Appendix 1 (as supplied by the authors): ONLINE SUPPLEMENT

Title Housing and tuberculosis in an Inuit village in Nunavik, Quebec: a case-control study

Authors Faiz Ahmad Khan, Greg J. Fox, Robyn S. Lee, Mylène Riva, Andrea Benedetti, Jean-Francois Proulx, Shelley Jung, Karen Hornby, Marcel A. Behr, Dick Menzies

Appendix to: Ahmad Khan F, Fox GJ, Lee RS, et al. Housing and tuberculosis in an Inuit village in northern Quebec: a case–control study. *CMAJ Open* 2016. DOI:10.9778/cmajo.20160049.

Table S1. Comparison of participants and non-participants.

(A) By TST category: new infection, prior infection, no infection [1].

TST category*		Age group (years)			Gender
	Total	0-14	15-29	≥ 30	Male
New infection¶					
Participants	88	30 (34.1%)	44 (50%)	14 (15.9%)	40 (45.5%)
Non-participants	100	32 (32.0%)	46 (46.0%)	22 (22.0%)	45 (60.8%)
Prior infection¶					
Participants	45	2 (4.4%)	16 (35.6%)	27 (60%)	22 (48.9%)
Non-participants	202	11 (5.4%)	46 (22.8%)	145 (71.8%)	98 (48.8%)
No infection					
Participants	67	45 (67.2%)	17 (25.4%)	5 (7.5%)	31 (46.3%)
Non-participants	193	117 (60.9%)	55 (28.6%)	20 (10.4%)	104 (54.2%)

(B) Amongst villagers with confirmed or probable TB disease

Disease group*		Age group (years)			Gender	TST category	
	Total	0-14	15-29	≥ 30	Male	New infection	Prior infection
TB disease							
Participants	44	12 (27%)	22 (50%)	10 (23%)	22 (50%)	35 (80%)	9 (20%)
Non-participants	25	6 (24%)	11 (44%)	8 (32%)	12 (48%)	18 (72%)	7 (28%)

TST, tuberculin skin test; TB, tuberculosis. Gender of 2 non-participants was missing, and age of 1 non-participant in the "No infection" category. *No significant differences were found between the characteristics of participants and non-participants in any sub-group ($p>0.05$ for all comparisons, using Fisher's exact test). ¶Includes individuals with or without disease. All individuals with disease had a positive tuberculin skin test.

Appendix to: Ahmad Khan F, Fox GJ, Lee RS, et al. Housing and tuberculosis in an Inuit village in northern Quebec: a case–control study. *CMAJ Open* 2016. DOI:10.9778/cmajo.20160049.

Table S2. Comparison of dwellings used for social activities to other dwellings in the village.

Variable	Dwellings used for social activities	Dwellings that were not used for social activities
Number of dwellings of this type in the village	18	178
Dwelling where at least one study participant lived from Nov 2011-Nov 2012, N(%)	10 (55.6%)	69 (38.8%)
Dwelling where ventilation was measured and at least one study participant lived from Nov 2011-Nov 2012, N(%)	4 (22.2%)	47 (26.4%)
Number of dwellings of this type visited weekly by at least one participant	18 (100%)	128 (71.9%)
Years since dwelling built (as of 2012), mean (±SD) †	16.4 (± 9.7)	17.4 (± 9.4)
Pulsed air heating, N(%)†	18 (100%)	137 (80.1%)
Number of rooms per dwelling, mean (±SD) †	6 (± 1)	5 (± 1)
ACH in bedrooms with heating off, mean (±SD) †	0.28 (± 0.11)	0.29 (± 0.19)
ACH in bedrooms with heating on, mean (±SD) †	1.16 (± 0.11)	1.29 (± 0.27)
ACH in living areas with heating off, mean (±SD) †	0.73 (± 0.18)	0.64 (± 0.22)
ACH in living areas with heating on, mean (±SD) †	1.60 (± 0.12)	1.68 (± 0.30)
Volume of bedrooms m^3, mean (±SD) †	29.9 (± 4.8)	28.2 (± 4.0)
Volume of living areas m^3, mean (±SD) †	70.9 (± 9.8)	66.1 (± 7.9)
Information on occupancy and TB Status of occupants-determined from cohort of 695 villagers ††		
Number of persons per dwelling, mean (±SD)	5.3(± 3.0)	5.0 (± 2.8)
Occupants per room, mean (±SD)	0.9(± 0.5)	1.0 (± 0.5)
Dwelling where at least 1 occupant had TB disease, N(%)	6 (33.3%)	41 (23.0%)
at least 1 occupant with smear-positive TB, N(%)	1 (5.6%)	10 (5.9%)

ACH, air changes per hour; SD, standard deviation

Occupant defined as person residing in the dwelling; Visitor defined as someone who visits the dwelling at least once weekly but does not reside there.

† Denominators exclude 7 dwellings that were not used for social activities for which information on housing characteristics could not be determined.

†† Data on number of persons per dwelling missing for 2 dwellings used for social activities. Data on number of rooms in the dwelling missing for 19 dwellings that were not used for social activities.

