Population Generation for Agent-based Simulations of Stroke Logistics Policies

A Case Study of Stroke Patient Mobility

Abdulrahman Alassadi, Fabian Lorig, and Johan Holmgren
Department of Computer Science and Media Technology
Faculty of Technology and Society, Malmö University
Malmö, Sweden
e-mail: alassadi@live.com, fabian.lorig@mau.se, johan.holmgren@mau.se

***Abstract*—For acute medical conditions, for instance strokes, the time until the start of the treatment is a crucial factor to prevent a fatal outcome and to facilitate the recovery of the patient's health. Hence, the planning and optimization of patient logistics is of high importance to ensure prompt access to healthcare facilities in case of medical emergencies. Computer simulation can be used to investigate the effects of different stroke logistics policies under realistic conditions without jeopardizing the health of the patients. The success of such policies greatly depends on the behavior of the individuals. Hence, agent-based simulation is particularly well-suited as it imitates human behavior and decision-making by means of artificial intelligence, which allows for investigating the effects of policies under different conditions. Agent-based simulation requires the generation of a realistic synthetic population, that adequately represents the population that shall be investigated such that reliable conclusions can be drawn from the simulation results. In this article, we propose a process for generating an artificial population of potential stroke patients that can be used to investigate the effects of stroke logistics policies using agent-based simulation. To illustrate how this process can be applied, we present the results from a case study in the region of Skåne in southern Sweden, where a synthetic population of stroke patients with realistic mobility behavior is simulated.**

Keywords-Agent-based Simulation; Synthetic Population; Population Generation; Policy Making; Mobile Stroke Unit.

I. Introduction

Certain medical conditions, for instance strokes, require rapid responses of medical professionals to prevent a fatal outcome and to facilitate the recovery of the patient's health. However, many victims of potentially life-threatening conditions such as strokes are often not surrounded by medical professionals when these conditions occur as they may appear suddenly and without prior indications. Accordingly, to prevent permanent damage from such acute medical conditions or even the patients passing away, it has become a goal of our society to provide adequate care to all citizens as quickly as possible. This includes the development of an advanced medical infrastructure that provides necessary health care services but also to ensure that this infrastructure is available for everyone. Some treatments of medical emergencies, however, require specialized equipment and personnel, which is only available in certain facilities (e.g., specialty hospitals). Especially in rural areas, there is a lack of such facilities. Hence, the planning and optimization of patient logistics is of high importance to ensure prompt access to healthcare facilities in case of medical emergencies.

In a study presented at the *IARIA DIGITAL 2021* conference on *Advances on Societal Digital Transformation*, Alassadi et al. [1] address the challenge of generating a realistic population of stroke patients, which takes travel behavior into account. Such an artificial population of stroke patients is required, for instance, in agent-based simulations (ABS) and allows for the assessing the effects of different stroke logistics policies. One example of such a policy is the optimal placement of Mobile Stroke Units (MSUs) across a region. The study uses aggregated and individual-based data from different sources, from which probability distributions can be derived to generate an artificial population of agents.

In their study, the authors focus on strokes, which is a common cause of mortality [2]. Every year, more than 1 million people in the European Union suffer from a stroke and the one-month case-fatality is up to 35% [3]. The occurrence of strokes is associated with the age of the individual and most of those suffering from a stroke are 70 years of age or older. Hence, as the number of people that are older than 70 years will increase, the number of strokes is also expected to increase [4], [5].

To investigate the suitability and effects of different stroke logistics policies for a specific region, computer simulation can be used. They allow for investigating different policies under realistic conditions without jeopardizing the health of the patients. Instead of conducting real-world studies, different policies and their potential effects can be studied and compared in an artificial system. To this end, simulation enable *what-if* analyses of different scenarios, to analyze how different circumstances affect the behavior of a system without actually interfering with the real-world system that shall be investigated.

The use of simulation in healthcare is well established. Barnes et al. [6], for instance, provide a comprehensive overview of how simulation can be applied in healthcare operations management and underline the successful application of simulation for evaluating policy alternatives. An increasing application of simulation in healthcare has also been identified by Almagooshi [7], e.g., for the analysis of patient flows, emergency departments, and treatment of, e.g., stroke.

The effects and success of logistics policies greatly depends on the behavior of individuals that are affected by the policy. Hence, to investigate how different policies affect the accessibility to healthcare services for patients, the behavior and routines of the potential patients need to be represented in the simulation model. We argue that ABS is particularly well-suited to investigate the effects of logistics policies, not only in terms of strokes but also for other acute medical conditions.

In ABS, an artificial population is generated, which consists of so-called *agents*. Agents are characterized by individual attributes such as age, gender, place of residence, and health state, and imitate human behavior using Artificial Intelligence. A major challenge when using ABS to analyze logistics policies, for instance for the treatment of strokes, is the generation of an artificial population of patients that imitate the behavior of the real-world population such that the effects of the policies can be investigated. This might include deterioration of health condition according to individual attributes such as, e.g., age, gender, but also the modeling of the patients' whereabouts. For instance, when investigating the placement of MSUs for the treatment of strokes, the locations of the MSUs should be determined such that the time to treatment can be reduced for all inhabitants of the region. For this purpose, Amouzad Mahdiraji et al. [8] studied the average time to treatment for different distributions of MSUs and showed that a small number of MSUs can indeed significantly reduce the time to treatment for most inhabitants in the region. For their study, the authors used demographic data on the inhabitants' place of residence for determining where the demand for emergency care occurs. However, this does not consider that individuals travel and might not be at home when having a stroke, for instance, due to leisure activities, shopping, or work. Yet, the spatial distribution of strokes potentially affects the suitability of different stroke logistics policies and, thus, might need to be considered when assessing their suitability.

In this article, we propose a process for generating an artificial population of potential patients that can be used to investigate the effects of stroke logistics policies using ABS. To illustrate how this process can be applied, we present the results from a study where a synthetic population of stroke patients with realistic mobility behavior is simulated. We apply the model to the region of Skåne in southern Sweden to investigate how travel behavior is expected to affect the spatial distribution of stroke patients. The proposed approach allows for testing different policies without jeopardizing the health of the patients. The generated synthetic population of stroke patients can be used, for instance, to assess different logistics policies, e.g., to compare different placements of MSUs and to assess how this affects the time to treatment. The process, however, is not only applicable for stroke patients but can also be adapted and applied for other acute medical conditions.

