Using agent-based modelling to test hypotheses on the role of neighbourhood social mechanisms in the development of small-area health inequalities

Kim Alexandra Zolitschka, Oliver Razum, Odile Sauzet

ABSTRACT

Background Small-area health inequalities may originate from differentials in the spatial distribution of environmental stressors on health. The role played by neighbourhood social mechanisms on small-area health inequalities is difficult to evaluate. We demonstrate that agent-based modelling (ABM) is a useful technique to overcome existing limitations. It allows testing hypotheses that social contagion has the potential to modify the effects of environmental stressors by reducing or increasing small-area health inequalities.

Methods Parameters defining the strength of the effect of social contagion on health behaviour were used together with a stochastic model to obtain for every year the health outcome of every agent based on health the previous year, environmental stressors and health behaviour. Unequal spatial distribution of stressors was operationalised with spatial correlation structure. We measured changes in health inequalities using parameters of the spatial correlation structure of health after 10 years. In a further round of simulations, social contagion depended on the environmental stressors.

Results A social contagion mechanism led to a reduction of small-area health inequalities together with an increase in the spatial reach of the effect of environmental stressors. An association between environmental stressor and social contagion mechanism led to a stronger localisation of the effect of environmental stressors.

Conclusions Hypotheses about the role of neighbourhood social mechanisms can be tested using ABM. The respective models provide a better understanding of mechanisms in the causal chain between environmental stressors and health inequalities. This can pave the way to the development of a new type of neighbourhood-based intervention informed by social mechanisms.

INTRODUCTION

Inequalities in health outcomes have been documented at regional level where higher mortality rates or lower life expectancy can be observed within less economically developed regions as well as within cities where noise or air pollution in an area is negatively associated with health. Consequently, small-area health inequalities may originate from differentials in the spatial distribution of environmental stressors by which we mean small-area factors which are known to, or may, affect health, such as air and noise pollution, lack of infrastructure and green spaces, but also socio-economic factors like unemployment. Social structures in a neighbourhood also affect health. However, the later are difficult to operationalise in particular because they are processes, rather measurable factors. The usual methods to assess the role of determinants of health inequalities at small-area (eg, multilevel modelling) are of limited utility when it comes to understanding how a complex network of factors and processes interact at the local level.

An example of complex relations is the interaction of social mechanisms at neighbourhood level with environmental stressors on health, and its role in the development of health inequalities. One such social mechanism is social cohesion which is often operationalised as a range of concepts, from trust in the neighbours to the degree of degradation in a neighbourhood. Another is social contagion in which ‘behaviours, aspiration and attitudes may be changed in contact with […] neighbours’. Social cohesion appears to mediate the effects of various neighbourhood characteristics on health, thus showing a role in the causal pathway between...
neighbourhood characteristics and health. But while some social mechanisms can be operationalised as attributes of a neighbourhood via individual perception (social cohesion), others are processes (social contagion) and more difficult to operationalise in order to be studied within the framework of regression analysis. Accordingly, in a recent literature review, we could not identify any studies with a focus on the effect of social contagion as a mechanism in the production of small-area health inequalities.

However, some evidence on the role of social contagion in a small-area context has been established. For example, Hedström et al showed how social contagion occurs in social networks with political interests. Pan et al showed how behaviour can spread in crowds and during emergency evacuations. These studies were performed using computer simulation.

A complex system modelling approach has been advocated by the Network on Complexity, Inequality and Health to better understand the effects of the interactions between processes and factors within a neighbourhood. The network advocates focussing on—among others—using tools that allow capturing outcomes produced by many interacting variables. These tools should model the dynamics arising from individuals interacting in different social subgroups and social networks as well as the dynamics from complex casual patterns. A tool that enables such an approach is agent-based modelling, a computer simulation method based on stochastic models. The principle of agent-based modelling (ABM) is that the action of interacting agents (individuals) is simulated according to some stochastic models which allow to observe outcomes at population level (eg, health inequalities). This approach has been advocated to further develop knowledge about contextual effects on health inequalities. ABM is commonly used to evaluate public health interventions with the aim to model the interaction with a range of factors and stakeholders. In this work we show how ABM can be used to better understand the role of certain processes within neighbourhood. In particular we want to test hypotheses about the role of social processes which cannot be operationalised as factors.

Focusing on small-area health inequality, we consider a micro–macro model approach in which actions at the individual (micro) level have effects at the population (macro) level. Using this framework, we aim to investigate if social contagion has the potential to modify the effects of environmental stressors, such as to reduce or increase small-area health inequalities due to a differential in the spatial distribution of those environmental stressors.

