Explaining the one third reduction in tuberculosis incidence in the UK, 2011-2015

Appendices (supplementary material)

A1: Imputation process for missing year of entry

For the 6.4% of TB cases with a missing year of entry to the UK, a linear model was formulated to predict year of entry on the basis of year of notification, age group and country of birth. Countries with less than 30 cases overall were grouped according to the WHO incidence groups as described in the main paper. All 2-way interactions and the 3-way interaction for the three variables were considered, and models compared via the BIC (Bayesian Information Criteria). The model with the lowest BIC included all 2-way interactions, and was used to generate predicted mean year of entry and forecast standard errors (SE) for the missing data. Imputed values were then randomly generated based on the predicted means and SEs, with the restriction that values for the year of entry could not be (a) greater than the year of notification, and (b) before the earliest birth year that the individual could be within for their age group. Invalid imputations were thus re-imputed until all missing values had valid predictions.

Imputation of missing values is typically conducted multiple times, with analyses being performed on each imputed dataset and the within- and between-imputation variability accounted for. This was considered an unnecessary complication in our analysis. Firstly, the amount of missing data was small. Secondly, data at the country/WHO incidence group consist of large counts, and at the aggregate level the individual difference between imputations would be averaged over. Thirdly, any minor differences that might occur at the aggregate level under random imputation at the individual level are far outweighed by uncertainties in the length of stay distributions that underpin our results.

The imputation of missing year of entry (compared to discarding cases with missing values) should help to minimise any bias induced by patterns of missing data in our results. Those in the oldest age group were more likely to be missing year of entry than younger ages, and tend to have earlier years of entry, which could distort the distribution of length of stay. In addition, the proportion of cases with missing data has decreased over time, from 8.7% in 2011 to 4.3% in 2015. By including only the complete case data, there would be a smaller decline in cases over the 2011-2015 period. Of course, the possibility of missing not at random (MNAR) data cannot be discounted, in that there are systematic differences in the missing values that cannot be recovered from the relationships with other variables in the observed data. Nevertheless, taking into account all considerations the imputation process is more likely to provide unbiased results.

A2: Derivation of denominators

In order to calculate TB rates in the UK per 100,000 person-years for different migrant groups (*recent*, *mid-term* and *long-term*/resident) denominators are required. Although the total number of non-UK born individuals from different countries are available from the Office for National Statistics (ONS), these data do not include information on how long migrants have been in the UK for. We therefore used information on the number of new migrants entering the UK each year, based on visa and asylum applications, and distributions of length of stay. In general, the number of people in a particular group is therefore obtained via the formula:

$$N_{g,y} = \sum_{t=L}^{U} M_{y-t} P(y,t)$$

Where $N_{g,y}$ is the number of migrants with length of stay group g in year y, M_{y-t} is the number of new migrants that arrived in year that arrived in year y-t, $P(y,t)$ is the probability that a migrant will still remain in year y after t years , and L and U are the lower and upper bounds for migrant group g; e.g., for mid-term migrants L=3 and U=7.

For example, in 2013 the *recent* migrant group will consist of new migrants that arrived from 2011-2013 (that have remained in the UK), the *mid-term* migrant group will consist of migrants that arrived between 2007-2010 (and remained in the UK) and the *long-term* migrant group is defined by the total number of people living in the

UK from that country minus the two other groups. Shown below is table illustrating the relationship between current year, year of entry (*y*-*t*) and length of stay (*t*).

Year of entry	Current year				
	2011	**2012**	**2013**	**2014**	**2015**
2005	6				
2006	5	6			
2007	4	5	6		
2008	3	4	5	6	
2009	2	3	4	5	6
2010	1	2	3	4	5
2011	0	1	2	3	4
2012		0	1	2	3
2013			0	1	2
2014				0	1
2015					0

Length of stay distributions were based on UK entry clearance visas and applied as described in Aldridge et al [1]. Data were available according to visa type (work, study, family/settlement) and year; it is therefore assumed that lengths of stay distributions for each visa type were the same for each country/WHO incidence group. Family/settlement visas and asylum applications were assumed to remain permanently in the UK, although the impact of different assumptions was explored in sensitivity analysis.

Length of stay distributions were defined via categories of <3 months, 3-6 months, 6-12 months, and 1-2, 2-3,3-4 and 4-5 years. The derived denominators were expressed as expected person years to match the definitions of length of stay in the notifications data: in particular, cases that occur in migrants in their first year in the UK would have an average length of stay of 4 months, assuming entry and developing TB are independent events.