The remainder of the article is structured as follows. Section II presents related work on the use of agent-based modeling and simulation in healthcare, on policy making for treatment of acute strokes, and on methods for population generation. In Section III, the process for generating a synthetic population is presented. In Section IV, a case study is presented where the proposed process is applied for generating a population of stroke patients with travel behavior. Section IV presents and discusses the results of the case study in Skåne, Sweden, and in Section V, conclusions are drawn, and future work is presented.

II. Background

Diagnosis and treatment processes in healthcare often include multiple consecutive steps and involve different specialists and caregivers. Planning and optimizing such complex processes are challenging and requires the comparison of potential configurations under different circumstances. Evaluating these processes in the real-world prior to their implementation might not only be costly and time consuming but also pose a danger to the patients' well-being. To overcome this, simulation can be used. By building a virtual model of the real-world, an artificial system can be created to investigate different scenarios and to observe the effects different measures and decisions might have on the process of care provision.

A. Agent-based modeling and simulation in healthcare

There exist different simulation paradigms, i.e., approaches for modeling and simulating phenomena or systems. In healthcare, as well as in other domains where humans are the object of investigation, individual-based simulation paradigms are often applied. An example is ABS, a form of microsimulation, which consists of the simulation of states and behavior of individuals over time [9]. Here, each individual is represented by an agent, an autonomous entity that, for example, imitates human-like behavior and reasoning. This includes the subjective perception of the environment but also the individual decision making based on the personal traits and characteristics of each individual, which leads to individual actions and behavior.

In logistics and production, for instance, the use of simulation is well established [10]. But also in healthcare, for instance in terms of the ongoing Covid-19 pandemic, the use of simulation is feasible [11]. Cabrera et al. [12] use simulation for designing a decision support system that can provide management support for emergency departments. This is achieved by analyzing the optimal staff configuration to minimize patients' waiting time and maximize patient throughput. A more extensive simulation model of hospital processes has been proposed by Djanatliev & German [13]. It combines individual-based simulation with system dynamics for analyzing different innovative workflows prior to their implementation, e.g., prostate cancer screening and effects of MSUs on onset-to-treatment times.

Simulations of stroke treatment were presented by, for instance, Monks et al. [14] and Chemweno et al. [15]. Monks et al. investigate clinical benefits of reducing delays in thrombolysis (alteplase) of AIS patients. They propose a discrete-event simulation model of stroke patients arriving at a large district hospital, where measures can be adopted to reduce in-hospital delays (e.g., prealert of paramedics) and where certain limitations of alteplase treatment (i.e., extension of treatment deadline from 3 to 4.5 hours and patient age over 80 years) can be relaxed. To assess and compare the benefits of policies for reducing waiting times, the authors model two treatment paths, the traditional treatment and one that takes measures into account for reducing in-hospital delays. The results show that an extension of the time window in combination with reduced delays can lead to 5-times increased thrombolysis rates. Chemweno et al. present a discrete-event simulation of the diagnostic path of patients in a stroke unit to investigate the effect of different test capacities. This is to overcome shortcomings of traditional queuing theory models, which cannot predict waiting times due to the complexity of treatment pathways and interrelationships between required resources. This allows for the assessment of policy changes in capacity profiles and test resources. The study outlines the effects different policies might have on waiting times, e.g., adding extra timeslots, shifting from MR to CT scans, and implementing joint timeslots.

B. Agent-based Simulations for Policy Making

Even though the use of agent-based approaches is established in many domains and disciplines, the practical use to facilitate policy making is still limited. Most applications pursue scientific purposes, which might be due to a series of reasons, e.g., lacking trust in the models or inappropriateness of models, difficulties in developing or using the models, and a lacking awareness that the method exists.

Ruppert et al. [16] argue that a thorough analysis of policy options is required before a suitable policy can be implemented. To this end, they argue that ABS often are complex and, thus, difficult to access by policy makers. Silverman et al. [17] underline that simulations are essential to address population health challenges, as they complement traditional epidemiological toolkits. They particularly outline that additional data sources need and especially GIS information need to be integrated such that models can serve as policy sandboxes and that ABS can become a valuable tool for healthcare-related policy making. By proposing a process for generating an artificial population of patients that can be used in ABS, we contribute to lowering the threshold for using simulation in policy making.

C. Policies for Treatment of Acute Strokes

What makes the example of strokes particularly well-suited for studying patient logistics policies is that there are two types of acute strokes. Acute ischemic strokes (AIS), where a clot or narrowed blood vessel blocks the flow of blood to the brain, and hemorrhagic strokes, caused by a burst blood vessel [18]. Both types of strokes require immediate treatment and delays negatively affect the patients' outcomes. Yet, the treatment of these two kinds of strokes differs greatly. To dissolve the blood clot and to restore the blood flow, an AIS needs to be treated with thrombolytic medication as quickly as possible. In case of hemorrhagic strokes, however, there is a contraindication for thrombolysis as it might kill the patient. Instead, the effect of blood thinners must be counteracted to control and stop the bleeding. Hence, making the right diagnosis is a vital first step for the efficient treatment of strokes.

Imaging of the brain, e.g., through CT or MRI scans, and specific laboratory tests are required to adequately diagnose the cause of a stroke. However, especially in urban areas, the access to such scanners and laboratories is limited and the patient needs to be transported to a suitable hospital, causing valuable time to pass. A stroke logistics policy that can be applied to address this challenge is the deployment of Mobile Stroke Units (MSUs), which are specialized ambulances with all equipment required to diagnose stroke patients. Through this, the time between the onset of symptoms and the beginning of treatment of the stroke can be shortened, which significantly can improve the prognosis of the patients. The feasibility of this concept and its capability to prevent brain damage of stroke patients was demonstrated by Walter et al. [19].

For the treatment of acute ischemic strokes, intravenous thrombolysis to dissolve the blood clot is the only approved reperfusion treatment [20]. However, according to Fassbender et al. [21], only less than 5% of the stroke patients receive this therapy. One potential explanation is that the critical time window of 3 hours is exceeded due to the transport to the hospital. To reduce the time to treatment, the use of Mobile Stroke Units (MSUs) was proposed, i.e., specialized vehicles that are equipped with devices required for adequately diagnosing and treating stroke patients. Walter et al. [19] compared the use of MSUs to hospital treatment and found that the time from alarm to therapy could be reduced from 76 to 35 minutes. Calderon et al. [22] analyzed the worldwide status of MSUs and compared different services outlining the success of the approach. The economic viability of MSU treatment was analyzed by Kim et al. [23], underlining its cost-effectiveness due to earlier provision of therapy.