METHOD

For this, we use an approach in which the only action of agents is to adapt their health behaviour to the health behaviour of their neighbours through social contagion. Health and health behaviour of agents is recalculated at every iteration of the model, representing 1 year. Health behaviours, the mechanisms of social contagion and stressors and their effect on health are independent from each other, and the same model applies to all agents. This approach enables us to isolate the effect of social contagion on health behaviour, which would be far more complex in a real world. We check if this mechanism can be either an underlying cause of, or a protector against health inequalities. In a second step, we investigate how some form of association between social contagion and stressor modify the results obtained in the first part.

In this section, we first describe the empirical background for the models used, then the baseline model, followed by the models used to obtain the health outcome of each agent at each year of iteration. Finally, we describe the method of assessment of the simulation, in particular how changes in health inequalities are measured.

Definitions and empirical background of the simulation model

Environmental stressors and health: By environmental stressor we mean a factor of the physical world which operates at small-area level and negatively affects health. In the simulations we do not need to specify what the stressor is, only its relation with health is assumed. We based our relational model on empirical observations: In an analysis of data from the German Socio-Economic Panel (SOEP), with over 10 000 survey participants (years 2004 to 2014), we showed in previous work that a range of factors are associated with the physical health component score (PCS) of the SF12 (standardised score with value ranging from 0 to 100, mean 50 and SD 10). The effect of a score based on perceived pollution, noise and lack of green spaces (range 1–5), adjusted for socio-demographic factors, was of 0.92 points of PCS (SE: 0.14).9

Health behaviour and health: Under health behaviour we include a range of personal actions which promote health (eg, engaging in physical activity, eating a healthy diet). The SOEP survey does not collect data on health behaviour so that we could not evaluate the relative effect of an environmental stressor and health behaviour on health using the same empirical data. We therefore had to estimate the effect of health behaviour on health and decided to set it five times higher than the effect of the environmental stressor. The variability of health behaviours is simulated with a spatial correlation structure with distribution parameters similar to those of health. It is to be noted that only the relative effect of health behaviour and stressor on health play a role, and not the actual distribution parameters.

Social contagion and health behaviour: Through social contagion, attitudes may be changed by contact with neighbours and Galston proposed it as one of the mechanisms at work in neighbourhoud effects. We operationalise the mechanism of social contagion in the following way: the health behaviour of an individual is additively modified yearly by a fraction of the mean behaviours of the (maximum eight) neighbours. This fraction is the varying parameter of our simulations.

Spatial distribution of environmental stressor: Our model for health inequality due to the unequal spatial distribution of health outcome is based on the spatial correlation structure of health outcome as highlighted in Breckenkamp et al. In our simulations, we assume that health inequalities are due to an unequal spatial distribution of environmental stressors. Therefore the spatial distribution of the environmental stressors is based on the typical distribution of health outcomes as seen in SOEP Data (unpublished) as well as in Breckenkamp et al.

Baseline

Agents are placed on a grid coordinate system with 600×600 points. Each point on the grid represents a possible position (place of residence) for a single agent. Twenty per cent of points (72 000) are left empty, so there remain 288 000 agents. At baseline, two variables are simulated using a normal distribution and a correlation structure: Environmental stressor and the health behaviour. The correlation structure of the health outcome (which is standard normally distributed at baseline)
is obtained by running five rounds of simulation without social contagion to create health inequalities due to the environmental stressor only. All distribution and spatial correlation parameters are given (after five rounds of simulations) in table 1.

### Agent-based modelling, first stage

In this first round the health behaviour and stressor are independent from each other and the variability of the results is evaluated. The simulations are based on a yearly calculation of the health outcome and yearly calculation of an indicator of health behaviour. The health outcome is a standard normal outcome obtained from each other and the variability of the results is evaluated.

The health outcome is recalculated according to the following linear stochastic model dependent on health in the previous year, effect of the stressor and effect of health behaviour of the agent.

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.

The health behaviour of an agent is affected via social contagion by the health behaviour of its neighbours and is yearly (ie, at each iteration) revaluated. The health behaviour in year \( t \) depends on the health behaviour in year \( t \) and on the health behaviour of a maximum of eight adjacent neighbours (a so called Moore neighbourhood):

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.

The process of social contagion for an individual agent is illustrated in figure 1. The dynamic of the simulation model is shown in figure 1.

Health \( H_{i,t+1} \) of agent \( i \) in year \( t + 1 \) is given by the following linear model:

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.

The health behaviour of an agent at baseline depends on the health of agent \( i \) and on the health behaviour of the \( n \)th neighbour in year \( t \).

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.

The health behaviour of an agent at baseline depends on the health of agent \( i \) and on the health behaviour of the \( n \)th neighbour in year \( t \).

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.

