# Weekly Report (5.2-5.15)\*

Yanpei Cai<sup>†</sup>

May 16, 2025

During the past two weeks, I have managed to learn how to calibrate our smoking contagion model with BOLFI. The code has been constructed in a structured way which can be extended to multiple cases. I have also managed to retrace the smoking trend data for the UK between 1974 and 2023 from the Office of National Statistics reports. There have been some doubts casted in these processes as well.

### Coding

#### Merge ABM into BOLFI

It took some time to change the structure of the original code which allows us to put our ABM into the elfi simulator. The simulator can now give data in the same format as in the observed data from the Office of National Statistics reports, allowing more flexibility in choosing summary statistics and measures of discrepancy.

#### **BOLFI**

Priors were chosen based on the results from Adarsh et el..

**Questions:** Since our observed data are multi-column data, I initially choose the summary statistics as calculating the differences for each entry, but it didn't work very well. I wonder what is the more reasonable way to choose summary statistics for our observed data?

## Smoking Trend Data from the Office of National Statistics

#### **Data Validity**

I retraced the smoking trend data between 1974 and 2023 from the Office of National Statistics reports and found there are some differences between Adarsh's data and mine (though they can be omitted in some senses).

#### **Data Trend**

Based on the observed data, it is obvious that the percentage of never smokers and quitters has kept increasing since 1974, while the percentage of smoker has kept decreasing from 45.6% to 10.5%.

<sup>&</sup>lt;sup>†</sup>Artificial Intelligence and its Applications Institute, School of Informatics, The University of Edinburgh **Email Address:** yanpei.cai@ed.ac.uk

<sup>\*</sup>This is a summary of the works was done in the past week.

Weekly Report (5.2-5.15)

**Questions:** Since there can only be state change from never smoker to smoker in our model, it is obvious that the simulated data from our model will disobey the trend shown in the real data. How do we explain this?