Proceedings of the 2022 Winter Simulation Conference
J. Galindo Jaramillo

USE OF SOCIAL DETERMINANTS OF HEALTH IN AGENT-BASED MODELS FOR EARLY DETECTION OF CERVICAL CANCER

Juan F. Galindo Jaramillo

UNICAMP – University of Campinas
Rua Paschoal Marmo, 1888
Jardim Nova Itália
Limeira, SP, 13484-332, BRAZIL

ABSTRACT

Cervical cancer is a treatable disease when detected at an early stage. Yet, in Brazil, most cases are detected at an advanced stage due to social conditions that impede periodical screening. This project aims to test strategies for maximizing the early detection of cervical cancer. An Agent-Based Model using social determinants of health is being created to simulate the conditions in which women live. Then, with Reinforcement Learning techniques, several strategies are tested. Early results show how improvements in schooling and income levels improve early-stage detection. These results may indicate the importance of improving social conditions for overall prevention.

1 INTRODUCTION

Cervical cancer is the fourth most frequent cancer type among women, responsible for more than 340,000 deaths in 2020, 90% of them in developing countries (WHO 2022). It is a treatable disease when detected at an early stage, yet around 80% of cases in Brazil are detected at an advanced stage (Vale et al. 2019). A systemic approach to cervical cancer must consider the social environment in which populations live. The Social Determinants of Health (SDoH) are social and economic elements that impact how people live (Almeida-Filho et al. 2003). As SDoH are helpful for the identification of the relationships between social conditions and diseases, they might be helpful as well for the creation of models to reduce the impact of diseases on a population and, thus, to improve active response in public health policy.

This work aims to use SDoH for cervical cancer to create an Agent-Based Model (ABM) of the evolution of cervical cancer among women to define strategies for the maximization of early-stage detection. The sources for the SDoH are two databases of the State of São Paulo, Brazil. A Reinforcement Learning algorithm (Q-Learning) is used to identify strategies.

2 MATERIALS AND METHODS

For the construction of the model, different stages are being performed. In the first stage, a Logit model was created fusing SDoH for ABM validation. In the second stage, the ABM was created. In the third stage, a Q-Learning algorithm will be implemented to maximize the rate of detection at an early stage.

In the first stage, two databases were used. The first dataset comes from the São Paulo Oncological Foundation (FOSP), containing 9502 records and six variables of cancer diagnosis and treatment data of hospitals in the state of Sao Paulo. The second dataset is published by the State of São Paulo Statistics portal (SEADE). It contains 645 records and 17 variables representing social, economic, and demographic data from the cities belonging to the state. The two datasets were combined. Then, data imputation was used to complete the missing variables (Jaramillo et al. 2021).

The Logit model was created having as a dependent variable the stage of the disease, with two values, corresponding to the early (stage I according to the International Federation of Oncology and Obstetrics) and advanced stage (stages II, III, and IV). The resulting multivariate Logit model corresponds to the combination of significative variables on univariate models. This model is made using R v4.2.

To build the ABM, Netlogo 6.2 is used. The ABM has two types of agents, women, and cities. Women develop cervical cancer and decide whether to go to health facilities represented within the city. Cities attend women according to their capacity. Cities' environmental conditions determine access to health facilities. In this first version, a population of 200 women was created into an 8x8 grid. For the model validation, the output of the ABM model is compared to the output of the Logit model. In the last stage, to test different policies, cities perform a Q-Learning algorithm to maximize the rate of cases detected early in their population. Python 3.8 is used for Q-Learning implementation.

3 EARLY RESULTS AND CURRENT STATE OF RESEARCH

The resulting Logit model had the following variables: women's age and schooling level; and cities' mean income, access to water and garbage collection, schooling and fecundity rates, and also cities' number of beds per 1000 inhabitants and the Social Responsibility Index - SRI (an index of the State of São Paulo that evaluates the level of well-being of the population). These variables are used in the ABM model.

In the ABM model, we add some variables. For women, the last time they were screened for cervical cancer is necessary to determine the stage at which cervical cancer can be developed. For cities, the number of women attended and the capacity of the hospital were added to the ABM model. When simulation begins, the space is divided into an 8x8 grid, making each grid square a city. A random SRI is assigned to each city, ranging from 1 to 5. Each city holds the capacity to attend a determined number of women using a uniform function with a maximum number equal to the city's SRI. Women are created and their age and location are defined randomly.

At each time step (one year), women's age and last screening are increased by one. According to the city's capacity, a number of women are screened. Depending on the number of years without screening, a woman's probability of getting cancer at an advanced stage is increased. Women decide to be screened depending on her scholarity and on the success rate on being attended during the last five years. At this moment, the ABM is being calibrated to match the Logit model. It currently shows a relationship between increasing income level and schooling with a higher screening rate. In the future, Q-Learning algorithms will be implemented to compare the effectiveness of policies in early rate detection.

ACKNOWLEDGEMENTS

This work was partially sponsored by Coordenação de Aperfeiçõamento de Pessoal de Nível Superior (CAPES), Brazil – under grant code 001.

REFERENCES

Almeida-Filho, N., I. Kawachi, A. P. Filho, and J. N. W. Dachs. 2003, dec. "Research on Health Inequalities in Latin America and the Caribbean: Bibliometric Analysis (1971–2000) and Descriptive Content Analysis (1971–1995)". American Journal of Public Health 93(12):2037–2043.

Jaramillo, J. F. G., G. M. Formigari, D. B. Vale, E. L. Ursini, and P. S. Martins. 2021. "Missing Data: Comparison of Multiple-Imputation Algorithms for Social Determinants of Health in Cervical Cancer Stage Detection". In 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 0509–0514.

Vale, D. B., C. Sauvaget, R. Muwonge, L. C. S. Thuler, P. Basu, L. C. Zeferino, and R. Sankaranarayanan. 2019, jan. "Level of human development is associated with cervical cancer stage at diagnosis". Journal of Obstetrics and Gynaecology 39(1):86–90.

World Health Organization.2022. "Cervical Cancer." World Health Organization. https://www.who.int/news-room/fact-sheets/detail/cervical-cancer, accessed 22 Feb. 2022.