Childhood cancer and nuclear power plants in Switzerland: A census-based cohort study

Webappendix 1 - Methods

Content

Aims and objectives (copied verbatim from the original study protocol)

This study will address the question whether residence in the proximity of a nuclear power plant is associated with an increased risk of childhood cancer, and whether this can be explained by confounding, particularly by other area-based risk factors for childhood cancer which might cluster around nuclear power plants. Our detailed objectives are as follows:

1) To determine if residence near a nuclear power plant is associated with an increased risk of childhood cancer in Switzerland; with the following outcomes of interest:
   - All cancers in children aged <16 years at diagnosis
   - All leukaemias in children aged <16 years at diagnosis
   - All cancers in children aged <5 years at diagnosis
   - All leukaemias in children aged <5 years at diagnosis

2) To determine if this could be explained by other risk factors for childhood cancer, including, for example, exposure to electromagnetic fields or industrial pollutants and chemicals, or differences in socio-economic status (SES). Based on the available literature on environmental risk factors for childhood cancer12 we will consider the following nine potential confounders:
   - Major roads
   - Electric power lines
   - Radio and TV transmitters
   - Degree of urbanization
   - Socioeconomic position
   - Pesticides from agriculture, golf courses, pollutants from industry
   - Natural ionising radiation
   - Average number of children per family
   - Distance to the next pediatric cancer center

Main exposure of interest will be address at the time of birth

Planned additional analyses (copied verbatim from the original study protocol)

a. **Confounding**: We will consider a selection of potential confounders in our analysis and adjust the main analysis for these confounders (see section 2.2.3.5 and Table 4). Potential confounding factors will be included in the model one at a time. Confounding will be an issue if there is a substantial change in the effect estimates due to one confounder, or if the effect estimates are consistently changing in the same direction for all potential confounders.
b. **Modelling of exposure assessment**: As described in section 2.2.3.4. (Definition of main exposure) modelling of exposure taking into account wind directions and data from NADAM will be explored. If this analysis yields a stronger association than the main analysis, this would indicate that a causal relation with radiation or another exposure emanating from the NPP might exist.
c. **Restricting the study area**: At a distance >50km from a NPP emissions of radiation are virtually zero. If childhood cancer incidence in this area is different than <50km from NPPs due to other factors than the emission from the NPPs, this could bias our analysis. We will therefore restrict our cohort to the population living <50 km from a power plant. The reference category will then contain the population living between 15 and 50 km from a NPP.
d. **Timing of exposure**: We will perform a sensitivity analysis using the place of residence at the time of diagnosis (instead of birth) in order to compare our results with the German KiKK study. This will allow us to evaluate the importance of the time of the exposure.
e. **Continuous exposure measure 1/r**: In order to compare our results with the KiKK study and to increase the power of the analysis we will use 1/r (r =distance between place of residence at birth and nearest NPP). Additionally we will take into account the fact that some of the children have moved their place of residence, by calculating a time-weighted average of 1/r (inverse distance to the next NPP). We will determine the inverse distance to the nearest NPP for each place of residence up to 6 months prior to the diagnosis, and calculate the time weighted average.

Cohorts

We defined two cohorts:

