LITERATURE REVIEW

MACS 30200, Dr. Evans

Sushmita V Gopalan

In 1990, India adopted one of the Millennium Development Goals to reduce infant mortality rate (IMR) by two-thirds by 2015. However, an IMR of 80 per 1,000 live births has only reduced to 35 per 1,000 live births in 2015, significantly higher than the target IMR of around 26 per 1,000 live births (UNICEF, 2015). Even though there existed, as of 2013, over 20 schemes in India that targeted a reduction in IMR (Press Information Bureau, 2013), decline in IMR has been consistently slowing down, suggesting that there is a need to go beyond disease-, programme- and sector-specific approaches (Claeson et al. 2000).

These schemes aim to achieve a variety of goals, having a range of different target populations. The Janani Suraksha Yojana, for instance, is aimed at promoting institutional deliveries. The Navjaat Shishu Suraksha Karyakram (NSSK) is a programme that trains healthcare providers in essential newborn care and resuscitation. One scheme offers pregnant mothers a Mother and Child Protection Card to help monitor healthcare service delivery through the Ministry of Women and Child Development. These schemes are not uniformly implemented across states. There appears to be no overarching logic to the collection of schemes nor have there been rigorous independent evaluations of the impact of each scheme.

I will first very briefly outline existing traditional literature on the determinants of infant mortality, specifically in the Indian context. I will then touch upon why it is important to build predictive models in order to narrowly target policy and then describe some ways in which that can be done.

DETERMINANTS OF INFANT MORTALITY IN INDIA

Income is generally considered to be the variable most correlated with infant mortality (Barbus, 2011, Hobcraft, 1985) . Maesham et al. (1999) trace out the role of income in changing infant mortality rates between 1975 and 1990. While income did have statistically significant influence on IMR, other factors such as technical progress and education levels had more substantial impact. Less than 25% of India's reduction in IMR in this period can be explained by income growth. When compared to other developing countries, they find that India's reduction in IMR over this time period is lower than would be predicted by the corresponding increase in income.

A World Bank report in 1999 finds that while the poorest Indian states have the worst infant mortality rates, the richest states do not have the best (Maesham et al. 1999). The states with the best indicators - Kerala and Tamil Nadu rank seventh and eleventh, respectively, in terms of per capita income. On the other hand, Delhi and Goa, which are the two states with highest per capita income do not even feature in the list of ten states with lowest infant mortality rates. Non-income factors such as maternal and child health interventions are found to play more significant roles in reducing infant mortality (Claeson et al. 2000).

Tamil Nadu forms an interesting case study to explore the impact of state-driven initiatives to improve maternal and child health services. Relative to the rest of India, infant mortality decreased rapidly in Tamil Nadu, from 80 in 1995 to 21 in 2007 (Padmanabhan et al. 2009). Concerted efforts to improve infrastructure such as standardizing a maternal death registration and audit, setting up certified obstetric and newborn-care centres, changes in incentive structure to attract medical officers to rural areas were found to be the primary driver of this reduction in IMR (Padmanabhan et al. 2009).

Consistent with this, Cleaeson et al. (2000) find that there is a significant positive relationship between lowered infant mortality rates and certain child health interventions like oral rehydration therapy, care seeking for acute respiratory infections, and immunization rates. Other important factors have been found to be nutrition status, age of the mother, employment status of the mother, whether or not the delivery was institutional, access to healthcare during pregnancy and time since previous birth (Saabneh, 2017)

PREDICTION MODELS

Lemon et al. (2003) outline two approaches traditionally employed to segment out part of a population that is at high-risk for a particular health condition. The first is to simply compute the likelihood of observing the health issue conditional upon belonging to a particular pre-defined subgroup of the population. While this is useful for descriptive purposes, it does not allow provide for a simultaneous consideration of several independent factors. The second is regression analysis, in this case, usually logistic regression because the outcome variable is dichotomous (Hosmer & Lemeshow, 2000). Regression analyses compute the *average* effect of an explanatory variable on our outcome of interest and hence, when policy is developed from these results, they are targeted at the average member of the population, without accounting for the fact that certain subgroups are disproportionately vulnerable to some health risks (Forthofer & Bryant, 2000). Even though we can explore the impact of interaction terms, interpretation becomes progressively more difficult as more variables are interacted together.

With growing evidence that the actual relationships between health outcomes and their explanatory variables are complex and nonlinear (Song et al. 2004), recent studies in epidemiology have begun use decision trees and other modern prediction methods for identifying high-risk groups vulnerable to bacterial infections among infants (Bachur & Harper, 2001), colon cancer (Camp & Slattery, 2002), coronary heart disease (Carmelli, et al. 2007), et cetera.

A decision tree is a prediction method, often used in classification and clustering; it is essentially a tree-structured graph which visually depicts a sequence of decisions and their consequences (Duman et al. 2009, Wu et al. 2009). Neural networks are a tool used in predictive analyses that are modeled on the way in which biological nervous systems process information (Liu et al. 2009, Rajan et al. 2009).

Tesfaye, et al. (2017) created a model to predict under-5 mortality using the Ethiopian demographic and health survey data. Breast-feeding, maternal education, family planning, preceding birth interval, occurrence of diarrhoea, father's education, birth weight and mother's age were found to be predictors of child mortality. They find that a pruned decision trees method has greater accuracy of prediction than a logistic regression approach or a decision tree without pruning, with an accuracy of 93.38% and area under ROC of 94.8%. This model was written into a web-based algorithm for use in areas without well-trained health professionals, where users can enter certain key measureable pieces of information and then the model classifies the child as being high-risk or low-risk.

