The Challenge of Communicating Causes of Death and their Uncertainty

Jon Bird[1] and Edward Fottrell[2]

[1] Department of Computer Science, City University London, Northampton Square, London EC1V 0HB, UK
jon.bird@city.ac.uk
[2] Institute for Global Health, University College London, London WC1N 1EH, UK
e.fottrell@ucl.ac.uk

Abstract. Understanding how people die is vital for the evaluation and planning of public health interventions and yet currently the majority of the world's deaths are not recorded. One way that this issue can be addressed is by carrying out verbal autopsies (VAs): interviewing relatives of the deceased to find out the symptoms and circumstances that led to death. To facilitate this process we have developed MIVA, a mobile phone app that combines data capture and Bayesian analysis to determine the probable cause of death. The aim of this paper is to introduce verbal autopsy to the visualization community as it is a real-world application area where visualizations could help users reason under uncertainty. A challenge we face is how to effectively communicate causes of death and their levels of uncertainty to two very different user groups: health organisations, who may be familiar with statistical data, and the relatives of a deceased person, who may be illiterate.

Keywords: Cause of death, verbal autopsy, InterVA, uncertainty

1 Introduction

Information about mortality patterns is crucial for evaluating public health interventions and prioritising limited resources [8]. Verbal autopsy (VA) remains the best available approach for assessing causes of death in low-resource settings where most deaths occur at home [6]. WHO has developed standard questionnaires that ascertain 60 causes of death with reasonable accuracy from a VA interview [7]. These interviews are typically carried out by trained fieldworkers using paper-based questionnaires who ask relatives or other close caregivers of a deceased person questions about the symptoms of the terminal illness (Figure 1). During a verbal autopsy some questions relating to accidents and systemic symptoms are always asked (for example, for fever and intestinal symptoms); other questions are asked depending on the age and sex of the deceased (for example, pregnancy-related questions will only be asked if the deceased was a female of child rearing age). VA interview data are later reviewed by two or three doctors who reach a consensus about the probable cause of death in a

Fig. 1. A verbal autopsy (VA) being carried out using MIVA in South Africa. The two field workers are asking a relative of the deceased questions about the symptoms and circumstances of death and on the basis of the answers they record on the mobile phone the app automatically calculates a probable cause of death using a Bayesian approach.

process known as physician-coded verbal autopsy (PCVA). Although it might only take a physician a few minutes to identify the probable cause of death from a VA questionnaire, because they usually do the evaluations in their spare time, it can sometimes take several years from the interview to the completion of the PCVA. Recently, computer-based probabilistic methods have been developed for analysing VA data that have high levels of agreement with PCVA [2], as well as being faster and cheaper. One widely used method is InterVA (http://www.interva.net) that uses Bayes' theorem to calculate the probability of different causes of death given the symptoms recorded in a VA interview.

We have developed MIVA (Mobile InterVA), a mobile phone app that enables VA interviews to be recorded, analysed using the InterVA algorithm and then stored on the phone or transmitted to a server using SMS or wi-fi (Figure 2). Question skip patterns are built in to the software, a feature that fieldworkers appreciated in field testing and ensures that all the relevant questions are asked. Moving away from paper questionnaires also removes the possibility of transcription errors at the data processing stage. MIVA calculates the probable cause of death as soon as the interview is completed which creates new opportunities for fieldworkers to both inform respondents about how their relative died as well as provide them with appropriate health information. However, this raises complex ethical issues, with different stakeholders holding markedly different views about whether respondents should be informed about the cause of death, and if so when and by whom [1]. For example, we have found that although respondents want to know how their relatives died and that fieldworkers would be happy to inform them, doctors are concerned that the confidentiality rights of the deceased might be compromised.

