Rosemary Kinuthia

Using Machine Learning to Predict and Explain Pre-Service Attrition of the Nursing Workforce in Kenya

Emory University Nell Hodgson Woodruff School of Nursing

Abstract

Nurses are the backbone of the healthcare system in Kenya, however, pre-service attrition is a major concern because not only does it add to the burden of the healthcare workforce shortage crisis, it also adds to the complexity of the challenges associated with efforts to improve population health outcomes without adequate resources. In Kenya, nurses provide the bulk of healthcare related services, therefore, pre-service attrition limits access to healthcare services. Attrition is also costly to healthcare systems. The purpose of this study is to use innovative machine learning methods to predict and explain pre-service attrition of the Kenyan nursing workforce. Data from the Kenya Regulatory Human Resources Information System (rHRIS) was used for this study. Supervised learning methods were applied to predict pre-service attrtion . K-nearest neigbour was used for regression. This study shows that using machine learning for prediction is useful for resolving workforce issues; it can be used to developing evidence-based interventions that are targeted to the factors linked to pre-service attrition.

Using Machine Learning to Predict and Explain Pre-Sevice Attrition of the Nursing Workforce in Kenya

**Background**

Human resources for health (HRH), also known as health workforce, is an essential pillar of health systems. An optimally functioning health system relies on health workers to alleviate the global burden of disease and meet the needs of communities by providing quality care. However, health systems are challenged by a chronic shortage of health workers. In 2013, it was estimated that there was a global needs-based shortage of over 17 million healthcare professionals including over 9 million nurses and midwives, 2.6 million physicians, and the remainder included other healthcare workers. The greatest shortage of health care professionals has been in regions that have the highest needs such as South East Asia and sub-Saharan Africa1.

The country of Kenya located in East Africa is not exempt from the workforce shortage crisis. According to the 2012 Kenya Nursing Workforce Report, the nurse to population ratio in the country is 51.5 nurses per 100,000 individuals in the nation’s population2. This ratio is lower than the critical threshold recommendation of 23 doctors, nurses and midwives per 10,000 population that was established by the World Health Organization’s (WHO) in order to attain the Millennium Development Goals (MDGs)3. The global burden of disease seems to be worsening as populations are living longer and sicker lives leading to an increasing demand for health resources 4. Thus, it is unlikely that the current workforce can provide and sustain adequate healthcare coverage needs for the population.

The Kenya health workforce largely comprises of nursing professionals who provide a majority of healthcare services to the population. However, due to the workforce shortage crisis, nursing professionals in Kenya are overburdened. There are efforts in place to work towards increasing the workforce. According to WHO, the nursing and midwifery workforce shortage will reduce from 9 million to 7.6 million, however, while it is predicted that the situation will improve globally, the nursing and midwifery shortage in Africa will actually worsen5.

One of the issues that may be contributing to the chronic nursing workforce shortage in Kenya is pre-service attrition of nurses i.e. students enroll and complete their nursing education training but do take the registration examination which grants a license for them to practice and enter the nursing workforce. A nursing workforce study conducted in Kenya identified that there was a 4.3% pre-service attrition between enrollment and registration in 1999, which later doubled to 8.2% in 2004 6. These findings are troubling and demonstrate a trend towards increasing pre-service attrition. The loss of nursing health workers before they enter the workforce exacerbates the chronic shortage and impedes optimal health.

While there have been studies conducted to assess general attrition of Kenyan health workers 7, 8 there is a gap in the literature when it comes to pre-service attrition because it is not well understood. Appiagyei et al identify pre-nursing as an issue and make a recommendation for further research to investigate factors contributing pre-service attrition in Kenya. Nurses are the backbone of the healthcare system in Kenya; therefore, to gain adequate health coverage in the country, it will be essential to obtain accurate estimates of pre-service attrition and gain insight on factors influencing pre-service attrition to inform scale-up on the nursing workforce in Kenya. The findings from this study will facilitate the development of evidence-based recruitment and retention strategies aimed at reducing chronic workforce shortages.

**Purpose of the study**

The purpose of this study is to use machine learning to predict and explain pre-service attrition of the nursing workforce in Kenya. Advances in machine learning and data science make it possible to predict attrition and also understand the key variables associated with turnover of nurses before they enter the workforce. Currently, Kenya country suffers from a chronic shortage of nurse providers, therefore, this data may be useful in the development of recruitment and retention strategies aimed at reducing the chronic workforce shortage, and also reduce healthcare costs resulting from nursing shortages.

**Research Question**

This study aims to address the following question:

* What are the main predictors of pre-service attrition of the nursing workforce in Kenya?

**Methods**

**Source of Data**

Data for this study were obtained from the Kenya Regulatory Human Resources Information System (rHRIS) which tracks training activities of health professionals for 8 regulatory bodies in Kenya, including: Nursing Council of Kenya, Kenya Medical Practitioners and Dentist Board, Kenya Medical Laboratory Technicians and Technologist Board, Clinical Officers Council, Pharmacy and Poisons Board, Kenya Nutritionist and Dieticians Institute, Radiation Protection Board and Public Health Officers and Technicians Council. For the scope of this study, analysis was limited to data on nurses ever trained and nurses ever registered until 02/27/2015.