Chen, et al. (2011) take a novel approach to building a predictive model for preterm births, one of the biggest causes of new-born deaths, using a combination of a neural network and a decision tree. They collected data on thousands of variables covering medical history, lifestyle factors, socio-economic variables for both parents and first used a neural network to identify the 15 most important factors that affect the likelihood of a preterm birth. Following this, Chen et al. used a decision tree to arrive at a set of rules for classification into high-risk and low-risk categories based on these fifteen variables. They find that multiple births, paternal drinking, and smoking, previous preterm births and low body weight for the mother are some of the best predictors of preterm births. They arrive at a set of ten different rules for classification, which are easy to interpret algorithmically, with precision ranging from 80% to 100%.

The gap in the literature where my research will fit in is building a predictive model to classify at-risk pregnant mothers and target policy efforts, in the Indian context. For this, I will be using a combination of data from the National Family Health Survey and the Sample Registration System in India. I intend to compare and contrast the predictive abilities of a logistic regression approach (Hosmer & Lemeshow 2000), a decision tree approach (Chen et al. 2011) and a single layer perceptron approach (Lemon et al. 2003).

REFERENCES

Bachur, Richard G., and Marvin B. Harper. "Predictive model for serious bacterial infections among infants younger than 3 months of age." *Pediatrics* 108, no. 2 (2001): 311-316.

Barbus, Alexandra. 2011. "DETERMINANTS OF INFANT MORTALITY." Europolis 5, no. 2: 154-178. Political Science Complete, EBSCOhost (accessed April 24, 2017).

Camp, Nicola J., and Martha L. Slattery. "Classification tree analysis: a statistical tool to investigate risk factor interactions with an example for colon cancer (United States)." *Cancer Causes and Control* 13, no. 9 (2002): 813-823.

Carmelli, Dorit, Heping Zhang, and Gary E. Swan. "Obesity and 33-year follow-up for coronary heart disease and cancer mortality." *Epidemiology* (2007): 378-383.

Chen, Hsiang-Yang, Chao-Hua Chuang, Yao-Jung Yang, and Tung-Pi Wu. "Exploring the risk factors of preterm birth using data mining." *Expert Systems with Applications* 38, no. 5 (2011): 5384-5387.

Claeson, Mariam, Eduard R. Bos, Tazim Mawji, and Indra Pathmanathan. "Reducing child mortality in India in the new millennium." *Bulletin of the World Health Organization* 78, no. 10 (2000): 1192-1199.

Duman, Fazil, Aykut Erdamar, Osman Erogul, Ziya Telatar, and Sinan Yetkin. "Efficient sleep spindle detection algorithm with decision tree." *Expert Systems with Applications* 36, no. 6 (2009): 9980-9985.

Forthofer, Melinda S., and Carol A. Bryant. "Using audience-segmentation techniques to tailor health behavior change strategies." *American Journal of Health Behavior* 24, no. 1 (2000): 36-43.

Hobcraft, John N., John W. McDonald, and Shea O. Rutstein. "Demographic determinants of infant and early child mortality: a comparative analysis." *Population studies* 39, no. 3 (1985): 363-385

Hosmer, David W., and Stanley Lemeshow. "Special topics." *Applied Logistic Regression, Second Edition* (2000): 260-351.

Liu, David, Yudie Yuan, and Shufang Liao. "Artificial neural networks for optimization of gold-bearing slime smelting." *Expert Systems with Applications* 36, no. 9 (2009): 11671-11674.

Measham, Anthony R., Krishna D. Rao, Dean T. Jamison, Jia Wang, and Alaka Singh. "Reducing Infant Mortality and Fertility, 1975-1990: Performance at All-India and State Levels." *Economic and Political Weekly* 34, no. 22 (1999): 1359-367.

Padmanaban, P., Parvathy Sankara Raman, and Dileep V. Mavalankar. "Innovations and challenges in reducing maternal mortality in Tamil Nadu, India." *Journal of Health, Population and Nutrition* (2009): 202-219.

Press Information Bureau. Government of India. Ministry of Health and Family Welfare. 2013. *Steps By Government To Reduce Infant Mortality Rate*. http://pib.nic.in/newsite/PrintRelease.aspx?relid=101932.

Rajan, K., Vennila Ramalingam, M. Ganesan, S. Palanivel, and B. Palaniappan. "Automatic classification of Tamil documents using vector space model and artificial neural network." *Expert Systems with Applications* 36, no. 8 (2009): 10914-10918.

Saabneh, Ameed. "The association between maternal employment and child survival in India, 1998–99 and 2005–06." *Asian Population Studies* 13, no. 1 (2017): 67-85.

Song, Xiaowei, Arnold Mitnitski, Jafna Cox, and Kenneth Rockwood. "Comparison of machine learning techniques with classical statistical models in predicting health outcomes." *Medinfo* 11, no. Pt 1 (2004): 736-40.

Tesfaye, Brook, Suleman Atique, Noah Elias, Legesse Dibaba, Syed-Abdul Shabbir, and Mihiretu Kebede. "Determinants and development of a web-based child mortality prediction model in resource-limited settings: A data mining approach." *Computer Methods and Programs in Biomedicine* 140 (2017): 45-51.

Wu, Li-Cheng, Jian-Xin Lee, Hsien-Da Huang, Baw-Juine Liu, and Jorng-Tzong Horng. "An expert system to predict protein thermostability using decision tree." *Expert Systems with Applications* 36, no. 5 (2009): 9007-9014.

United Nations Children's Fund 2015, *Fact Sheet: Infant and Child Mortality in India: Levels, Trends and Determinants*, viewed 20 April 2017, <http://unicef.in/CkEditor/ck_Uploaded_Images/img_1364.pdf>