The health behaviour of an agent at baseline depends on the health of agent \( i \) and on the health behaviour of the \( n \)th neighbour in year \( t \).

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.

The health behaviour of an agent at baseline depends on the health of agent \( i \) and on the health behaviour of the \( n \)th neighbour in year \( t \).

\[
H_{i,t+1} = H_{i,t} + \beta_S S_i + \beta_B B_{i,t+1} + \epsilon_{i,t+1}
\]

Where \( H_{i,t} \) is the health of agent \( i \) in year \( t \) is, \( \beta_S S_i \) is the effect on agent \( i \) of the stressor, and \( \beta_B B_{i,t+1} \) the effect of the health behaviour of agent \( i \) in year \( t + 1 \). \( \epsilon_{i,t+1} \) is a uniformly distributed term between \(-0.5 \) and 0.5. The error variance was chosen after calibration (simulation run with different regression coefficients) to provide stable models.
health outcome spatially correlates with the health outcomes of its neighbours. Reported parameters are the partial sill (spatial variance of health outcome), the nugget effect (non-spatial variance), the relative structured variability (RSV, ratio of partial sill to the total variance, measure of the part of the total variance which is spatially structured), and the range (two observations at distances above this value are virtually uncorrelated).

If the RSV declines relative to baseline after 10 iterations, then the part of the total variance of health, which is spatially structured, is smaller than at baseline. This would indicate that the spatial effect on health are smaller as well as the resulting health inequalities. Moreover, if the range increases after 10 iterations, then the range at which the environmental stressor as an effect, increases, thus reducing the inequalities inherent to living in an area of high stressor compared with an area of low stressor.

RESULTS
Results of the simulations
After five initial runs, the spatial correlation structure for health acquired an identical spatial correlation structure as the stressor with the parameters given in Table 1. This correlation structure defines the baseline health inequalities to form the reference for comparisons 10 years later (after 10 model iterations). We observed further changes in the spatial correlation structure of health without social contagion after 10 iterations following the initial runs: the RSV for health decreased from 54% to 43% (reference).

When we increased the effect of social contagion on health behaviour (contagion factor), we observed an increased reduction in the spatial correlation structure of health between the initial value and after 10 iterations. The RSVs was as high 43% for the control model (no contagion) after 10 iterations but declined from 54% at baseline to 4% after 10 iterations for a contagion factor 0.4, thus observing practically no correlation.

For a contagion factor of 0.05, we observed a reduction in the spatial correlation structure of health outcomes by modifying health behaviour over time, thus affecting health inequalities. Moreover, if the range increases after 10 iterations, then the range at which the environmental stressor as an effect, increases, thus reducing the inequalities inherent to living in an area of high stressor compared with an area of low stressor.

Relation between simulation results and empirical observations
Spatial health inequalities are due in part to a differential in the spatial distribution of factors affecting health, for example, noise and air pollution.

In addition, social contagion among neighbours, a small-area mechanism, plays a role in the spatial distribution of health outcomes by modifying health behaviour over time, thus affecting health inequalities. The starting point of the simulations was that health inequalities were solely due to an environmental stressor and to a smaller extent to health behaviour. After 10 years of social contagion operating, we observed reduced inequalities (with a degree varying with the intensity of social contagion).

| Table 2 | Mean and SD of the health outcome taken over 20 samples of 30,000 positions for each simulation (second phase, in italics) and including 100 repeated simulations for the first phase after a 5-year initialisation and 10 years of follow-up |
|---------|-------------------------------------------------|----------------|----------------|----------------|
|         | Partial sill | Nugget | RSV | Practical range |
| Initial values | 0.543 | 0.461 | 0.541 | 184 |
| Positive control | 0.431 (0.003) | 0.569 (0.003) | 0.430 (0.003) | 182 (1.64) |
| Social contagion | | | | |
| 0.001 | 0.428 (0.010) | 0.573 (0.009) | 0.428 (0.009) | 185 (7) |
| 0.005 | 0.429 (0.012) | 0.572 (0.011) | 0.429 (0.011) | 185 (6) |
| 0.01 | 0.361 (0.005) | 0.637 (0.005) | 0.361 (0.005) | 193 (4) |
| 0.05 | 0.219 (0.007) | 0.784 (0.007) | 0.291 (0.007) | 201 (6) |
| 0.1 | 0.104 (0.002) | 0.899 (0.002) | 0.104 (0.002) | 230 (6) |
| 0.2 | 0.041 (0.002) | 0.962 (0.002) | 0.041 (0.002) | 304 (20) |
| 0.4 | 0.041 (0.006) | 0.962 (0.002) | 0.041 (0.003) | 301 (23) |
| Alternative models | | | | |
| 0.05×0.1×stressor | 0.265 (0.013) | 0.733 (0.012) | 0.266 (0.012) | 182 (9) |
| 0.05×(−0.1)×stressor | 0.267 (0.010) | 0.733 (0.011) | 0.267 (0.008) | 186 (11) |
| 0.05 behaviour×stressor | 0.230 (0.009) | 0.768 (0.008) | 0.230 (0.008) | 195 (9) |
| RSV, relative structured variability. | | | | |
Any spatial correlation structure in health outcomes empirically observed only partially reflects the effect of a differential in spatial distribution of small area factors affecting health. Other mechanisms such as social contagion contribute to a reduction of these effects. Moreover, increasing the strength of social contagion will weaken the spatial correlation structure in health outcomes.