The success of MSUs and their effect on treatment times also depends on where they are located. Rhudy et al. [24] visually analyzed data of MSU dispatches and the occurrence of strokes to optimize service provision. For Sydney, Australia, Phan et al. [25] searched for optimal locations for MSUs by investigating travel times from suburbs to each potential MSU hub. For a similar purpose, Amouzad Mahdiraji et al. [8] developed an agent-based model that allows for analyzing the benefits of different MSU configurations. In their study, Amouzad Mahdiraji et al. investigated the average time to treatment for different distributions of MSUs and showed that a small number of MSUs can significantly reduce the time to treatment for most inhabitants in the region. Moreover, agent-based simulation can also be used to assess other stroke logistics policies, e.g., whether patients should be brought to the closest hospital or to a specialized thrombectomy center [26]. To assess this, Al Fatah et al. developed a simulation model of logistical operations of stroke patients, i.e., whether patients should be transported to the closest hospital or towards a stroke center. The results showed that those patients that require special treatment indeed benefit from being transported in the direction of a stroke center whereas those who do not require specialist treatment benefit from being transported to the closest hospital.

None of the presented approaches takes travel behavior into consideration when investigating and optimizing locations and service designs of MSUs. Instead, individuals are assumed to stay at their home location, which can be derived from census data or randomly selected using Monte Carlo approaches [26].

D. Population Generation

To generate realistic results when applying agent-based simulation, individuals and their behavior must be modeled in a realistic way. Especially when modeling a larger population of individuals, it is important that the relevant features of the artificial population, e.g., age distribution or employment status, correspond to those of the original population. However, due to privacy reasons, data on each individual's properties is usually not available. The challenge associated with synthetic population generation is that aggregated data, e.g., census data, and disaggregated personal data need to be combined to model each individual, such that the characteristics of the modeled population correspond to the used input data [27], [28]. In transportation, for instance, population generation is used to model individual demand for mobility services [29], [30].

III. Population Generation for Simulating Stroke Logistics Policies

The success and suitability of using ABS for policy making strongly depends on how realistically the behavior of the individuals is modelled. In this regard, it is relevant to use appropriate data and to identify all aspects of human behavior and decision-making that are important for the scenarios that shall be investigated as well as to adequately model them. Especially when using ABS for policy making, simpler models of human behavior can limit the adaptability of the individuals and, thus, limit the significance of the simulation results. This is also in line with the trend from simple and abstract models (KISS approach) to more data-based and descriptive models, taking a wider range of evidence into account (KIDS approach) [31].

Modelling human behavior often includes the use of real-world data, e.g., socio-demographic census data, for generating a realistic population. Chapuis & Taillandier [27] identify two main difficulties when generating synthetic populations, i.e., *expansion* and *harmonization* of available data, and discuss different approaches for generating synthetic populations for ABS. In particular, the authors distinguish between *synthetic reconstruction* and *combinatorial reconstruction*. Synthetic reconstruction is mostly based on probability distributions that the generated individuals need to correspond to whereas the approach of combinatorial optimization is to draw individuals from a sample and to (iteratively) modify this sample until it satisfies a given fitness criterion.

Based on a literature review, Chapuis & Taillandier conclude that population generation processes tend to be poorly described. To overcome this, we propose a process for generating a synthetic population of individuals that can be used to assess stroke logistics policies. To demonstrate the feasibility of this process, we apply it to a case study of stroke patient mobility in Skåne, Sweden.

There exists a number of different approaches for generating synthetic populations that can be used in agent-based simulation models [32]. Some approaches reply on the availability of disaggregated data, from which samples are drawn and optimized. In our study, some of the available data sources contain aggregated data, i.e., census data and stroke data, which is due to data protection reasons.

Thus, we pursue a conventional approach in this study as, for instance, described by Barthelemy & Cornelis [33]. In this approach, aggregated data, that in average describes certain characteristics of the true population, is merged with disaggregated data from a sample to generate a disaggregated dataset of a synthetic population. After identifying all relevant socio-demographic characteristics of the population, the underlying distribution of the true populations' characteristics is estimated such that the distribution of the characteristics is preserved. Following this, individuals are selected from the disaggregated sample and added to the synthetic population in accordance with the estimated distribution. The goal of this approach is not to provide a one-to-one mapping of the real population but a matching of individuals to maintain the distribution of the relevant characteristics and attributes of the original population.

For the assessment of stroke logistics policies, we first need to identify the important aspects of human behavior that need to be included. One determining factor for the prevalence of health conditions is the age of the individual and as we want to analyze logistics policies, the location of the individuals is important as well. Hence, we integrate census, travel survey, and health data to model the occurrence of a medical condition at specific locations.

The presented process is inspired by Bissett et al. [34] and consists of three major steps that gradually add complexity to the synthetic population (see Figure 1). First, all individuals are generated using socio-demographic census data to define the constant attributes, i.e., those that do not change over time such as age, gender, or municipality of residence. Census data is often aggregated on a grid-base, e.g., 1x1 km grid cells. Hence, the grid cells need to be assigned to municipalities using GIS data. These constant attributes are important when using data from different datasets for the population generation as common attributes across the datasets, e.g., age or municipality, are required for the matching process [33].