This is a pressing ethical issue that we are currently investigating and in our field trials of MIVA we have not made the probable cause of death available to fieldworkers or respondents. However, we are currently working with medical ethicists and anticipate that under certain circumstances it will be appropriate to inform relatives about the probable cause of death. We will then need to effectively communicate the uncertainty associated with the probable cause of death, which could be particularly challenging given that sometimes the relatives of the deceased may be illiterate [3]. Another very different user group that can benefit from VA data is managers in health organisations. Although these individuals are generally well educated, their understanding of uncertainty may be limited [5] and they may not have the tools required to help them make decisions about resource allocation on the basis of mortality information. We believe that visualizing the uncertainty of causes of death is potentially beneficial for both of these very different user groups.

In the rest of this paper we give an describe the InterVA algorithm and the main sources of uncertainty in the causes of death calculated by this algorithm. Our aim is to introduce VA to the visualization community as it is a real world application area where there is decision-making and reasoning under uncertainty carried out by very different user groups. Furthermore, there is potential for effective uncertainty visualizations to make a positive contribution to public health in low resource settings.

2 InterVA

InterVA-4 uses a Bayesian approach to calculate the probability of 60 causes of death, based on the answers to 265 questions about the symptoms and circumstances of death contained in the World Health Organisation verbal autopsy standard [7]. In mathematical terms:

$$P(C|I) = \frac{P(I|C)P(C)}{P(I|C)P(C) + P(I|!C)P(!C)} \tag{1}$$

where $P(C)$ is the prior probability of a cause of death, $P(I|C)$ is the probability of a symptom (or other indicator) given a cause of death and $P(!C)$ is the probability of not C, or $(1 - C)$.

The algorithm assumes that the total conditional probability of the set of causes of death $C_1 ... C_m$ is unity and it repeatedly applies the same calculation for each indicator $I_1 ... I_n$ that applies to a particular cause of death:

$$P(C_i|I_{1...n}) = \frac{P(I_j|C_i)P(C_i|I_{0...n-1})}{\sum_{i=1}^{m} P(C_i|I_{0...n-1})} \tag{2}$$

Applying equation 2 results in the probability of most causes of death reducing while a few increase in probability. InterVA displays the most probable cause

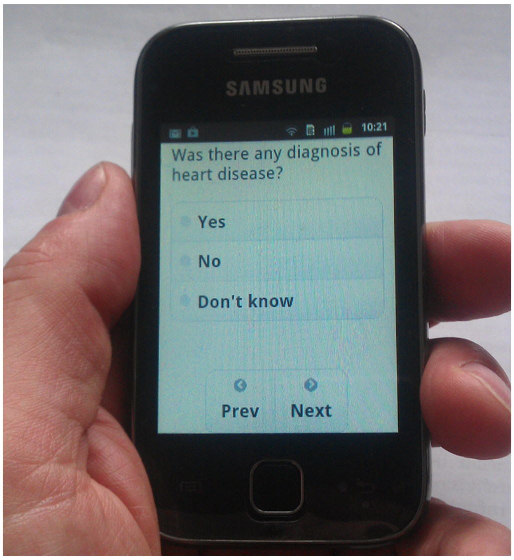

Fig. 2. MIVA (Mobile InterVA) is a verbal autopsy app that runs on mobile phones and combines data capture and analysis. It displays the WHO verbal autopsy standard questions [7], with automatic skipping, and when the interview is finished uses the InterVA algorithm to calculate the probable cause of death.

of death if it has a probability greater than 0.4, and up to two other causes if their probability is at least 50% of the probability of the highest cause.

