Supplementary material

S1. Calculations on probability of remote LTBI

Estimation of remote LTBI (P_{remote}) (%) included probability of infection in country of origin (P_{origin}), in Sweden (P_{Sweden}) and during travels to high-endemic countries >1 year (P_{travel}), stays in high risk environments such as hospital, refugee camp and/or prison >3 months (P_{high}) and previous known exposure to contagious pulmonary TB (P_{expo}) and was calculated as:

$$P_{remote} = 1 - ((1 - P_{origin}) * (1 - P_{Sweden}) * (1 - P_{travel}) * (1 - P_{high}) * (1 - P_{expo}))$$

P_{origin} was calculated for three separate time-periods (-1990, 1990-2000 and 2000-) as:

$$(1 - ((1 - (P_{-1990} / 100)) * (1 - (P_{1990\text{-}2000} / 100)) * (1 - (P_{2000\text{-}} / 100)))) * 100$$

Probability for each time-period was calculated as: $(1 - ((1 - (ARI / 100))^{years})) * 100$

Years; number of years in country of origin from birth to immigration to Sweden for each time-period.

ARI; annual risk of infection = TB incidence / Dye´s constant [1-3]. TB incidences in country of origin at year 1990, 2000 and at time of immigration were used for each time-period, respectively (http://www.who.int/tb/country/data/profiles/en/ and http://publichealthintelligence.org/content/global-burden-tuberculosis-1990-2013). Dye´s constant was adjusted according to improved clinical management and availability to effective TB drugs over time e.g. before 1990 we estimated that one active TB case transmitted the disease to ten individuals over one year, giving us a Dye´s constant of 100 (TB cases / 100000 x 10 yearly x 1 year = TB cases / 10000 = ARI (%) / 100) and 200 for 1990-2000 and 400 after 2000 [4].

P_{Sweden} was calculated as: $(1 - ((1 - (ARI / 100))^{years})) * 100$

Years; number of years after immigration for foreign born and years from birth for Swedish born until inclusion in study (2010).

ARI; based on TB incidence in Sweden in 2010 and Dye´s constant at 400.

P_{travel} was calculated as: $(1 – ((1 – (ARI / 100))^{years})) * 100$

Years; number of years on travel.

ARI; based on TB incidence and Dye´s constant corresponding to respective country and time period.

P_{high} was calculated as: $(1 – ((1 – (ARI / 100))^{years})) * 100$

Years; number of years in high risk environment.

ARI; based on TB incidence estimated to 300/100 000 and Dye´s constant at 400.

P_{expo} was estimated according to previously published data [5] e.g. close contact to smear microscopy positive (SM+) 35%, SM- 10% and SM unknown 20%, while not close (casual) contact to SM+ 10%, SM- 2% and SM unknown 5%.

Example: Man born 1975 in Eritrea. Immigration to Sweden 2010. Soldier, imprisoned two years. Brother (not close contact) treated for pulmonary TB in the 90´s.

P_{origin}	Probability of TB infection due to origin from and living in Eritrea 1975 – 2010: For respective time periods 1975 – 1990, 1990 – 2000 and 2000 – 2010 the yearly TB incidences per 100 000 were 242 (year 1990), 157 (year 2000) and 100 (year 2010, time of immigration). The respective Dye´s constants were 100, 200 and 400, giving respective ARIs of 2.42, 0.79 and 0.25. The probability for each time period and the summarized probability was calculated as described above, giving a summarized probability of remote LTBI based on origin of 37.6%.
P_{Sweden}	No additional probability until time of inclusion in study (2010).
P_{travel}	No additional probability from travels to high endemic countries more than one year.

P_{high} Added probability from staying in prison 2 years: Estimated TB incidence (per 100000) was 300 and Dye´s constant 400, giving an ARI of 0.75 and an added probability of 1.5%.

P_{expo} Added probability from previous exposure to brother with pulmonary TB: Estimated probability of infection after casual contact to TB with unknown SM status was estimated to 5% as described by Grzybowski et al. [5].

Summarized probability of remote LTBI:

$P_{remote} = (1 – ((1 – (37.6 / 100)) * (1 – (1.5 / 100)) * (1 – (5 / 100)))) * 100 = 41.6\%$

S2. Calculations on probability of recent LTBI

Estimations of recent LTBI (P_{recent}) (%) included contagiousness of index case, proximity and time of exposure to index and environmental factors and was calculated as:

$P_{recent} = 1 – ((1 – P_{recent}\,day) * (1 – P_{recent}\,night))$

P_{recent} day and P_{recent} night was calculated as: $1 – (1 – P_i)^n$ according to daytime and night time parameters, respectively.

n; estimated number of breaths by individual contact during estimated time of TB exposure and was calculated as: RR * 60 * t * d.