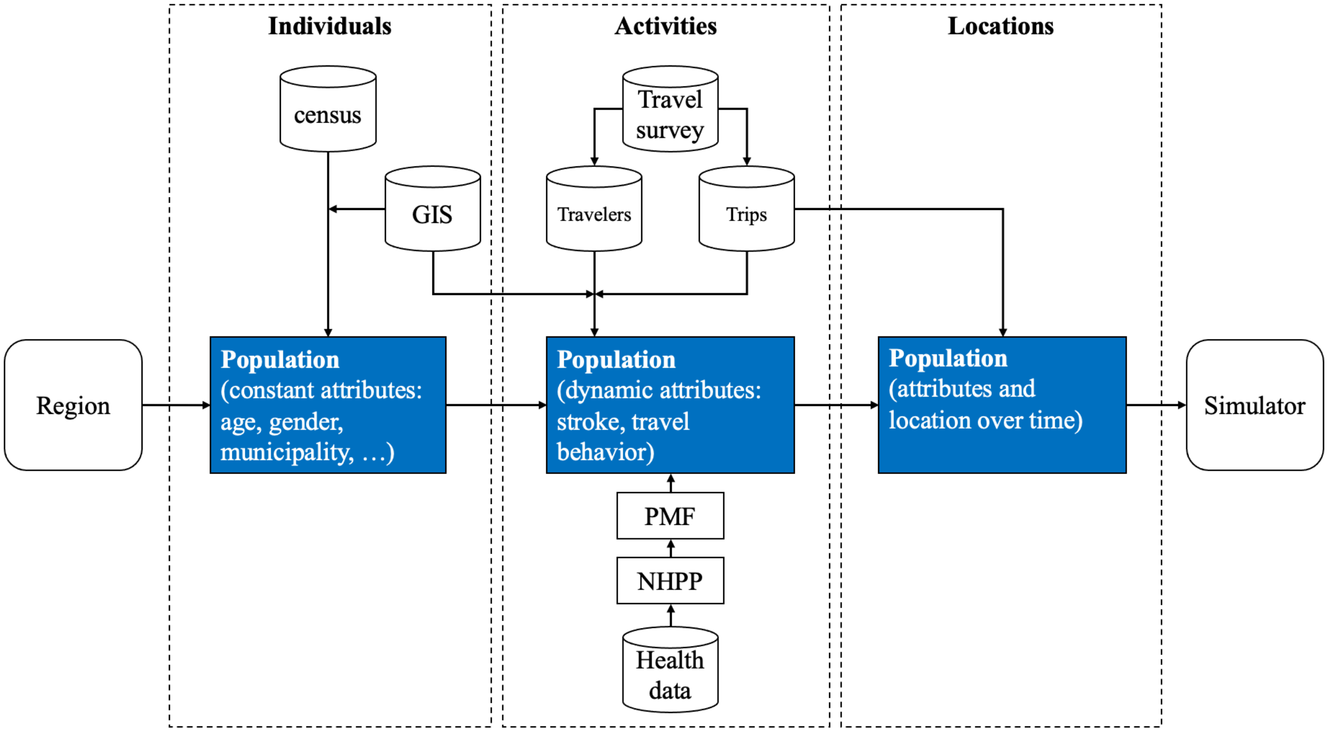


Figure 1. Process for generating a synthetic population of patients for investigating stroke logistics policies.

Individuals are then equipped with activities, which affect personal attributes that change over time, i.e., mobility patterns that affect the individuals' locations and stroke occurrences that affect their health condition. In this step, disaggregated data is used to complement the aggregated census data according to the conventional approach described by Barthelemy & Cornelis. For modeling mobility, travel survey data can be used, that consists of real-world data on travelers (e.g., age and home municipality) and the trips they have taken (e.g., origin, destination, and trip length). Joining this data with GIS data allows for generating time-dependent travel demand between zones of a municipality as well as between municipalities. The second dynamic attribute, which is coupled to the occurrence of a stroke, can be modeled by estimating the probability distribution of strokes from health data such that it can be matched according to age groups and municipalities. This is achieved by means of probability mass functions (PMF) and a non-homogeneous Poisson process (NHPP) (cf. Section IV).

Finally, based on the travel demand, the location of each individual can be calculated over time. This is required to be able to determine where an individual is located when a stroke occurs.

IV. An Agent-based Model for Generating a Population of Stroke Patients with Travel Behavior

To allow for a more dynamic and realistic assessment of health logistics policies, we proposed a process for population generation. The resulting synthetic population can be used as input to an agent-based model to investigate the effects of different policies. In this article, we use the example of stroke patients to show how such an analysis can be performed, taking travel behavior of individuals into account.

In this section, we demonstrate how we generate a synthetic population of potential stroke patients, combining socio-demographic census data, data on real strokes cases from a healthcare provider, and travel data from a transport service provider. This allows for the simulation of the dynamical spatial and the temporal component of stroke occurrence and treatment.

For the study, we selected Skåne, a region in southern Sweden. Skåne consists of approximately 1.4 million inhabitants, that live in 33 municipalities with a total area of nearly 11 000 km^2. In Skåne, there are 9 hospitals with emergency departments that can treat acute strokes. In 2015, there were 3 973 stroke incidents recorded in Skåne, out of which 3 830 patients also live in Skåne. Moreover, 12 patients that have their place of residence in Skåne were treated in the neighboring counties Kronoberg, Blekinge, or Halland and most of the patients are 45 years of age and older. Based on data from the regional healthcare provider, *Södra Sjukvårdsregionen* (Southern Health Care Region; SHR), we derived the daily distribution of strokes per hour (see Figure 2). Most strokes are reported during the afternoon with most of the incidents occurring around 4 p.m.

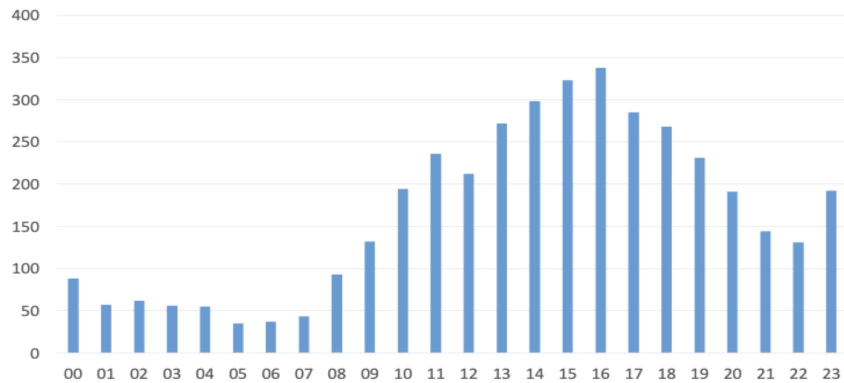


Figure 2. Total number of strokes per hour extracted from data of the regional healthcare provider.

To model travel behavior, we used data from a regional travel survey (*Resvaneundersökning för Skåne; RVU*) that was conducted in Skåne in 2013 [35]. As part of this study, travelers were asked about their traveling habits and the resulting dataset contains information on approximately 56 000 distinct trips. This includes, for instance, the origin, destination, duration, and purpose of the trip but also socio-demographic data on the travellers, e.g., age, gender, and place of residence.