The level of uncertainty associated with each cause of death is due to a number of factors, including the available evidence and the accuracy of the prior probabilities built into the system. There is no single source of data that associates the signs and symptoms that led to death with causes. The prior probabilities $P(C)$ and $P(I|C)$ are therefore determined by an expert panel of physicians who work in population health and are chosen for their knowledge about different geographical regions (e.g. Africa and Asia) and medicine specialisms (e.g. respiratory doctors and paediatricians). Based on their medical experience these physicians meet and come to a consensus about what probability to assign to each of the prior probabilities. They use letters ranging from 'A+', corresponding to 0.8, 'A' (0.4), 'A-' (0.2) to 'E' (0.00001), where each rank is approximately half the probability of the preceding higher one. The expert panel first determines the prior probabilities of the 60 causes of death ($P(C)$). Estimating $P(I|C)$ is a far more time consuming task and involves assigning a prior probability to each of the 15,900 cells in a 60 x 265 array ($P(I_1|C_1)...P(I_{265}|C_{60})$, where each column represents a possible cause of death ($C_1...C_{60}$) and each row is an indicator ($I_1...I_{265}$).

InterVA is designed to be used world wide and sensitivity analysis has demonstrated that the algorithm is robust to relatively large variations in the prior probabilities [4]. There are two causes of death, malaria and HIV, that have a high baseline prevalence in some settings and low prevalence in others and therefore local variations need to be taken into account. MIVA enables users to set the baseline of both of these causes to one of three levels: very low, low or high.

Another source of uncertainty is the InterVA model itself. First, it only includes signs and symptoms that are recognisable and reportable by relatives of the deceased. Second, a limited number of causes of death are included in the model and often these are broad categories. For example, the model does not distinguish between different diarrhoeal diseases as this can only be determined by knowing the bacteria underlying the disease. Third, the model only considers the presence of symptoms, that is, questions that were answered 'yes', and does not take in to account the absence of symptoms. Finally, the model is based on a simplifying assumption that the responses to each of the indicators are independent.

3 Summary

Data on mortality patterns are essential for health planning and evaluation. VA is currently the best tool available for recording deaths in resource limited settings. InterVA uses a Bayesian approach to analyse VA data and determine the probable cause of death. It is faster and cheaper than PCVA but there is uncertainty associated with the calculated cause(s) of death. This uncertainty can be seen as a strength of the model compared to physician review which generates a single 'certain' cause of death and where all uncertainty information is lost. A current challenge is to find ways to communicate the uncertainty associated with causes of death to the very different user groups who use VA data and to support their reasoning and decision making. We hope that the visualization community can help address this challenge and contribute to making a positive difference to public health.

References

1. Bird, J., Byass, P., Kahn, K. and Fottrell, E.: A Matter of Life and Death: Practical and Ethical Constraints in the Development of a Mobile Verbal Autopsy Tool. In 31st International Conference on Human Factors in Computing Systems, pp. 1489–1498 (2013)
2. Byass, P., Fottrell, E., Dao, L.H., Berhane, Y., Corrah, T., Kahn, K., Muhe, L., Van, D.D.: Refining a probabilistic model for interpreting verbal autopsy data. Scandinavian Journal of Public Health 34, pp. 26–31 (2006)
3. Medhi, I., Prasad, A. and Toyama, K.: Optimal Audio-Visual Representations for Illiterate Users. In 16th International Conference on World Wide Web, pp. 873–882 (2007)
4. Fottrell, E., Kahn, K., Tollman, S. and Byass, P.: Probabilistic Methods for Verbal Autopsy Interpretation: InterVA Robustness in Relation to Variations in A Priori Probabilities. PLoS One 6(11) (2011) e27200.

5. Logan, R.L. and Scott, P.J.: Uncertainty in Clinical Practice: Implications for Quality and Costs of Health Care. The Lancet 347, pp. 595–598 (1996)
6. Mathers, C. D., Ma Fat, D., Inouue, M., Rao, C. and Lopez, A. D.: Counting the Dead and What They Died From: An Assessment of the Global Status of Cause of Death Data. Bulletin of the World Health Organization 83, pp. 171–177 (2005)
7. The 2012 WHO Verbal Autopsy Standard (2012) http://www.who.int/healthinfo/statistics/verbalautopsystandards/en/
8. WHO Fact Sheet. Civil Registration: Why Counting Births and Deaths is Important (2007) http://www.who.int/mediacentre/factsheets/fs324/en/index.htm