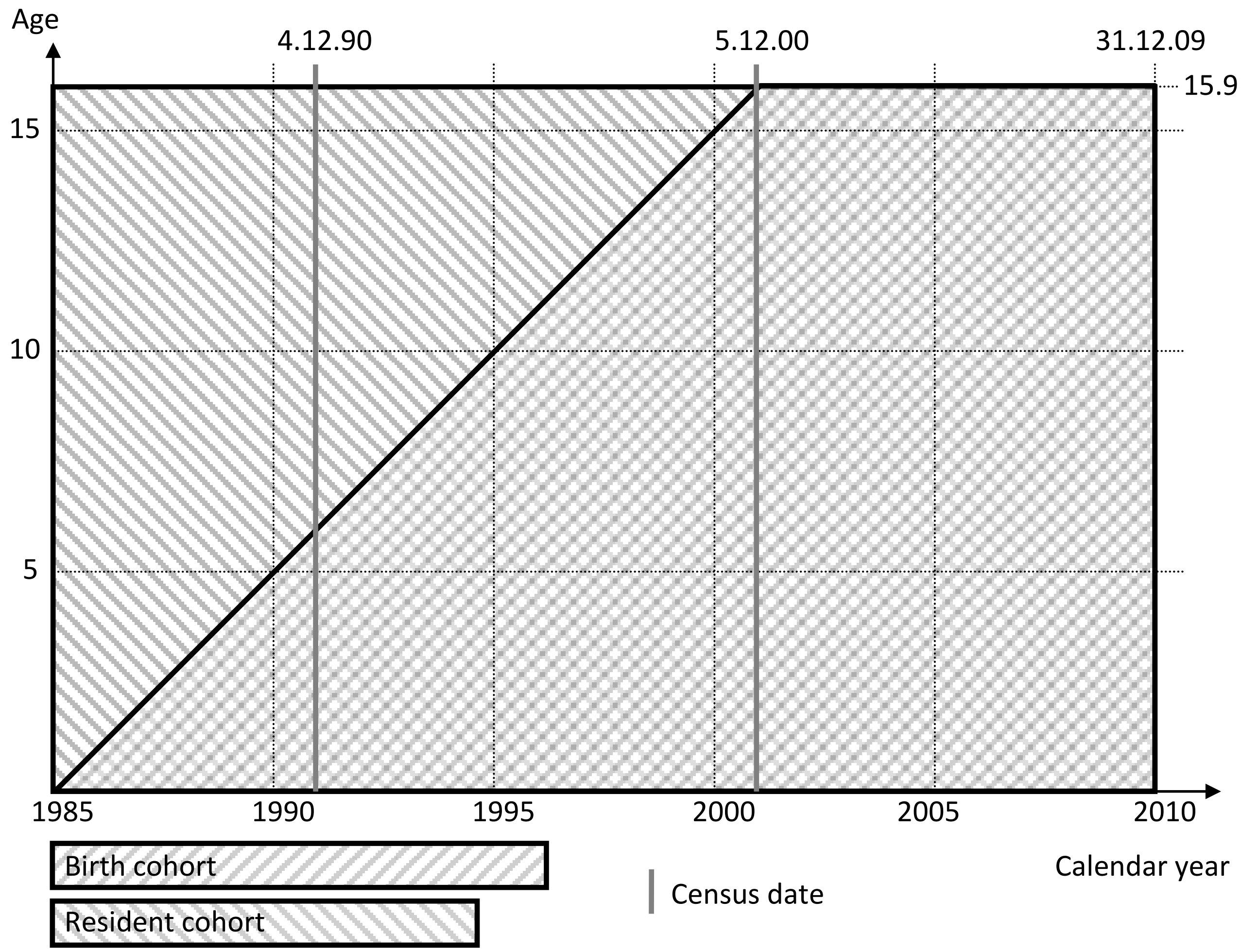
Birth cohort

- Born in Switzerland between 1.1.1985 and 31.12.2009
- Follow-up time begins on date of birth and ends when the child becomes 16 years old

Resident cohort

- Resident in Switzerland for some time between 1.1.1985 and 31.12.2009
- Follow-up time begins when a child becomes a resident of Switzerland and ends when the child becomes 16 old or leaves Switzerland.

Figure 1: Populations used in the CANUPIS study on a timeline



Children born before 1985 were excluded from the birth cohort because the ascertainment of addresses at birth and back calculation of person-years (to 1970 for children diagnosed at age 15 in 1985) proved difficult.

Calculating person years

Our analysis required calculating person years at risk (PY) for each combination of the predictors included in the Poisson regression models (exposure, sex, age, calendar year and potential confounder). Let $x_t(u)$ represent the number of PY for a particular combination representing the age group *u* and calendar year *t*, i.e. $x_t(u)$ is the number of PY in year *t* of children who were aged *u* in *t* for a given sex and distance category from NPPs. We computed $x_t(u)$ using numbers using data from the censuses in 1990 and 2000. This involved 3 steps which are explained in the 3 sections below.

The same method for computing PY was used in the birth and the resident cohort. However, in the birth cohort (main analysis), we only computed PY for <1year olds, $x_t(0)$, using the steps below. We then considered the cohort size as fixed as the cohort grew older so that PY at later ages, $x_t(u)$($u > 0$), were set equal to $x_t(0)$. Thus, we ignored attrition caused by death or emigration. For the resident cohort, $x_t(u)$ was variable and calculated anew for each age using the steps below.

1. Computing PY for the census years 1990 and 2000

We first calculated the number of PY in the years 1990, $x_{90}(u)$, and 2000, $x_{00}(u)$, using census data. The censuses represent the Swiss population on the reference dates 4/12/1990 and 5/12/2000 respectively. For each child registered in a census, we calculated the PY contributed to each age group during the 365 days preceding the reference date. For instance, a child recorded in the 1990 census with date of birth 15/May/1987 would have contributed 162/365 PY (from 4/12/1989 to 14/May/1990) to the age group *u* = 2, 203/365 PY (15/May/1990-3/12/1990) to the age group *u* = 3 and 0 PY to all other age groups. To obtain $x_{90}(u)$ and $x_{00}(u)$, we summed all contributions to age group *u* from the respective censuses.