Finally, for generating a realistic population, we used a census dataset from *Statistiska centralbyrån* (Statistics Sweden; SCB), the Swedish government agency for statistics. The SCB dataset includes, for instance, information of the density and age of the population of Skåne. Yet, this data only provides information on the permanent residence of individuals and not on their actual location. To allocate the anonymized trips of the RVU dataset to actual individuals from the SCB dataset, we randomly match the datasets based on the individuals' age group and home municipality.

For modeling the inter-arrival time of stroke incidents, we used a non-homogeneous Poisson process (NHPP) [36]. In contrast to ordinary Poisson processes, that are used to model events that occur with a fixed average rate of arrivals (λ), the rate of arrivals can vary over time in an NHPP where $\lambda(t)$ is the rate function for time segment t for all $t \in [0, t_0]$ and $\lambda_u(t)$ is the maximum number of actions in a time series with $0 \leq \lambda(t) \leq \lambda_u(t)$. By this means, we can explicitly model the accumulation of stroke events during the afternoon. In our NHPP, we divide each day into 24 time segments, each equipped with a specific probability that a stroke occurs during this hour in relation to the number of strokes occurring per day.

Based on the generated number of daily stroke incidents, we define two probability mass functions (PMF) to distribute strokes across age groups and municipalities. These two distributions are then used to generate stroke incidents. Each generated stroke event consists of the patient's age group, municipality, day of the year, and time of the day. This dataset is then matched with the population dataset, to predetermine the stroke patients as well as the point in time when the stroke will occur.

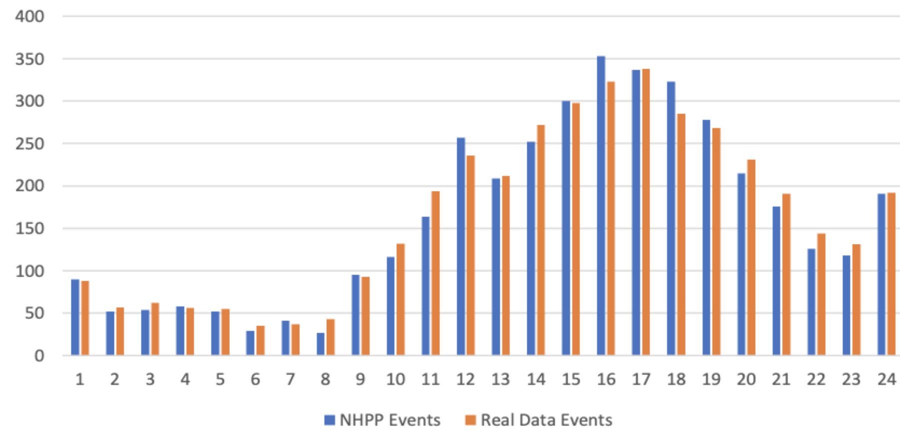
When executing the model, the travel behavior, i.e., each trip of an individual, and the resulting locations of each individual of the population is simulated over time. The generated NHPP events define when the individual stroke incidents occur, and, at the generated time of each stroke incident, the individual's current location can be determined based on the simulated trips.

In recent years, ODD (Overview, Design concepts, and Details) protocols have been used to describe the structure and the dynamics of agent-based models in a standardized document [37]. They provide more detailed insights into the model and the underlying assumptions, which can be relevant for the interpretation of the results as well as for replicating experiments. The ODD of the model presented in this article can be found in [38].

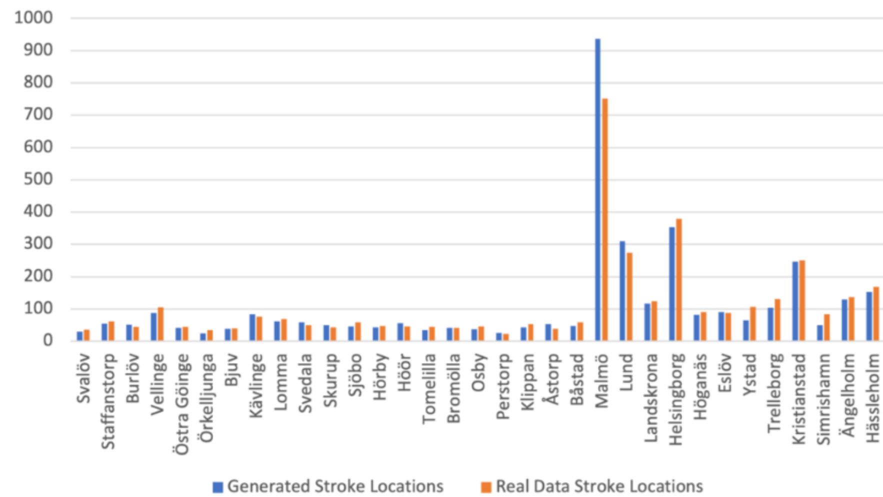
V. Results of the Scenario Study in Skåne

We implemented the agent-based model of stroke patient travel behavior in the *Repast Simphony* simulation framework [39]. In the simulation, each time unit (tick) corresponds to 1 minute in reality. Hence, each day is simulated as 1440 ticks. For each tick, it is determined whether an individual will go on a trip and move to another location. When the predetermined stroke events occur, it is checked whether the individual is on a trip, to determine where the stroke occurred.

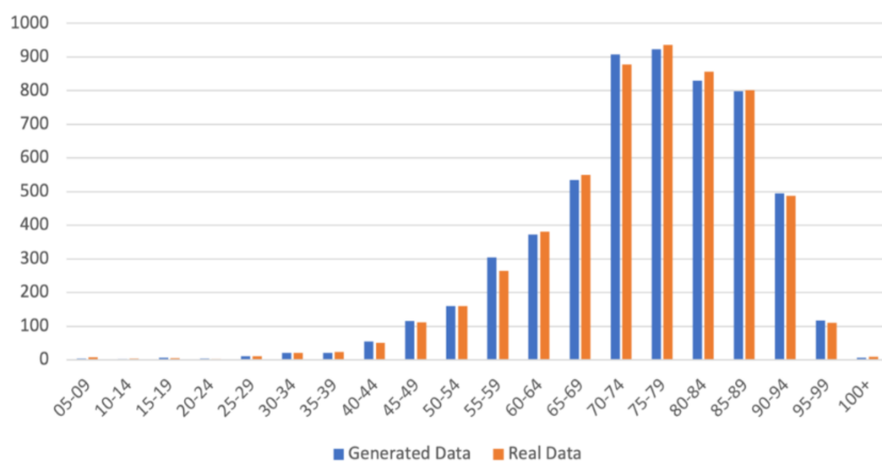
The probability distributions that we extracted from the dataset are shown in Figure 3. The charts show the real data (orange) in comparison to the NHPP events we calculated (blue). There are only minor deviations from the original data for the stroke incidents per hour, per municipality, and per age group. To quantify how well the artificial data replicates the original data, the Relative Squared Error (RSE) and Root Relative Squared Error (RRSE) are given as measures. No significant deviation of the generated data from the real data can be observed considering the hour of the day (RRSE: 0.160) and the age group (RSSE: 0.046).