**Variables**

The dependent variable for this study was attrition. The following independent variables from two rHRIS datasets (Nurses Ever Registered and Nurses Ever Trained) were analyzed for this study.

* Attrition, Age, Gender, County, FacilityAgent, Course

**Inclusion and Exclusion Criteria**

With the exception of duplicate occurrences, all data in the subset of the merged data were analyzed for this analysis.

**Statistical Analysis**

De-identified data from the rHRIS datasets were used to conduct a secondary data analysis to predict and explain pre-service attrition of the nursing workforce in Kenya. Data tidying of The Nurses Ever Registered and Nurses Ever Trained datasets was conducted by staff of Emory University Kenya Health Information System Program. For purposes of this study, the two datasets were merged to create one dataset with 105,327 nurses (rows) and 6 features (columns). A new column titled ‘attrition’ was added to identify whether the nurse registered with the Nursing Council of Kenya (NCK). Those nurses who did not register with NCK are identified with a “Yes” in the Attrition column. A subset of the original data containing the following variables (Attrition, Age, Gender, County, FacilityAgent, Course) was created for the analysis:

All statistical analyses were performed using R programming software. Logistic regression using glm function was used for automated machine learning to model pre-enrollment nursing attrition.

In preparation for modeling the data were changed to objects that the required package can interpret. The dataset was also split into training and test sets. A target was set (in this case “Attrition” and the features (independent variables) were used to model the prediction. Predictions were made on the test set and the model was evaluated using peformance assessment.

**Human Subjects Consideration**

The study protocol for the original rHRIS primary data collection was reviewed and approved by the Emory University Institutional Review Board. De-identified data and permission to use the data were obtained from the staff of Emory University Kenya Health Information System Program in Microsoft Excel format. Confidentiality of data was maintained at all times.

**Findings**

**Discussion**

**Limitations**

Innovative machine learning methods were used to make accurate predictions about pre-service nursing attrition, however, the data used for this study was not current. The most recent data on registration available was from 12/27/2015. There is a possibility that the state of nursing in Kenya may have evolved given that attrition is an area of high interest for the Kenyan Ministry of Health. However, given that not much has been published on this topic, this manuscript contributes to evidence in this area of research.

**Recommendations**

This study highlights the importance of addressing pre-service attrition in the nursing workforce. It will be necessary for nursing schools and the Kenyan government to develop policies to recruit qualified students to pursue a nursing career and provide support and engagement to ensure that they enter the workforce upon graduation. It will also be essential to develop mechanisms to sustain the nursing workforce after entry. It may be necessary to explore ways to incentivize nursing as a profession in Kenya address the workforces shortages in Kenya. For example, the government could explore options for subsidized tuition in the universities and pay differentials for nurses willing to work in areas with highest shortages.

**Implications for nursing practice and research**

Research in this area shows great potential and is important to the future of nursing practice, research, and knowledge. As one of the most trusted professionals in healthcare, nurses are privileged with the unique opportunity to engage with individuals and communities on an intimate level that other healthcare professionals do not. Nurses are also in the frontline and play a significant role in providing the bulk of healthcare services in under-resourced areas. Therefore, nurses should be more engaged in advocacy and influence policy development to advance the nursing profession. Because they are in the forefront and are most informed on factors that influence attrition in the profession, nurses should also be encouraged to engage in research that helps to alleviate the workforce burden.

**Conclusion**

In summary, there is room for significant contribution to improving workforce shortages in nursing and other healthcare professions. Use of machine learning to explain and predict is practical and should, therefore, be encouraged. This method provides accurate predictions which can be used to advance scale-up of training and education for nurses and prevent attrition. Addressing attrition to reduce workforce shortages has potential to improve health outcomes. Future studies should explore the significance of using evidence-based informed by machine learning techniques.

References

1. World Health Organization. Global strategy on human resources for health: Workforce 2030. WHO: <http://apps.who.int/iris/bitstream/10665/250368/1/9789241511131-eng.pdf?ua=1>; 2016.

2. Kenya Ministry of Health NCoK. Kenya Nursing Workforce Report: The Status of Nursing in Kenya, 2012. Kenya: 2012.

3. World Health Organization. Achieving the health-related MDGs. It takes a workforce! n.d. [cited 2018 April 15]. Available from: <http://www.who.int/hrh/workforce_mdgs/en/>.

4. GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388(10053):1459-544.

5. World Health Organization. Global Health Observatory (GHO) data n.d. [cited 2018 April 18]. Available from: <http://www.who.int/gho/health_workforce/nursing_midwifery_density/en/>.

6. Appiagyei AA, Kiriinya RN, Gross JM, Wambua DN, Oywer EO, Kamenju AK, et al. Informing the scale-up of Kenya's nursing workforce: a mixed methods study of factors affecting pre-service training capacity and production. Human resources for health. 2014;12:47.

7. Gross JM, Rogers MF, Teplinskiy I, Oywer E, Wambua D, Kamenju A, et al. The impact of out-migration on the nursing workforce in Kenya. Health Serv Res. 2011;46(4):1300-18.

8. Chankova S, Muchiri S, Kombe G. Health workforce attrition in the public sector in Kenya: a look at the reasons. Human resources for health. 2009;7:58.

**```**