2. Computing PY before, between and after census years

We calculated the number PY in the years between the censuses (1990 < *t* < 2000) by linear interpolation and PY in the years before (*t* < 1990) and after the censuses (*t* > 2000) by linear extrapolation from the numbers of PY obtained for the census years. This consisted of computing a weighted average of the PY values in 1990 and 2000 using following weights:

$w_{t,90} = \frac{2000 - t}{10}$ Weight applied to PY in 1990 (1 in 1990, 0 in 2000),

$w_{t,00} = \frac{t - 1990}{10}$ Weight applied to PY in 2000 (0 in 1990, 1 in 2000).

We compared two different methods of inter-/extrapolation. In the first method (inter-/extrapolation by age group) which was used in the main analysis, the number of PY in *t* for children aged *u*, $x_t(u)$, was computed from the number of PY of children aged *u* in the census years, $x_{90}(u)$ and $x_{00}(u)$. In a sensitivity analysis we compared results using this method to an alternative method (inter-/extrapolation by cohort) in which $x_t(u)$ was computed from the number of PY lived by the cohort of children aged *u* in *t* in census years, $x_{90}(u - [t - 1990])$ and $x_{00}(u - [t - 2000])$. Obviously this method cannot be computed for children born after 2000 because they are not represented in any of the censuses and for children born in 1991-2000 we used the number of PY of their respective cohorts in 2000 without interpolating as they were not represented in the 1990 census. Formulas and examples are given below and an illustration of the two methods is given in Figure 2.

1.1. Inter-/extrapolation by age group (main analysis):

Defined for all combinations of *u* and *t*:

$$x_t(u) = w_{t,90}x_{90}(u) + w_{t,00}x_{00}(u)$$

For instance, the number of PY for 6 year olds in 1996, $x_{96}(6)$, was computed by linear interpolation of the number of PY for 6 year olds in 1990, $x_{90}(6)$, and 6 year olds in 2000 $x_{00}(6)$ (Figure 2 below).

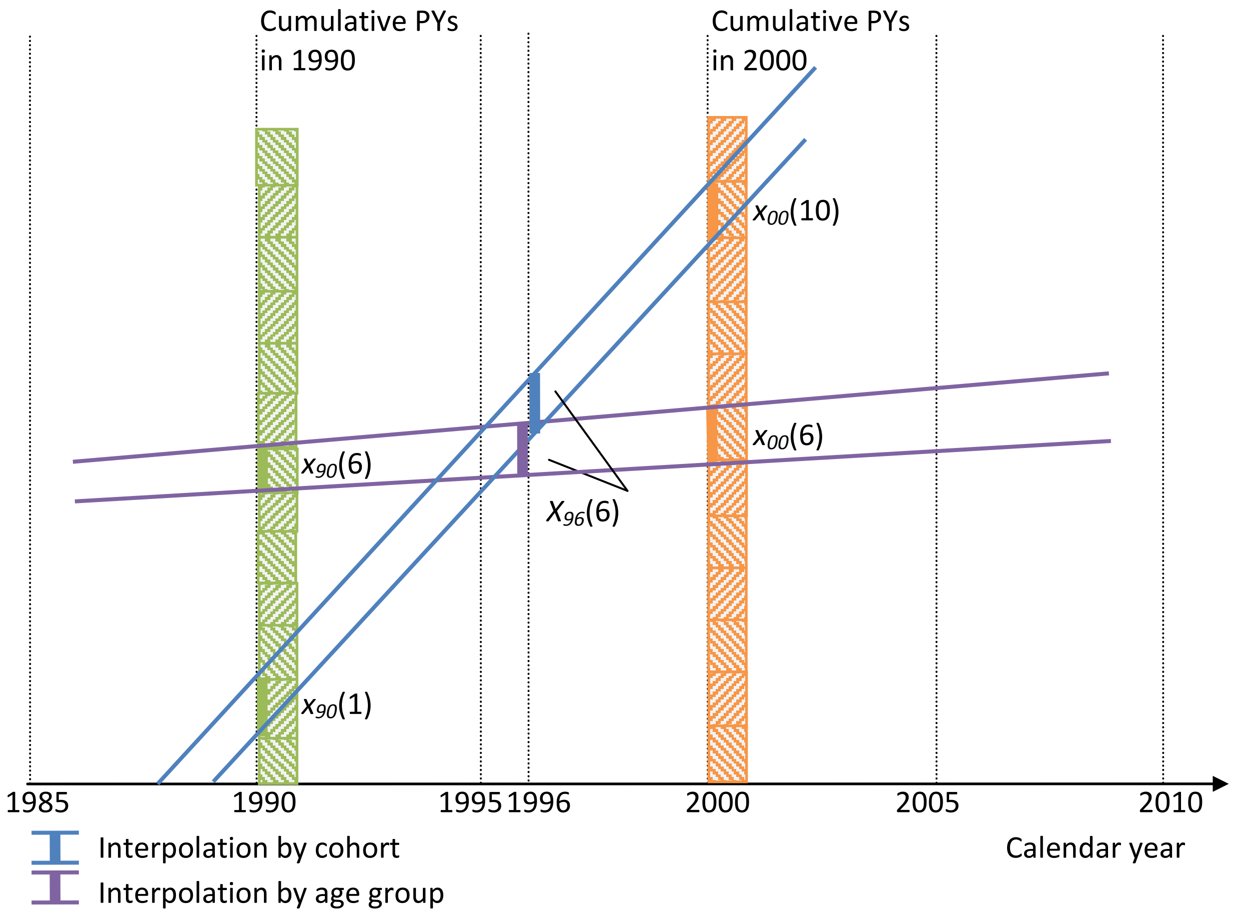
1.2. Inter-/extrapolation by cohort (sensitivity analysis):