(a) hour of the day (RSE: 0.026, RRSE: 0.160)



(b) municipalities in Skåne (RSE: 0.067, RRSE: 0.259)



(c) age group (RSE: 0.002, RRSE: 0.046)

Figure 3. Distribution of stroke incidents: (a) per hours of the day (b) per municipality in Skåne (c) per age group. For each distribution, the Relative Squared Error (RSE) and Root Relative Squared Error (RRSE) are given as measures of the quality of the generated data.

Only for the municipalities in Skåne, a difference can be observed for Malmö (RRSE: 0.259). This might be due to a bias, which results from Malmö's role as center of the region and as the city has notably more inhabitants compared to all other cities and municipalities in Skåne.

Instead of simulating all inhabitants of Skåne, we only simulated the trip activities of those individuals that were predetermined to suffer from a stroke. This is to reduce the computational complexity of the simulation. To reduce the effect of stochastic variations in the results, we replicated the simulation five times and calculated the average values from these runs.

On average, 3 912 strokes occur in our simulation. The results indicate that 3 839 (98.1%) strokes occur at home whereas 73 (1.9%) strokes happen while the individual is on a trip and at another location. To check the plausibility of these results and to validate the study, we compare them to existing data. In the RVU travel data, only 15% of the recorded trips are performed by individuals that are 65 years of age or older, the main risk group for suffering from a stroke. Out of these trips, only 35% are taken in the afternoon, which is the time of the day where the occurrence of a stroke is most likely.

Moreover, we analyzed the dataset of stroke incidents from SHR. Out of 3 842 stroke incidents of patients that live in Skåne, which were recorded within SHR, 3 830 actually got their treatment in Skåne. 3 106 (80.84%) of the patients that got a treatment in Skåne also got it within their municipality or at the hospital that is responsible for their municipality. Out of the remaining 736 patients (19.16%) that did receive their treatment at another hospital, 497 patients live in municipalities where the responsible hospital does not provide emergency services around the clock. Of the remaining 239 patients, 80 were treated at Skåne University Hospital, which also provides highly specialized treatments for severe cases, 57 received treatment at private facilities, whose exact location is unknown, and 59 were treated at a hospital in a neighboring municipality, which might be due to the patients living closer to the hospital in the neighboring municipality. In total, only 46 patients (1.2%) receive their treatment obviously outside their home municipality, where it can be assumed that they were traveling. This corresponds to the results of our simulation.

Figure 4 provides a visualization of the simulation results across the simulated region, Skåne. The occurrence of strokes is shown using a heatmap where the intensity of the red color indicates a higher (denser) occurrence of stroke incidents whereas brighter or completely white areas indicate that less strokes occurred or that no strokes occurred at all. Overall, the shape of the heatmap corresponds to the distribution of inhabitants across the region. In the larger cities, an accumulation of strokes can be observed (e.g., the Malmö-Lund area in the south-western part of Skåne) as well as around other larger cities such as Helsingborg (northwest), Kristianstad (northeast), as well as Trelleborg and Ystad (southern coast).

The crosses indicate the location of the hospitals. It can be observed that there are some accumulations of strokes in the central part of Skåne (Höör and Hörby) as well as in the eastern (Simrishamn) and northern (Osby and Örkelljunga) parts of the region, which do not have hospitals close by. It is in particular such areas that are dependent on sophisticated logistics policies and that can, for instance, benefit from the availability of MSUs.

Figure 5 complements Figure 4 and shows a more detailed view on the Malmö-Lund area. On the left-hand side, the heatmap of this area is shown and each red dot represents one 1x1 km grid cell as defined by the census data used for the population generation process. On the right-hand side, an ordinary map of the same area is shown. Especially in the densely populated areas of Malmö close to the seaside but also in Limhamn (west), Oxie (center), as well as Arlöv, Lomma, and Staffanstorp (north), an accumulation of stroke incidents can be observed. This underlines the plausibility of the generated stroke incidents and shows that the distribution of the occurrences corresponds to the actual residences of the population.

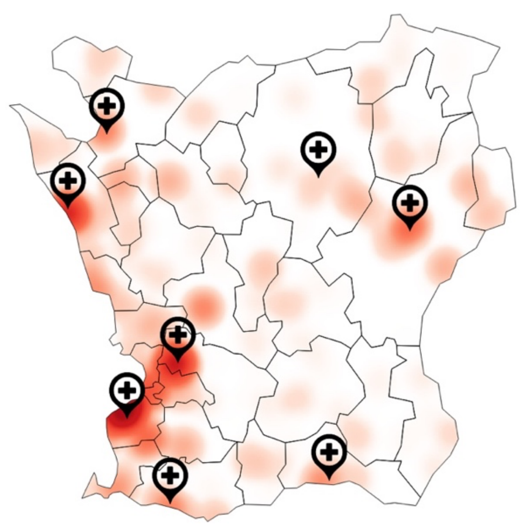


Figure 4. Heatmap of where strokes occur and locations of the hospitals that provide treatment for stroke patients.