**DISCUSSION**

We used ABM to test the hypothesis that social contagion affects the development of health inequalities due to a differential in the spatial distribution of environmental stressors. While there have been a number of studies using ABM to investigate health behaviour, to our knowledge the first attempt to understand the effect of social mechanisms in the neighbourhood on small-area health inequalities using a simulation approach. We chose linear stochastic models to define the relationship between stressors, social contagion and health. Social contagion had an effect on the health behaviour, but social contagion did not have a direct effect on health but affected a measure of health behaviour that, in turn, did affect health. Social contagion did not affect the spatial correlation structure of health outcomes empirically observed only partially reflects the effect of a differential in spatial distribution of small area factors affecting health. Other mechanisms such as social contagion contribute to a reduction of these effects. Moreover, increasing the strength of social contagion will weaken the spatial correlation structure in health outcomes.

Our model implies that health inequalities slightly decreased over time even in the absence of social contagion because health behaviour was also spatially structured. Compared with the scenario with no social contagion, an annual modification factor of 1% of the health behaviour via social contagion, when social contagion and environmental stressors are independent from each other, leads to a decrease of 16% in the part of the variance of the health outcome which is spatially structured. This proportion increases to 90% for a modification of 20%: in this case, only little spatial correlation structure of health outcomes remains. The practical range (distance between neighbours beyond which the outcomes are no more correlated) does increase from 180 to 300 m. This means that the effect of the environmental stressor on health inequalities could become spatially wider-reaching (i.e., there is less clustering of health outcomes), as well as weaker when less of the variability of health outcomes is spatially structured.

In the two scenarios with a dependence of the social contagion mechanism on the environmental stressor, a slightly greater reduction of the RSV is seen. This time, however, it is associated with a reduction of the range, meaning that the effect of the stressor becomes more localised. Understanding the effects of one particular process necessitates the isolation of this process in a highly simplified model. Thus, the effects seen can be attributed to the process of interest. Because we wanted to test a hypothesis, we used a simple model in which heterogeneities in the population did not play a role to be able to isolate the consequences of various strengths of effect of social contagion on health inequalities. In our simulations, social contagion did not have a direct effect on health but affected a measure of health behaviour that, in turn, did affect health.

Agent-based modelling shows potential for studying the role played by social mechanisms in the production of small-area health inequalities. It can be used to test hypotheses about models at the micro level by checking their effects at the macro level. Further models should...
Investigate thresholds of micro level model parameters to show either a relevant effect at macro level, or to fit with observed empirical data.

Compare the effects on competing models to see which model best explains the empirical data.

Our study has several limitations. First, we did not consider competing models for social contagion because it was beyond the scope of our investigation to see if social contagion could have a modifying effect on small-area health inequalities. Then we based our model parameters on empirical effects only in a limited way. A next step should be to design agent-based models based or confronted to empirical data as mentioned above.

Social contagion can have positive as well as negative effects on health behaviour depending on the locally dominant health behaviour. However, reducing small-area health inequalities is not about improving health for a few but about reducing the impact of neighbourhood effects on individual health. In models where the process of social contagion was dependent on the stressor, no reduction to the mean occurred (the variance of health behaviour remained unchanged) but health inequalities were reduced.

A better understanding of the role social contagion performs in attenuating the effects of environmental stressors will help better understand how small-area based interventions might work toward reducing small-area health inequalities. This can be achieved by developing intervention to facilitate the positive aspects of social contagion within a neighbourhood. ABM can help reach this goal.

**Contributors** OS and OR conceptualised the study and contributed to the manuscript. KAZ performed, and analysed the simulations, and wrote a first draft of the manuscript. OS is the guarantor of this work.

**Funding** This work was supported by Deutsche Forschungsgemeinschaft (DFG), grant number 336562538.

**Disclaimer** The funder played no role in study design, analysis and reporting.

**Competition interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. Not applicable.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**ORCID ID** Odile Sauzet http://orcid.org/0000-0002-1029-8846

**REFERENCES**