Defined only for $u \geq t - 2000$:

$$x_t(u) = \begin{cases} w_{t,90}x_{90}(u - [t - 1990]) + w_{t,00}x_{00}(u - [t - 2000]) & \text{if } u \geq t - 1990 \\ w_{t,00}x_{00}(u - [t - 2000]) & \text{if } t - 2000 \leq u < t - 1990 \end{cases}$$

For instance, the person years for 6 year olds in 1996, $x_{96}(6)$, was computed by linear interpolation of the number of PY for 1 year olds in 1990, $x_{90}(1)$, and 10 year olds in 2000, $x_{00}(10)$ (Figure 2 below).

Figure 2: Illustration of PY interpolation by age group and by cohort for the example of 6 year olds in 1996



Using $w_{t,90} + w_{t,00} = 1$ it is easy to see that both these methods represent linear inter/extrapolation of the two values, x_{90} and x_{00} (*u* omitted):

$$\begin{aligned} x_t &= w_{t,90}x_{90} + w_{t,00}x_{00} \\ &= x_{90} - (1 - w_{t,90})x_{90} + w_{t,00}x_{00} \\ &= x_{90} + w_{t,00}(x_{00} - x_{90}) \end{aligned}$$

3. Adjusting for population fluctuations

Because the inter-/extrapolations above are linear, they cannot account for non-linear trends in the population throughout the study period. We obtained national population levels by age group, sex, and calendar year from the Federal Office of Statistics and used these to adjust for non-linear trends as explained in the paragraphs below.

Note that these corrections only adjust for population trends on a national level. Population trends at a local level may not have been parallel and may have shifted the spatial distribution of the population in relation to NPPs. We could not fully account for this because population numbers were not available at a sufficiently high spatial resolution for the study period. However shifts in the spatial distribution were accounted for to the extent that they were captured by the censuses and by linear inter-/extrapolations from census values.

Population adjustments in inter-/extrapolation by age group:

In the computation of $x_t(u)$, the values $x_{90}(u)$ and $x_{00}(u)$ were first multiplied by indices $I_{t,90}(u,u)$ and $I_{t,00}(u,u)$ respectively, where $I_{t,yy}(u,u)$ represents the ratio between the population aged *u* in year *t* and the population aged *u* in the census year *yy*.

Population adjustments to inter-/extrapolation by cohort:

In the computation of $x_t(u)$, the values $x_{90}(u\text{-}[t\text{-}1990])$ and $x_{00}(u\text{-}[t\text{-}2000])$ were first multiplied by indices $I_{t,90}(u,u\text{-}[t\text{-}1990])$ and $I_{t,00}(u,u\text{-}[t\text{-}2000])$ respectively, where $I_{t,yy}(u,u\text{-}[t\text{-}yyyy])$ represents the ratio between the Swiss population aged *u* in *t* and the population aged *u*-[*t*-*yyyy*] in the census year *yy*.