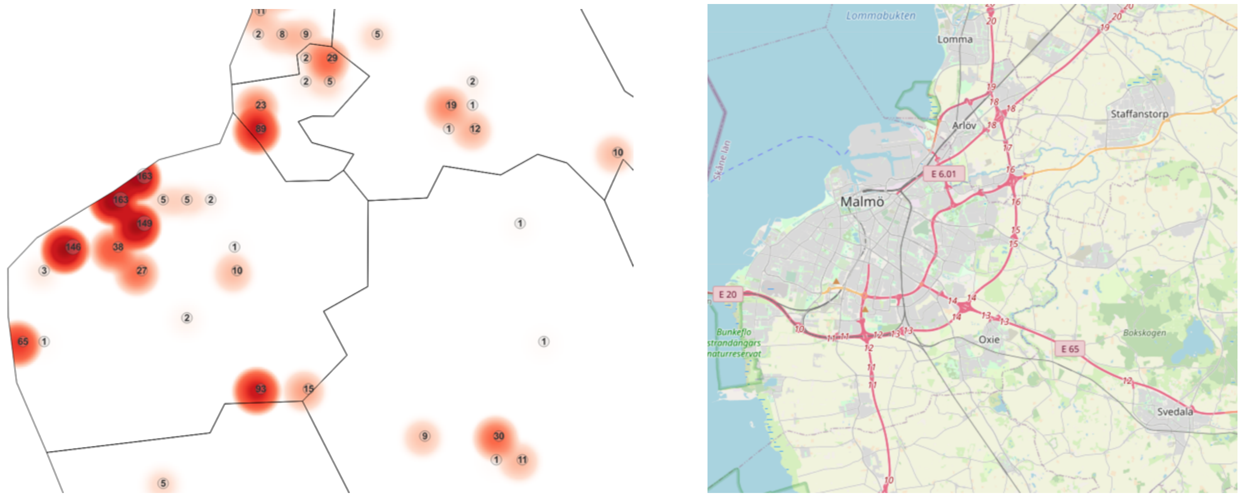


Figure 5. Heatmap of where strokes occur in and around Malmö municipality (left) and OpenStreetMap of the municipalities (right).

VI. Conclusions

In this article, we propose a process for the generation of synthetic populations, which can be used for investigating the effects of stroke logistics policies. The synthetic population corresponds to the original population in terms of the distribution of their relevant attributes that are required for analyzing a particular health condition, in this case, of stroke patients. This includes the age of the individuals, their place of residence (municipality) but also the imitation of their travel behavior as well as of the occurrence of strokes. Such synthetic populations enable the use of ABS, a simulation paradigm that uses Artificial Intelligence for modelling human behavior, which is a crucial factor when investigating the effects of policies.

In particular, we demonstrate the generation of a realistic population of stroke patients, which also takes travel behavior into account. This allows for the assessment of different stroke logistics policies, such as the optimal placement of MSUs across a region. For this purpose, we used aggregated and individual-based data from different sources, from which we derived probability distributions that were then used to generate an artificial population of agents.

To demonstrate the feasibility of the presented approach, we used data from the region of Skåne in southern Sweden. In the presented study, we simulated the travel behavior of stroke patients to investigate where strokes occur. Through this, a better understanding of the spatial distribution of stroke occurrence is achieved. This is relevant, for instance, for the optimal distribution of MSUs, such that the time to treatment is reduced for stroke patients.

Our results show that the generated artificial population corresponds to the real data in terms of the time of the day, at which strokes occur, the distribution of strokes across the municipalities, and the age group of the patients. In total, approximately 1.9% of the strokes occur while the individual is on a trip and not in their municipality of residence. This observation corresponds to data on strokes that was provided by the healthcare provider. Hereby, were able to show by means of simulation that traveling only has a minor impact on where strokes occur and, thus, for policy making in stroke logistics.

The generated artificial population of stroke patients is based on socio-demographic, healthcare, and travel data of the investigated region, to ensure the realistic representation of the real-world population. Yet, the presented process of population generation can also be applied to other regions and for other medical conditions, assuming that the required input data is accessible. This facilitates the conducting of agent-based simulation studies for investigating the effects different stroke logistics policies might have. It also increases the credibility of the simulation results such that conclusions can be drawn regarding the real world.

As part of future work, we plan to incorporate the results from the population generation into simulations for assessing and comparing different policies for stroke logistics. Moreover, we intend to investigate and include seasonality effects into the model, i.e., tourists that come to the region and changed travel behavior during weekends.

Acknowledgment

We would like to thank *Södra sjukvårdsregionen* for the provision of the anonymized data on the occurrence of strokes in Skåne. We would also like to thank the anonymous reviewers of DIGITAL 2021 for their fruitful and valuable comments that helped us to improve this work. This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program – Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation

REFERENCES

[1] A. Alassadi, F. Lorig, and J. Holmgren, "An Agent-based Model for Simulating Travel Patterns of Stroke Patients," in *DIGITAL 2021–Advances on Societal Digital Transformation, 14-18 November 2021, Athens, Greece*, 2021, pp. 11–16.

[2] eurostat, "Causes of death statistics." [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Causes_of_death_statistics [retrieved: May 2022]

[3] Y. Béjot, H. Bailly, J. Durier, and M. Giroud, "Epidemiology of stroke in Europe and trends for the 21st century," *La Presse Médicale*, vol. 45, no. 12, pp. e391–e398, 2016.

[4] H. A. Wafa, C. D. Wolfe, E. Emmett, G. A. Roth, C. O. Johnson, and Y. Wang, "Burden of stroke in Europe: thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years," *Stroke*, vol. 51, no. 8, pp. 2418–2427, 2020.

[5] E. Stevens, E. Emmett, Y. Wang, C. McKevitt, and C. Wolfe, "The burden of stroke in Europe: the challenge for policy makers," 2017, [Online]. Available: https://www.stroke.org.uk/sites/default/files/the_burden_of_stroke_in_europe_-_challenges_for_policy_makers.pdf [retrieved: May 2022]

[6] S. Barnes, B. Golden, and S. Price, "Applications of agent-based modeling and simulation to healthcare operations management," in *Handbook of healthcare operations management*, Springer, 2013, pp. 45–74.

[7] S. Almagooshi, "Simulation modelling in healthcare: Challenges and trends," *Procedia Manufacturing*, vol. 3, pp. 301–307, 2015.

[8] S. Amouzad Mahdiraji, O. Dahllöf, F. Hofwimmer, J. Holmgren, R.-C. Mihailescu, and J. Petersson, "Mobile stroke units for acute stroke care in the south of sweden," *Cogent Engineering*, vol. 8, no. 1, p. 1874084, 2021.

[9] P.-O. Siebers, C. M. Macal, J. Garnett, D. Buxton, and M. Pidd, "Discrete-event simulation is dead, long live agent-based simulation!," *Journal of Simulation*, vol. 4, no. 3, pp. 204–210, 2010.

