Predicting residential radon concentrations in Finland: Model development, validation, and application to childhood leukemia [1]

by Atte Nikkilä, MD, PhD,[2] Hannu Arvela, PhD, Juha Mehtonen, MS, Jani Raitanen, MS, Merja Heinäniemi, PhD, Olli Lohi, MD, PhD, Anssi Auvinen, MD, PhD

1. *Supplementary tables*
2. *Correspondence to: Atte Nikkilä, Faculty of Medicine and Health Technology, Tampere University Arvo Ylpön katu 34, 33520 Tampere, Finland. [E-mail: atte.nikkila@tuni.fi]*

Supplementary table 1 – The percentage of missing data on each predictor before multiple imputation

	Apartments	**Hosues**
Building material	0.1%	0.7%
Basement	0%	0%
Number of floors	0.4%	1.1%
Soil's uranium concentration	4.3%	4.1%
Elevation	10.6%	12.6%
Exhaust fan	0.0%	0%
Soil permeability	35%	3.3%
Formation by ice-age	0%	0%
Year of completion	0.1%	1.4%
Floor area	0.0%	0.5%
Total area	31%	36%
Total volume	20%	34%
Median radon (county)	0.1%	-
Median radon (postal code)	-	3.5%

For some predictors (exhaust fan, basement, formation by ice-age) missing data and other-group were combined as the distinction between them was not possible and we did not impute for those values.

Supplementary table 2 – The characteristics of cases and controls before any exclusions

	Cases *(n=1093)*	**Controls** *(n=3279)*	**OR (95% CI)**
Gender			
female	48.0% (525)	48.0% (1575)	
male	52.0% (568)	52.0% (1704)	
Large for gestational age			
no	86.7% (788)	90.1% (2493)	
yes	13.3% (121)	9.9% (275)	1.44 (1.14, 1.81)
missing	184	511	
Mother's smoking during pregnancy			
no	83.1% (742)	84.5% (2296)	
yes	16.9% (151)	15.5% (420)	1.15 (0.94, 1.42)
missing	200	563	
Down syndrome			
no	96.3% (1053)	99.9% (3277)	
yes	3.7% (40)	0.1% (2)	60 (14.5, 248)
Parents' education			
Mother			
Upper secondary	48.5% (530)	50.6% (1659)	ref.
Bachelor's degree	22.3% (244)	23.1% (756)	1.02 (0.84, 1.23)
Master's or doctor's degree	10.2% (112)	9.8% (321)	1.11 (0.87, 1.42)
missing	18.9% (207)	16.6% (543)	
Father			
Upper secondary	52.0% (568)	51.4% (1685)	ref.
Bachelor's degree	15.2% (166)	16.2% (532)	1.09 (0.74, 1.14)
Master's or doctor's degree	10.0% (110)	10.2% (334)	0.98 (0.79, 1.31)
missing	22.8% (249)	22.2% (728)	
Parents' socioeconomic status			
Mother			
Self-employed	7.7% (84)	8.3% (273)	ref.
Upper level employees	16.1% (176)	15.7% (514)	1.11 (0.83, 1.50)
Lower level employees	34.8% (380)	34.5% (1130)	1.09 (0.83, 1.44)
Manual workers	21.4% (231)	20.6% (674)	1.11 (0.83, 1.47)
others	18.2% (199)	20.3% (664)	0.97 (0.72, 1.31)
missing	2.1% (23)	0.7% (24)	
Father			
Self-employed	13.9% (152)	12.0% (395)	ref.
Upper level employees	17.6% (192)	18.2% (596)	0.85 (0.66, 1.08)
Lower level employees	18.3% (197)	17.9% (587)	0.87 (0.68, 1.12)
Manual workers	34.0% (372)	35.0% (1148)	0.86 (0.69, 1.07)
others	12.4% (135)	14.3% (469)	0.75 (0.58, 0.98)
missing	4.1% (45)	2.6% (84)	
Age at leukemia diagnosis, years			
0 – 2	14.3% (156)		
2 – 7	55.5% (605)		
7 – 15	33.4% (332)		
Leukemia type			
pre-B-ALL	75.6% (826)		
pre-T-ALL	5.9% (64)		
unclassified ALL	1.8% (20)		
AML	13.6% (149)		
other	3.1% (34)		

The reported ORs and their respective confidence intervals are from an univariate conditional logistic regression model. The non-binary variables were treated as factors and the reference categories are marked with "ref".
An alternate version of this table has been previously published (Nikkilä et al. Haematologica. 2018)

Supplementary Table 3 – The odds ratios and their confidence intervals from conditional logistic regression analyses on the effect of predicted indoor radon concentration on childhood leukemia and its several subgroups

		Cumulative indoor radon exposure (Bq/m³-years)				Average indoor radon concentration (Bq/m³)			
		Log-linear		*Random forests*		*Log-linear*		*Random forests*	
	N	OR	95 % CI	OR	95 % CI	OR	95 % CI	OR	95 % CI
All subjects	*922*	1.06	0.59, 1.92	1.02	0.99, 1.05	0.93	0.42, 2.05	1.01	0.98, 0.98
By leukemia subtype									
ALL	*806*	1.32	0.67, 2.60	0.99	0.38, 2.58	1.03	1.00, 105	1.02	0.98, 1.05
Others	*183*	0.42	0.09, 1.89	0.51	0.09, 2.87	0.98	0.89, 1.07	0.99	0.89, 1.10
pre-B ALL	*735*	1.59	0.74, 3.38	1.11	0.39, 3.18	1.03	1.00, 1.06	1.02	0.98, 1.05
By age-group (years)									
2 – 5.99		3.53	0.80, 15.5	2.86	0.52, 15.9	1.03	1.00, 1.07	1.02	0.99, 1.06
6 – 15		0.79	0.40, 1.57	0.67	0.25, 1.77	0.98	0.93, 1.04	0.98	0.93, 1.04

Only subjects with non-zero exposure were included. All estimates are from adjusted models.