exposed n=270,505
Patients active between 2003-2009 that met THIN minimum data acceptability standards, flagged A or C, with a MED or AHD code for DM or TB

Unexposed n=22,019,934
Patients active between 2003-2009 that met THIN minimum data acceptability standards, flagged A or C

Exposed n=229,978
Remove anyone that dies or transfers out before 2003, or that has implausible values for age or sex.

DM cohort n=224,508
MED or AHD code for DM

TB cohort n=5,470
MED or AHD code for TB

PTB cohort n=1,589
MED or AHD code for PTB

EPTB n=1,005
MED or AHD code for EPTB
Table 1 Sex, age, SES and smoking distribution within all cohorts

<table>
<thead>
<tr>
<th>Variables</th>
<th>DM Cohort</th>
<th>TB Cohort</th>
<th>PTB Cohort</th>
<th>EPTB Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>(%)</td>
<td>Number</td>
<td>%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>224,508</td>
<td>100</td>
<td>5,470</td>
<td>100</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>123,264</td>
<td>54.9</td>
<td>2,698</td>
<td>49.3</td>
</tr>
<tr>
<td>Female</td>
<td>101,244</td>
<td>45.1</td>
<td>2,772</td>
<td>50.7</td>
</tr>
<tr>
<td><strong>Age Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-15</td>
<td>3,045</td>
<td>1.3</td>
<td>394</td>
<td>7.2</td>
</tr>
<tr>
<td>16-30</td>
<td>8,696</td>
<td>3.9</td>
<td>1,035</td>
<td>19</td>
</tr>
<tr>
<td>31-45</td>
<td>29,576</td>
<td>13.1</td>
<td>375</td>
<td>19.3</td>
</tr>
<tr>
<td>46-60</td>
<td>63,977</td>
<td>28.5</td>
<td>1,109</td>
<td>20.3</td>
</tr>
<tr>
<td>61-75</td>
<td>80,235</td>
<td>35.7</td>
<td>1,268</td>
<td>23.2</td>
</tr>
<tr>
<td>76+</td>
<td>38,979</td>
<td>17.4</td>
<td>584</td>
<td>10.7</td>
</tr>
<tr>
<td><strong>Townsend quintile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>45,608</td>
<td>20.3</td>
<td>774</td>
<td>14.1</td>
</tr>
<tr>
<td>2</td>
<td>44,518</td>
<td>19.8</td>
<td>800</td>
<td>14.6</td>
</tr>
<tr>
<td>3</td>
<td>45,121</td>
<td>20.1</td>
<td>1,017</td>
<td>18.6</td>
</tr>
<tr>
<td>4</td>
<td>43,283</td>
<td>19.3</td>
<td>1,193</td>
<td>21.8</td>
</tr>
<tr>
<td>5</td>
<td>32,502</td>
<td>14.5</td>
<td>1,252</td>
<td>22.9</td>
</tr>
<tr>
<td><strong>Smoker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>62,288</td>
<td>27.7</td>
<td>1,449</td>
<td>26.5</td>
</tr>
<tr>
<td>Past</td>
<td>119,619</td>
<td>53.3</td>
<td>2,236</td>
<td>40.9</td>
</tr>
<tr>
<td>No</td>
<td>42,601</td>
<td>19.0</td>
<td>1,785</td>
<td>32.6</td>
</tr>
</tbody>
</table>
### Table 2 IRRs, 95%CI and P-values for the occurrence of DM amongst people with TB

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Unadjusted IRR (95%CI)</th>
<th>P-value</th>
<th>Adjusted IRR (95%CI)(^1)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB</td>
<td>DM</td>
<td>6.38 (5.69-7.16)</td>
<td>&lt;0.001</td>
<td>5.65 (5.19-6.16)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PTB</td>
<td>DM</td>
<td>8.03 (6.90-9.34)</td>
<td>&lt;0.001</td>
<td>5.74 (5.08-6.49)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>EPTB</td>
<td>DM</td>
<td>5.42 (4.46-6.58)</td>
<td>&lt;0.001</td>
<td>4.66 (3.94-5.51)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DM</td>
<td>TB</td>
<td>1.62 (1.37-1.92)</td>
<td>&lt;0.001</td>
<td>1.50 (1.27-1.76)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DM</td>
<td>PTB</td>
<td>1.74 (1.30-2.32)</td>
<td>&lt;0.001</td>
<td>1.24 (0.93-1.64)</td>
<td>0.137</td>
</tr>
<tr>
<td>DM</td>
<td>EPTB</td>
<td>1.56 (1.09-2.24)</td>
<td>0.016</td>
<td>1.43 (0.99-2.07)</td>
<td>0.055</td>
</tr>
</tbody>
</table>

\(^1\)Adjusted for age, sex, region, Townsend score and smoking status.