[10] A. A. Tako and S. Robinson, "The application of discrete event simulation and system dynamics in the logistics and supply chain context," *Decision Support Systems*, vol. 52, no. 4, pp. 802–815, Mar. 2012.

[11] F. Lorig, E. Johansson, and P. Davidsson, "Agent-Based Social Simulation of the Covid-19 Pandemic: A Systematic Review," *JASSS*, vol. 24, no. 3, p. 5, 2021.

[12] E. Cabrera, M. Taboada, M. L. Iglesias, F. Epelde, and E. Luque, "Optimization of Healthcare Emergency Departments by Agent-Based Simulation," *Procedia Computer Science*, vol. 4, pp. 1880–1889, 2011.

[13] A. Djanatliev and R. German, "Prospective healthcare decision-making by combined system dynamics, discrete-event and agent-based simulation," in *2013 Winter Simulations Conference (WSC)*, 2013, pp. 270–281.

[14] T. Monks, M. Pitt, K. Stein, and M. James, "Maximizing the population benefit from thrombolysis in acute ischemic stroke: a modeling study of in-hospital delays," *Stroke*, vol. 43, no. 10, pp. 2706–2711, 2012.

[15] P. Chemweno, V. Thijs, L. Pintelon, and A. Van Horenbeek, "Discrete event simulation case study: Diagnostic path for stroke patients in a stroke unit," *Simulation Modelling Practice and Theory*, vol. 48, pp. 45–57, 2014.

[16] T. Ruppert, J. Bernard, A. Ulmer, H. Lücke-Tieke, and J. Kohlhammer, "Visual access to an agent-based simulation model to support political decision making," in *Proceedings of the 14th International Conference on Knowledge Technologies and Data-driven Business*, 2014, pp. 1–8.

[17] E. Silverman *et al.*, "Situating agent-based modelling in population health research," *Emerging Themes in Epidemiology*, vol. 18, no. 1, pp. 1–15, 2021.

[18] R. A. Grysiewicz, K. Thomas, and D. K. Pandey, "Epidemiology of ischemic and hemorrhagic stroke: incidence, prevalence, mortality, and risk factors," *Neurologic clinics*, vol. 26, no. 4, pp. 871–895, 2008.

[19] S. Walter *et al.*, "Bringing the hospital to the patient: first treatment of stroke patients at the emergency site," *PloS one*, vol. 5, no. 10, p. e13758, 2010.

[20] E. Berge *et al.*, "European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke," *European stroke journal*, vol. 6, no. 1, p. I–LXII, 2021.

[21] K. Fassbender *et al.*, "'Mobile stroke unit' for hyperacute stroke treatment," *Stroke*, vol. 34, no. 6, pp. e44–e44, 2003.

[22] V. J. Calderon, B. M. Kasturiarachi, E. Lin, V. Bansal, and O. O. Zaidat, "Review of the mobile stroke unit experience worldwide," *Interventional neurology*, vol. 7, no. 6, pp. 347–358, 2018.

[23] J. Kim *et al.*, "Economic evaluation of the Melbourne mobile stroke unit," *International Journal of Stroke*, vol. 16, no. 4, pp. 466–475, 2021.

[24] J. P. Rhudy Jr *et al.*, "Geospatial visualization of mobile stroke unit dispatches: a method to optimize service performance," *Interventional neurology*, vol. 7, no. 6, pp. 464–470, 2018.

[25] T. G. Phan, R. Beare, V. Srikanth, and H. Ma, "Googling location for operating base of mobile stroke unit in metropolitan Sydney," *Frontiers in neurology*, vol. 10, p. 810, 2019.

[26] J. Al Fatah, A. Alshaban, J. Holmgren, and J. Petersson, "An agent-based simulation model for assessment of prehospital triage policies concerning destination of stroke patients," *Procedia Computer Science*, vol. 141, pp. 405–412, 2018.

[27] K. Chapuis and P. Taillandier, "A brief review of synthetic population generation practices in agent-based social simulation," 2019.

[28] N. Fournier, E. Christofa, A. P. Akkinepally, and C. L. Azevedo, "An integration of population synthesis methods for agent-based microsimulation," in *Proceedings of the Annual Meeting of the Transportation Research Board, Washington, DC, USA*, 2018, pp. 13–17.

[29] N. Huynh, M.-R. Namazi-Rad, P. Perez, M. J. Berryman, Q. Chen, and J. Barthelemy, "Generating a synthetic population in support of agent-based modeling of transportation in Sydney," 2013.

[30] S. Dytckov, F. Lorig, P. Davidsson, J. Holmgren, and J. A. Persson, "Modelling Commuting Activities for the Simulation of Demand Responsive Transport in Rural Areas.," in *VEHITS*, 2020, pp. 89–97.

[31] B. Edmonds and S. Moss, "From KISS to KIDS–an 'anti-simplistic'modelling approach," in *International workshop on multi-agent systems and agent-based simulation*, 2004, pp. 130–144.

[32] O. E. Ramadan and V. P. Sisiopiku, "A critical review on population synthesis for activity-and agent-based

transportation models," *Transportation Systems Analysis and Assessment*, 2019.

[33] J. Barthelemy and E. Cornelis, "Synthetic populations: review of the different approaches," 2012.

[34] K. R. Bissett, J. Cadena, M. Khan, and C. J. Kuhlman, "Agent-based computational epidemiological modeling," *Journal of the Indian Institute of Science*, vol. 101, no. 3, pp. 303–327, 2021.

[35] Region Skåne, "Resvaneundersökning i Skåne," 2013. [Online]. Available: https://utveckling.skane.se/publikationer/rapporter-analyser-och-prognoser/resvaneundersokning-i-skane [retrieved: May 2022]

[36] P. W. Lewis and G. S. Shedler, "Simulation of nonhomogeneous Poisson processes by thinning," *Naval research logistics quarterly*, vol. 26, no. 3, pp. 403–413, 1979.

[37] V. Grimm *et al.*, "The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism," *Journal of Artificial Societies and Social Simulation*, vol. 23, no. 2, 2020.

[38] A. Alassadi, "An Individual-based Simulation Approach for Generating a Synthetic Stroke Population," 2021. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-41895 [retrieved: May 2022].

[39] Repast Simphony, Available: https://repast.github.io/ [retrieved: May 2022]