RR; repiratory rate, defined as 14/min during daytime and 16/min during nighttime [6]

t; time (hours) exposed in daytime and defined as eight hours in nighttime if sleeping together.

d; number of days exposed was calculated from total period and frequency of interaction between contact and index case, starting from one month before any symptoms by index until end of interaction or hospitalization of index.

P_i ; probability of the individual contact to inhale *Mtb* existing in the air volume with each breath was calculated as: $(Dn^+ * S * V_i * K) / (E * V_d)$

Dn^+ ; estimated production of *Mtb* containing droplets (Dn) per minute was calculated as: C_{Mtb} * Dn

C_{Mtb} ; *Mtb* concentration in sputum was calculated as average number of colony forming units (CFU) per µL of 2 – 3 sputum samples per index patient. To verify accuracy, microscopy slides were revised by the same microbiologist at the TB laboratory, Karolinska to define the absolute number of *Mtb* per µL [7, 8].

Dn; production of droplet nuclei (Dn) was estimated to 200/min daytime with an 150/min added for an estimated 20% time of speech. During nighttime production was estimated to 100/min. With documented cough another 150/min was added daytime and nighttime [6, 9-13].

E; elimination of Dn^+ was calculated as: $1 – E = (1 – E_v) * (1 – E_p)$.

E_v ; estimated elimination rate of Dn^+ through ventilation/min was calculated from defined numbers of air exchanges/h e.g 0.4 times/h for general buildings e.g. households, schools, offices etc and 5 times/h for hospitals. $E_v = 0.4 / 60 = 0.007$ for general buildings. $E_v = 5 / 60 = 0.083$ for hospitals.

E_p ; elimination of Dn^+ through precipitation/min was estimated from a half-life of 30 min for mixed droplet nuclei of 1-10 µL [14] and was calculated as: $(1 – E_p)^{30} = 0.5 \rightarrow E_p = 0.023$.

E = 0.029 for general buildings.

E = 0.104 for hospitals

S; saturation of Dn^+ in air volume was calculated as: $S = S_{ss} * S_f$

S_{ss}; estimated number of Dn^+ in the air at steady state i.e. how many times more Dn^+ are present in the air as compared to what is produced by the index case every minute. At steady state production of Dn^+ into the air equals elimination of Dn^+ from the air i.e. E = 1 and as such $S_{ss} = 1 / 0.029 = 34.07$ for general buildings.

$S_{ss} = 1 / 0.104 = 9.59$ for hospitals.

S_f ; saturation factor i.e. estimated degree of full saturation (steady state) was calculated as time to 90% saturation, which was reached within 1h for general buildings and 2h for hospitals. As such, S_f was approximated to 1.0 in case of contact more than 8h and 0.75 in case of contact less than 8h.

V_i ; estimated alveolar volume (l) in one normal inhalation by the contact [15] was calculated as: (VT – ADS) / 1000

VT; estimated volume (ml) in one normal inhalation was calculated as: 6 * IBW

IBW; Ideal Body Weight (kg) by Hamwi GJ Formula 1964 (http://www.calculator.net/ideal-weight-calculator.html)

ADS; Anatomic dead space (ml) [16] was calculated as: $7.585 * \text{height (cm)}^{2.363} * 10^{-4}$

V_d ; volume of distribution was the estimated volume of air (liter) in premises e.g. household, classroom, office, ward room etc where main exposure took place and was calculated as:

A * h * 1000

A; area of premises (m^2)

h; height of premises (2.4m)

K; calculated constant = 1/5000000 was applied for all individuals to adjust for unknown parameters such as numbers of *Mtb* containing droplets, *Mtb* per droplet, *Mtb* reaching the alveoli, *Mtb* needed to establish infection etc. and validated according to previous published data on probability of transmission in contact tracing [5, 17, 18].

Example: Man born 1975 in Eritrea (same as above), 179 cm and 86 kg. He has been staying with his wife in a friend's apartment since his arrival in Sweden. His wife is pregnant and has been sick with fever, night-sweats and slight coughing for one and a half month and was recently diagnosed with SM+ pulmonary TB. They have not been sleeping together during the period when index case has been symptomatic.