A3: Visa applications and asylum applications granted, 2005-2015

There have been overall declines in total numbers of visas granted and asylum applications over time, but with peaks around 2009-2011 for many countries (India, Pakistan, Bangladesh and 151-250 per 100,000 WHO incidence group). For all but the <40 and 40-150 per 100,000 incidence countries there has been a decline in total visas granted/asylum applications from 2011 to 2015, but declines have been in different visa/asylum types for different countries.

Permanent (family/dependent) visas granted have declined for all higher TB-incidence countries except India, with particularly marked declines for Somalia, while asylum applications have increased for all countries except Pakistan and Somalia. Work and study visas granted have also declined for many countries, with the number of study visas granted for applicants from Pakistan, Bangladesh and India showing large decreases, but the number of work visas for applicants from India showing small increases .

The composition of granted visa types also differs markedly: India is the largest single contributor to visa granted, but the majority are work visas. This is in contrast to Somalia, for which the majority of visas granted are permanent visas or asylum applications. This has some impact on the derived denominators for the cases: permanent visas are assumed to remain part of the UK population in subsequent years, whereas work and study visas are of generally shorter duration. The distributions used here indicate an average duration of around two and a quarter years for work/study visas.

A4: Pre-entry screening

Table A4. Numbers of cases testing positive for TB in pre-entry screening according to country of origin and year, and the estimated number of TB cases averted in the UK each year based on the proportion of visa applications that are granted, 2011-2015.

Country/ WHO incidence group per 100,000	Pre-entry screening cases									
	2011		2012		2013		2014		2015	
	Cases identified	UK cases averted	Cases identified	UK cases averted	Cases identified	UK cases averted	Cases identified	UK cases averted	Cases identified	UK cases averted
<40	0	0	0	0	0	0	0	0	0	0
40-150	0	0	0	0	1	1	162	151	59	54
151-250	17	14	20	16	44	34	50	40	70	55
>250	2	2	3	2	42	37	60	50	112	92
India	0	0	0	0	9	8	79	72	116	104
Pakistan	53	40	37	24	25	17	12	8	22	12
Bangladesh	11	7	7	5	6	4	3	2	2	1
Somalia	1	1	0	0	7	5	3	2	1	0
Total	84	64	67	47	134	106	369	325	382	318

A5: Predicted and observed changes in TB rates in recent and mid-term migrants based on WHO-estimated TB incidence in their country of origin

Table A5. Predicted TB cases in recent and mid-term migrants based on WHO-estimated incidence in their country of origin and observed TB cases (notifications in the UK plus cases averted through PES) for 2011 and 2015 and the observed/predicted ratio, 2011 and 2015. Incidence rate ratios (IRR) based on the WHO rates and observed IRRs are shown.

	2011			2015				
Country/ WHO incidence group per 100,000	WHO predicted UK cases	Observed	Obs/Pred Ratio	WHO predicted UK cases	Observed	Obs/Pred Ratio	WHO predicted IRR	Observed IRR
<40	61.3	39	0.64	46.3	31	0.67	0.91	0.95
40-150	337.6	234	0.69	314.6	242	0.77	0.89	0.99
151-250	248.9	328	1.32	152.1	176	1.16	0.95	0.84
>250	1094.8	341	0.31	579.2	290	0.5	0.84	1.35
India	880.2	930	1.06	469.3	507	1.08	0.9	0.92
Pakistan	385.1	436	1.13	243.8	216	0.89	0.99	0.78
Bangladesh	118.2	128	1.08	55.3	52	0.94	1	0.87
Somalia	55.9	142	2.54	28.6	52	1.82	0.99	0.71

A6 Sensitivity analyses

Table A6. Sensitivity analysis of total TB notifications, cases identified by pre-entry screening (PES), population sizes and estimated impact of changes in migration, TB rates and PES on TB notifications between 2011 and 2015 under shorter length of stay and longer length of stay assumptions.