Appendix to: Ahmad Khan F, Fox GJ, Lee RS, et al. Housing and tuberculosis in an Inuit village in northern Quebec: a case–control study. *CMAJ Open* 2016. DOI:10.9778/cmajo.20160049.

Methods Supplement

Section 1. Questionnaire data

Administered questionnaires assessed nutrition, housing, socio-economic and lifestyle factors.[1] Parents provided responses for children under the age of 12.

- *Smoking status*: was assigned based on responses on a self-administered questionnaire completed only by adult participants, and serum cotinine [2,3] which was measured in all (children and adults).

- *Annual personal income*: participants were asked to classify their annual personal income into one of three categories: less than $20,000; $20,000-40,000; or greater than $40,000. For participants under the age of 18, the annual personal income of the parent or guardian was used.

- *Identification of dwellings used for social activities:* The addresses of these houses were provided to the research team by community members and clinic staff. All other dwellings were classified as “not used for social activities”.

- *Visiting of dwelling:* In one section, participants were asked to list three dwellings they visited at least once weekly; this information was used to determine if they had visited one of the dwellings used for social activities, at least one weekly.

- *Identification of dwellings where visitors could have been exposed to someone with active TB*: We used two methods to identify dwellings where study participants could have been exposed to someone with active TB. From the questionnaires completed by

Appendix to: Ahmad Khan F, Fox GJ, Lee RS, et al. Housing and tuberculosis in an Inuit village in northern Quebec: a case–control study. *CMAJ Open* 2016. DOI:10.9778/cmajo.20160049.

participants with confirmed or probable active TB, we determined which dwellings they visited at least weekly. From data collected as part of medical and public health investigations, we identified the addresses of where all persons diagnosed with confirmed or probable TB had lived between November 2011 and November 2012.

Section 2. Housing & ventilation assessment

Data on the year of construction, heating method, dwelling type, and layout were provided by the two municipal/regional institutions which owned these dwellings. Room dimensions were measured by study personnel and verified using floor plans provided by the housing bureau.

"Ventilation" is defined as the exchange of indoor air for outdoor air per unit of time. In the community where the study was undertaken, nearly all dwellings are heated using "forced air" (also called "pulsed air") heating systems — when the heating system is on and the temperature is lower than the thermostat setting, outdoor air is heated and delivered indoors; as a result, ventilation is higher compared to when the heating is off. (The exceptions to use of forced air systems are apartments, which are heated using radiators.) Because of this, we measured ventilation with heating off and with heating on. Measurements were taken in at least one bedroom (where the participant slept) and in the common living area (this was the living room and the kitchen in all houses visited).

To measure ventilation, all windows and doors were closed – including doors leading to all indoor rooms and doors leading to the outside. Pure CO_2 was used as a tracer gas [5,6] and released at a high rate to achieve a peak of 1500ppm above baseline (corresponding, on

average, to a release time of 6 seconds/m^2 in bedrooms, and 9 seconds/m^2 in living areas). Two motorized fans were used throughout the release and measurement to ensure mixing of CO_2 throughout the entire room. CO_2 was measured continuously with a portable device with electro-chemical detectors (Q-Trak Plus, TSI) which logged the concentration every minute. Recordings were taken for a total of 20 minutes. During the recording, study staff noted if any doors or windows were opened. All occupants were asked to vacate the rooms where ventilation was being measured; and when this was not possible, the number of people present, and the number of people entering or leaving were noted, and used to adjust estimates of ventilation (see below).

Because CO_2 concentration falls at an exponential rate, the following equation is used to calculate ventilation based on the changes in CO_2 concentration over time ("Standard Test Method for Determining Air Change in a Single Zone by Means of a Tracer Gas Dilution," ASTM E741-11):

$$\ln[CO_2]_i(t) = [CO_2]_0 - \beta \times t_i$$

where $[CO_2]_i(t)$ is the concentration at time i, $[CO_2]_0$ is the CO_2 concentration at time zero, and β is air changes per hour (ACH), which is the number of times in one hour that the total volume of air in a room is replaced by air from outside the room.

The above relationship is applicable if CO_2 concentration has reached equilibrium throughout the room. We took several precautions to ensure that CO_2 measurements were at equilibrium when calculating ventilation. First, as mentioned earlier, two fans were running during CO_2

release and measurement to ensure mixing of air. Second, CO_2 recordings during the first three minutes of measurement were discarded. Third, we also discarded CO_2 measurements that were outliers and unduly influencing the ACH; this was done using simple linear regression (with $\ln[CO_2]$ as the dependent variable, and time as the independent variable) to identify observations with Cook's distance greater than 4/n (where n is the number of CO_2 measurements) [7].

Dwellings in this community are grouped into five main types, each of which has between one and 3 sub-types. Dwellings of the same type share similar construction characteristics and layouts; sub-types differ in size and room number. We prioritized taking ventilation measurements in dwellings where persons with active TB had lived and dwellings of the same type as the gathering houses. When ten measurements had been performed in the most common sub-types, we performed an interim analysis of ventilation data to decide whether we had enough measurements for each sub-type. This was based on an analysis of the variance between and within buildings- when the variance of ventilation measurements between different buildings of the same sub-type was less than the variance of repeated measurements within the same building, we concluded we did not need further measures in buildings of that type. For all subsequent participants who lived in buildings of that type, they were assigned model-specific average ventilation measures (average values of all ventilation measurements from all buildings of that type). If ventilation was measured in the dwelling of a participant, these directly measured values were assigned to that participant.