Data sources and definitions of variables used in analyses

Exposure and potential confounders	Data source	Variable definition
Nuclear power plant (NPP)	Geo-coded manually using the fixpoint data service (FPDS) of the Federal Office of Topography (swisstopo) http://map.fpds.admin.ch	Nearest distance (categories: <5km, 5-10km, 10-15km, >15km)
Other nuclear facilities	Geo-coded manually using the fixpoint data service (FPDS) of the Federal Office of Topography (swisstopo) http://map.fpds.admin.ch	Nearest distance (<5km, 5-10km, 10-15km, >15km)
Planned but not built NPPs	Geo-coded manually using the fixpoint data service (FPDS) of the Federal Office of Topography (swisstopo) http://map.fpds.admin.ch	Nearest distance (<5km, 5-10km, 10-15km, >15km)
Background ionising radiation including cosmic, terrestrial and artificial sources.	Swiss Radiation Maps [1] Contact person is Benno Bucher from the Swiss Federal Nuclear Safety Inspectorate (ENSI).	Modelled dosage in area grid cell (2x2km); categories: - Cosmic: <50, 50-60, 60-65, >65nSv/h - Terrestrial: <50, 50-85, 85-110, >110nSv/h - Artificial: <10, 10-30, 30-50, >50nSv/h - Total radiation (sum of cosmic, terrestrial and artificial): <100, 100-150, 150-195, >195nSv/h
Radio and TV transmitters	Modelled exposure maps from the Federal Office of Communications (BAKOM). Contact person is René Vogt.	Modeled field strength in area grid cell (<100, 100-106, >106 dBµV/m)
Multi-track railway lines	Geodaten © swisstopo Federal Office of Topography (swisstopo). Contact person is Raffael Bovier.	Nearest distance to multiple track railway line (<50, 50-200, 200-600, >600m)
High voltage power lines (≥200kV)	Federal Inspectorate for Heavy Current Installations (ESTI). Contact person is Urs Huber.	Nearest distance (<50, 50-200, 200-600, >600m)

Major road or highway	Geodaten © swisstopo Federal Office of Topography (swisstopo). Contact person is Raffael Bovier.	Nearest distance to main road (class 1) (<20m, 20-50, 50-200, >200m) or to highway (<40m, 40-100, 100-400, >400m)
Fruit tree cultivation	Arealstatistik 1985 & 1997 Data owner: Swiss Federal Statistical Office (BFS). The University of Bern holds a license for these data at the Institute of Geography. Contact person is Hubert Gerhardinger.	Nearest distance to a land use cell (200x200m) classified as fruit cultivation (<50, 50-100, 100-200, >200m)
Grape vine cultivation	Arealstatistik 1985 & 1997 Data owner: Swiss Federal Statistical Office (BFS). The University of Bern holds a license for these data at the Institute of Geography. Contact person is Hubert Gerhardinger.	Nearest distance to a land use cell (200x200m) classified as grape vine cultivation (<100, 100-250, 250-500, >500m)
Golf course	Arealstatistik 1985 & 1997 Data owner: Swiss Federal Statistical Office (BFS). The University of Bern holds a license for these data at the Institute of Geography. Contact person is Hubert Gerhardinger.	Nearest distance to a land use cell (200x200m) classified as golf course (<750, 750-1500, 1500-3000, >3000m)
Paediatric cancer centre	Geo-coded manually using the fixpoint data service (FPDS) of the Federal Office of Topography (swisstopo) http://map.fpds.admin.ch	Nearest distance (<5, 5-15, 15-30, >30km)
Socioeconomic status	Swiss Federal Statistical Office (BFS). Data available from the Swiss National Cohort (SNC).[2] Contact person is Adrian Spörri.	Community level index (population quintiles)
Average number of children aged 0-15 years per household	Swiss Federal Statistical Office (BFS). Data available from the Swiss National Cohort (SNC).[2] Contact person is Adrian Spörri.	Community level index (population-quintiles)

Degree of urbanization	Swiss Federal Statistical Office (BFS). Data available from the Swiss National Cohort (SNC).[2] Contact person is Adrian Spörri.	Community level variable (urban, agglomeration, rural)

References

1. Rybach L, Bachler D, Bucher B, Schwarz G. Radiation doses of Swiss population from external sources. *J Environ Radioact* 2002;**62**:277-86.
2. Bopp M, Spoerri A, Zwahlen M, et al. Cohort Profile: the Swiss National Cohort--a longitudinal study of 6.8 million people. *Int J Epidemiol* 2009;**38**:379-84.