LONGER MIGRANT LENGTH OF STAY (larger recent/mid-term migrant population, smaller long-term migrant population)

	2011			2015			Impact on notifications			
Population	**Notified TB**	**PES TB**	**Population (1000s)**	**Notified TB**	**PES TB**	**Population (1000s)**	**Migration**	**TB rates**	**PES**	**Total**
Recent/ mid-term migrants (non-EU)	2514	64	2354	1248	318	1922	-710 (-25.8%)	-420 (-15.3%)	-305 (-11.1%)	-1266 (-52.1%)
Long-term migrants (non-EU)	3366	0	2874	2513	0	3540	1066	-1420	0	-853 (-35.1%)
EU-born	308	0	2582	469	0	3202	100 (4.6%)	43 (2.0%)	0 (0.0%)	161 (6.6%)
UK-born	2221	0	54787	1751	0	55642	35	-497	0	-470 (-19.4%)

SHORTER MIGRANT LENGTH OF STAY (smaller recent/mid-term migrant population, larger long-term migrant population)

	2011			2015			Impact on notifications			
Population	**Notified TB**	**PES TB**	**Population (1000s)**	**Notified TB**	**PES TB**	**Population (1000s)**	**Migration**	**TB rates**	**PES**	**Total**
Recent/mid-term migrants (non-EU)	2514	64	1358	1248	318	1035	-922 (-34.8%)	-152 (-5.8%)	-305 (-11.5%)	-1266 (-52.1%)
Long-term migrants (non-EU)	3366	0	3682	2513	0	4260	676	-1251	0	-853 (-35.1%)
EU-born	308	0	2582	469	0	3202	100 (4.6%)	43 (2.0%)	0 (0.0%)	161 (6.6%)
UK-born	2221	0	54787	1751	0	55642	35	-497	0	-470 (-19.4%)

Methods

Length of stay distributions for migrant groups are stratified into categories of less than 3 months, 3-6 months, 6 months to 1 year, 1-2 years, 2-3 years, 3-4 years and 4 years or more. In the base analysis, those arriving in the

UK are assumed to stay for the upper bound of their visa duration: 3 or 6 months, 1, 2, 3, 4 or 5 years for study and work visas, and permanently for family/settlement visas and asylum seekers.

We assessed the impact of changing our assumed distribution for length of stay based on longer and shorter distributions for the average time that migrants with different visa types remain in the UK. For the *longer* scenario, we assumed that all those migrating to the UK with work and study visas would remain in the UK for the duration category above that of their application: e.g., migrants on 3 month visas remain for 6 months, those on 6 month visas remain for 1 year and so on. In addition, 25% of all migrants on work/study visas were assumed to remain permanently in the UK. For the *shorter* scenario, we assumed that all those migrating to the UK with work and study visas would remain in the UK for the duration category below that of their application (or at least 3 months): e.g., those on 6 month visas remain for 3 months, those on 1 year visas for 6 months, and so on. In addition, 25% of all migrants with visa types that were of more than one year were assumed to leave the UK after 1 year; and 25% of all migrants on family/settlement visas were also assumed to leave the UK after 1 year (rather than remain permanently in the UK).

Results

Changing the distribution of lengths of stay to be shorter or longer primarily affects the size of the mid-term migrant group. Most migrants are assumed to stay for 1-2 years, so the size of the recent migrant group remains similar regardless of average length of stay. The size of the long-term migrant group is relatively large, so does not change markedly if new migrants have longer or shorter lengths of stay. IRRs for 2015 vs. 2011 for recent and long-term migrants therefore show little change. For mid-term migrants the results change in magnitude but not direction. India and Pakistan are the most affected, with the IRR increasing if lengths of stay are shorter, and reducing if lengths of stay are longer. For India, the IRR for mid-term migrants with longer lengths of stay was estimated to be 0.68 (0.59, 0.78); with shorter lengths of stay the estimate was 0.95 (0.82, 1.09). For Pakistan, the IRR for mid-term migrants with longer lengths of stay was estimated to be 0.71 (0.58, 0.87); with shorter lengths of stay the estimate was 0.86 (0.71, 1.06).

The contributions of migration, changes in TB rates and PES to the overall reduction in TB in the UK from 2011 to 2015 also showed little change under different lengths of stay assumptions, although the relative importance of changes in migration and TB rates were altered if lengths of stay are longer. .For instance, the estimated contribution of changes in TB rates in recent/mid-term migrants increased from 7.4% to 15.3% while the impact of migration was somewhat smaller.

References

1. Aldridge RW, Zenner D, White PJ, *et al.* Tuberculosis in migrants moving from high-incidence to low-incidence countries: a population-based cohort study of 519,955 migrants screened before entry to England, Wales, and Northern Ireland. *Lancet* 2016; **388**: 2510–8.