Appendix to: Ahmad Khan F, Fox GJ, Lee RS, et al. Housing and tuberculosis in an Inuit village in northern Quebec: a case–control study. *CMAJ Open* 2016. DOI:10.9778/cmajo.20160049.

Air change rates were estimated by analyzing the CO_2 recordings using multi-level multivariable regression, with $\ln[CO_2]$ as the dependent variable and time as the independent variable whose estimated slope is the ACH. These models included interaction terms to account for the effects on ventilation of the room (living area vs bedrooms) and heating condition (on vs off), plus the number of persons that were in the room during the recording, as well as a random-effect for the ACH between dwellings. This approach had several advantages. First, it ensured that dwelling-specific ACH estimates accounted for the variance of the ACH measurement in each dwelling. Second, all recordings from dwellings of the same type were used to estimate a weighted average ACH for each type. Third, multi-level regression allowed us to adjust for the number of persons who were present during the recording- this was important because people are sources of CO_2, and the presence of a continuous source of CO_2 release during the recording would slow the rate of decline in CO_2 concentration, and if not controlled for, would lead to underestimation of ventilation.

Four ACH estimates were calculated for each dwelling: with heating off in living areas, with heating on in living areas, with heating off in bedrooms, and with heating on in bedrooms. These ACHs were calculated as the sum of: the weighted average ACH of dwellings of the same type and the applicable interaction terms for the effects of the room and heating status (also for dwellings of the same type). For dwellings where ACH was directly measured, we also added the dwelling's random effect estimate.

Section 3. Details on statistical methods

a. Analysis of ventilation as a risk factor for TB infection and disease: We analyzed the ventilation level in the common living area of participants' dwellings — since all household occupants shared these areas, and many occupants slept in these areas. The maximal ventilation (with heating on) and the minimal ventilation (with heating off) were analyzed separately.

b. Multivariable modeling: For adjusted analyses we used the method of generalized estimating equations for logistic regression, with clustering at the dwelling level. For each objective, we first calculated univariable odds ratios (OR) for: (1) characteristics of the participant's dwelling (method of heating, occupancy, ACH, and room volume of the living area); (2) variables related to potential exposures to persons with TB (at least weekly visiting a dwelling that was used for social activities [yes vs. no], and having lived with someone who had smear-positive TB [yes vs. no]); and, (3) characteristics we considered well-established determinants of TB (age, gender, cigarette smoking, and socioeconomic status). In our preliminary multivariable model, we included all variables associated in univariable analysis with a p-value <0.5. Using a step-wise approach we excluded variables if their removal improved the multivariable model's fit as indicated by a lower quasi-likelihood information criterion (QICu) [8]. This approach was taken in order to minimize over-specification given our limited sample size. For the first objective, we hypothesized that occupancy would only be associated with risk of infection in dwellings inhabited by smear-positive persons, and assessed this in univariable analysis using stratification, and in multivariable analysis by including an interaction term. Interactions were considered significant if their associated p-values were <0.05. For disease, the interaction term was kept in our final model despite an increase QICu because it was statistically significant, and

for consistency with our model for infection. We considered associations to be significant if the 95% confidence interval (95% CI) of the odds ratios excluded 1.0.

References Methods Supplement

1. Fox GJ, Lee RS, Lucas M, Ahmad Khan F, Proulx JF, Hornby K, et al. Inadequate Diet is Associated with Acquiring Mycobacterium tuberculosis Infection in an Inuit Community: A Case-Control Study. *Ann Am Thorac Soc*. 2015.

2. Benowitz NL, Bernert JT, Caraballo RS, Holiday DB, Wang J. Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. *Am J Epidemiol*. 2009; **169**(2): 236-48.

3. Available from: http://www.salimetrics.com/assets/documents/Spit_Tips_-_Cotinine_Guidelines.pdf Accessed June 6, 2014

4. Statistics Canada. December 2012. Available from: http://www.statcan.gc.ca/concepts/definitions/dwelling-logement01-eng.htm Accessed July 9, 2014

5. Menzies R, Schwartzman K, Loo V, Pasztor J. Measuring ventilation of patient care areas in hospitals. Description of a new protocol. *Am J Respir Crit Care Med*. 1995; **152**(6 Pt 1): 1992-9.

6. Persily AK. Evaluating building IAQ and ventilation with indoor carbon dioxide. Transactions-American Society Of Heating Refrigerating And Air Conditioning Engineers. 1997; **103**: 193-204.

7. Rawlings JO, Pantula SG, Dickey DA. Applied regression analysis a research tool. Springer texts in statistics. 2nd ed. New York: Springer; 1998.

8. Hardin JW, Hilbe JM. Generalized estimating equations. Boca Raton, Fla.: Chapman & Hall/CRC; 2003.

