**Cohorts**

|  |  |  |  |
| --- | --- | --- | --- |
|  | SNC | SAPALDIA | EPIC\_NL |
| Population | 8.5 million  (whole Swiss) | 5,000 – 10,000 | 16,440 |
| Home address, | Y | Y | Y |
| move history, reason to move | Y | Y | N |
| Education | Y | Y | level |
| Ethnitciy | Y | Y | N |
| Occupation | y | y | N |
| Time spent at home and work | maybe | yes | N |
| Traffic mode, duration | limited | yes | N |
| Income | no | yes | n |
| Access to garden | no | yes | n |
| Work address | no | yes | n |
| Notes: | Administrative  Census-based  Aggregated level | Self-reported | demographic characteristics.  physical activity |

Contact Ayoung about SAPALDIA, with a research plan and see how data can be shared.

SNC on the Swiss TPS side.

Detailed software, algorithm, capacity implementation

EPIC-NL may be run on server of Utrecht.

**General design and aims:**

In this study, we will focus on the SPALDIA cohort for model development, activity pattern discovery, testing, and the scalability of the developed method. From the ABM (agent-based modelling) developed using SPALDIA, we will gradually add in more assumptions, and patterns/relationships derived from the dataset, to scale it to data-sparse situations.

Specifically, the following goals are to be achieved for ABM-based exposure assessment:

1. Scale the ABM model developed using SPALDIA to the national level exposure assessment.
2. Quantify uncertainties.
3. Compare exposure assessed with ABM vs. time geography models (e.g. distance decay).

For A, we plan to Find relationships between occupation, work address, education, ethnicity vs. activity patterns (travel duration and traffic model), model them (e.g. time geography, linear regression model). Then make assumptions, and use the models identified to replace the information we know and test the mode. The model we developed is probabilistic, for B. For example,

* We may find strong relationship between distance to travel vs. travel mode, and the relationship can be modelled (e.g. the probability of travelling by car increases with the distance).
* We may find relationship between occupation and ethnicity vs. travel mode, e.g. most Dutch students travel by bike or train.

For B, as information becomes sparse, we expect higher uncertainty, and it is important to provide uncertainty measures to the ABM model and the model indicating relationships found in A. As mentioned in A, this will be achieved through developing probabilistic models and use the SPALDIA dataset to validate our model.

For C, we will compare models developed using SPALDIA, using ABM, time geography functions, and hybrid models.

With the study matures, we may also enable ABM modelling of human-environment-health interactions.

**Specific design:**

***Time step:***

Focusing on weekdays for model development and validation, day and night, hourly for day time and average for night time.

***Activity:***

* *Weekday:*

Home, work, commuting. [simulations and added randomness on go to supermarket, gyms etc., with a schedule, but this part cannot be validated].

* *Weekend, holidays:*

This part cannot be validated, a purely imagined model may have little scientific values.

***Buffers and indoor-outdoor ratios:***

We can still use buffers to represent activities at home or at work locations. The indoor-outdoor ratio can be used if the information such as “access to garden” is available. If the SPALDIA is from random sampling, this part of information can be used to decide indoor-outdoor ratio, and this can be dynamic for different seasons and occupations.

***Routes:***

Routes will be extracted from OpenStreetMap. Routes can be chosen from the shortest distance routes (or routes with supermarkets on them).

**Notes:**

The model developed for SPALDIA will not be a very complicated model, mainly focusing on duration, traffic mode, home and work locations, and then find relationships between them and information available in the national cohort (SNA, EPIC-NL).

**Other data**

<https://data.geo.admin.ch/ch.bfs.gebaeude_wohnungs_register/>

Air pollution maps:

Static maps. Dynamic maps, Swiss part.

GPS tracks:

GPS track data from Roel/Gerard and exposomeNL for ABM.

Will buy 1000 tracks, for measurements campaign, 2 weeks, usb recharge, mid Januaryl

GPS tracks from Benjaming: E-CARAS, not in SPALDIA? Not processed. 40 subjects, healthy, elderly (above 50).

**Involving GPS tracks for updating probability.**