Probability of recent LTBI, daytime:

$Dn+ = C_{Mtb} * Dn$ = 1543 *Mtb*/µl * 200 Dn/min = 308 600 Dn+/min

$S = S_{ss} * S_f$ = 34.07 (household) * 1.0 (>8h contact) = 34.07

V_i = (VT – ADS) ml = (6 * 76.3 kg) – ($7.585 * 179cm^{2.363} \times 10^{-4}$) = 459 – 159 = 300 ml = 0.3 L

V_d = Area * 2.4 *1000 = 108 m^2 * 2.4 * 1000 = 259 200 L

K = 1/5000000

n = RR * 60 * t * d = 14 * 60 * 4 (h/day) * 56 days = 188 160

$P_i = (Dn+ * S * V_i * k) / V_d$ = (308 600 * 34.07 * 0.3) / (259 200 * 500 000) = 2.43E-6

$P_{recent\ day} = 1 – (1 – P_i)^n = 1 – (1 – 2.43E\text{-}6)^{188160}$ = 0.367 = 36.7%

Summarized probability of recent LTBI, daytime and night time:

$P_{recent} = 1 – ((1 – P_{recent}\,day) * (1 – P_{recent}\,night)) = 1 – ((0.633) * (1)) = 0.367 = 36.7\%$

References

1. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC: **Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project**. *JAMA* 1999, **282**(7):677-686.
2. Dye C: **Breaking a law: tuberculosis disobeys Styblo's rule**. *Bull World Health Organ* 2008, **86**(1):4.
3. Dye C, Bassili A, Bierrenbach AL, Broekmans JF, Chadha VK, Glaziou P, Gopi PG, Hosseini M, Kim SJ, Manissero D *et al*: **Measuring tuberculosis burden, trends, and the impact of control programmes**. *Lancet Infect Dis* 2008, **8**(4):233-243.
4. Kompala T, Shenoi SV, Friedland G: **Transmission of tuberculosis in resource-limited settings**. *Curr HIV/AIDS Rep* 2013, **10**(3):264-272.
5. Grzybowski S, Barnett GD, Styblo K: **Contacts of cases of active pulmonary tuberculosis**. *Bull Int Union Tuberc* 1975, **50**(1):90-106.
6. Douglas NJ, White DP, Pickett CK, Weil JV, Zwillich CW: **Respiration during sleep in normal man**. *Thorax* 1982, **37**(11):840-844.
7. Lohmann EM, Koster BF, le Cessie S, Kamst-van Agterveld MP, van Soolingen D, Arend SM: **Grading of a positive sputum smear and the risk of Mycobacterium tuberculosis transmission**. *Int J Tuberc Lung Dis* 2012, **16**(11):1477-1484.
8. Andersen AA: **New sampler for the collection, sizing, and enumeration of viable airborne particles**. *J Bacteriol* 1958, **76**(5):471-484.
9. Johnson GR, Morawska L: **The mechanism of breath aerosol formation**. *J Aerosol Med Pulm Drug Deliv* 2009, **22**(3):229-237.
10. Nicas M, Nazaroff WW, Hubbard A: **Toward understanding the risk of secondary airborne infection: emission of respirable pathogens**. *J Occup Environ Hyg* 2005, **2**(3):143-154.
11. Fennelly KP, Martyny JW, Fulton KE, Orme IM, Cave DM, Heifets LB: **Cough-generated aerosols of Mycobacterium tuberculosis: a new method to study infectiousness**. *Am J Respir Crit Care Med* 2004, **169**(5):604-609.
12. Fennelly KP, Jones-Lopez EC, Ayakaka I, Kim S, Menyha H, Kirenga B, Muchwa C, Joloba M, Dryden-Peterson S, Reilly N *et al*: **Variability of infectious aerosols produced during coughing by patients with pulmonary tuberculosis**. *Am J Respir Crit Care Med* 2012, **186**(5):450-457.
13. Jones-Lopez EC, Namugga O, Mumbowa F, Ssebidandi M, Mbabazi O, Moine S, Mboowa G, Fox MP, Reilly N, Ayakaka I *et al*: **Cough aerosols of Mycobacterium tuberculosis predict new infection: a household contact study**. *Am J Respir Crit Care Med* 2013, **187**(9):1007-1015.
14. Rieder HL: **Epidemiologic Basis of Tuberculosis Control**: International Union Against Tuberculosis and Lung Disease; 1999.
15. ERS: **Multi-Ethnic reference values for spirometry for the 3–95 year range**. *European Respiratory Journal* 2012, **40**(6):1324-1343.
16. Hart MC OM, Cook CD: **Relation between anatomic respiratory dead space and body size and lung volume**. *Journal of Applied Physiology* 1963, **18**(3):519-522.
17. Fennelly KP, Jones-Lopez EC: **Quantity and Quality of Inhaled Dose Predicts Immunopathology in Tuberculosis**. *Front Immunol* 2015, **6**:313.
18. Gupta N, Kumar R, Agrawal B: **New Players in Immunity to Tuberculosis: The Host Microbiome, Lung Epithelium, and Innate Immune Cells**. *Front Immunol* 2018, **9**:709.