Robust Identification of Determinants of Physical Activity Behaviour Change

Arjen Hommersom[1,2] **Simone C.M.W. Tummers**[1] **Catherine A.W. Bolman**[1] **Roger Bemelmans**[3] **Lilian Lechner**[1]

[1]Open Universiteit, Heerlen, The Netherlands
[2]Radboud University, Nijmegen, The Netherlands
[3]Zuyd University of Applied Sciences, Heerlen, The Netherlands

1 EXTENDED ABSTRACT

Physical Activity (PA) is associated with important health benefits, for example prevention of chronic illnesses such as cancer, cardiovascular disease, and diabetes (see e.g. Bullard et al. [2019]). The World Health Organization therefore recommends adults to engage in PA of moderate intensity for at least 150 minutes every week, spread over several days WHO [2010]. About 40% of European adults do not achieve recommended levels of PA, so there is much room for improvement Marques et al. [2015]. In specific higher-risk subgroups, e.g., cancers survivors, only 29 to 47% meet the guidelines Blanchard et al. [2008].

Interventions have been designed to stimulate PA of different target groups by influencing relevant psycho-social determinants based on theoretical behavioural explanation frameworks. In the Active4Life project, we integrated data from five different RCT intervention studies, and used this data to learn Bayesian network models to obtain insight into the most important determinants and how they interact. The dataset contains data from 5975 participants of which 4405 received an intervention. The dataset contains 51 variables which includes 11 determinants that were measured at 4 time points (at baseline, after 3 months, 6 months, and 12 months), several demographic variables, the intervention variable, and variables which represent the amount of PA at several time points.

A search-and-score based structure learning algorithm was used to learn a linear Gaussian Bayesian network from data. This search process was modified to be subject to time constraints, i.e., an arc $X_t \rightarrow Y_{t'}$ is not allowed if $t' < t$ and several constraints to ensure that demographic variables and the intervention variable do not contain determinants as parents. To handle the missing data, the structural expectation maximization (SEM) algorithm was used Friedman [1998]. This algorithm iteratively combines structure learning with the estimation of missing values based on observed data and the model learnt in a certain iteration. In previous work, we have shown that SEM outperforms various imputations methods in terms of goodness-of-fit Tummers et al. [2020] on this dataset. However, SEM is commonly known to be quite unstable due to its local optima. From the application point-of-view, this is problematic, because the structure is most relevant for learning about the main paths from intervention to behaviour change.

For improving the reliability of identifying arcs using structure learning algorithms, Efron's Bootstrap has been proposed Friedman et al. [1999], where an arc is included in the graph if it is contained in the learned graph in $k\%$ of the bootstrap samples, where $k\%$ exceeds some fixed threshold. While it is has been shown that this is an effective approach when dealing with synthetic data generated from a Bayesian network, it is not clear if this is a good approach when dealing with complex and noisy real-world data. To obtain some insight into the usefulness of this approach in case of missing data, we ran experiments with the ALARM network (containing 37 variables), from which datasets were sampled with missing data.

To illustrate, we chose a dataset with 1000 cases, where we ensured that 30% of the data are missing not at random (MNAR). Missings were generated such that higher values of the variables have a higher chance of being missing Xia et al. [2017]. While we do not know whether missings are MNAR in the PA application, it could be expected that motivational characteristics of the participants are correlated to filling out questionnaires. We evaluated the performance of learning the correct structure from this dataset (repeated 10 times). Without bootstrapping, the results are worse (mean tp: 20.5, fp: 19, fn: 25.5) compared to fully observed data (mean tp: 27.9, fp: 16.4, fn: 18.1). The high number of false positives are most undesirable for this application, because these may lead to a conclusion that some determinants are relevant for behaviour change, whereas in reality they are not. The bootstrapping results are shown in Figure 1. We observe that with a threshold of 0.65, we find much lower false positives with comparable number of tp and fn (mean tp: 20.1, fp: 12.8, fn: 25.9) to the network learned without bootstrapping. Additionally, with a higher threshold (0.95),

Accepted for the 16th Bayesian Modelling Applications Workshop (@UAI2022) (BMAW 2022).

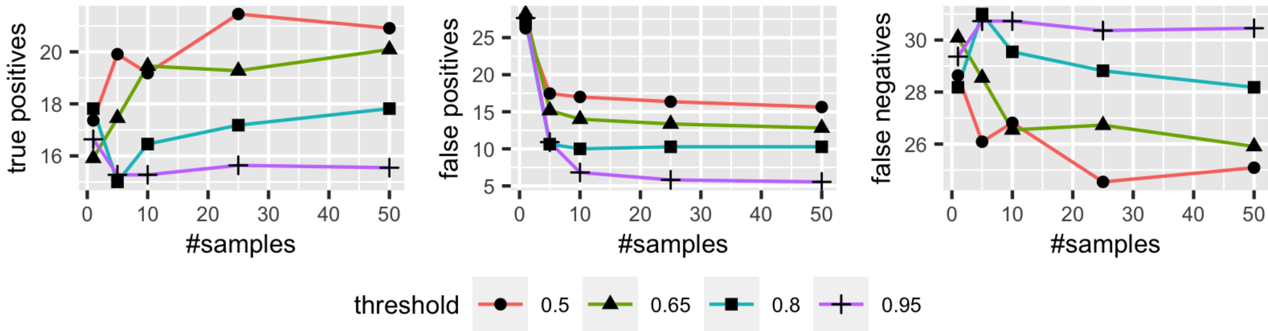

Figure 1: The effect on bootstrapping for learning the structure of the ALARM network from data with MAR missing data. The x-axis shows the number of bootstrap samples, and the y-axis is the mean tp, fp, and fn, respectively.

the false positive rate is very small.

A practical issue in real-world applications is that we do not known how many bootstraps samples should be generated. We propose the following idea to measure robustness: given a number of bootstrap samples n and a fixed threshold, one first learns two structures G_n and H_n each from n (independent) bootstrap samples. In case G_n and H_n provide robust estimates, one expects that $G_n \approx H_n$, which we can measure by the structure Hamming distance (SHD) between G_n and H_n. Then, to select n, we increase the number of n until the SHD converges. On the PA data, we observed that it usually converged between 100-150 bootstrap samples in the whole dataset and subsets of the data.

The Bayesian network model that we found show the complex structure in which PA interventions influence PA through pathways of determinants. The study reveals the significance of determinants for which previous research has only found limited evidence, for example self-efficacy and social influence concepts. Therefore, this research may provide new insights into the mechanism of PA behaviour change. However, in order to obtain robust estimates of these models, considerable effort should be taken to ensure that the end-result is stable and reproducible. In future research, we would like to investigate further ways to obtain robust estimates of the structure in the presence of different types of missing data mechanisms.

Acknowledgements

This work is part of the research programme Active4Life with project number 546003005, which is financed by ZonMw.

References

Christopher M Blanchard, Kerry S Courneya, and Kevin Stein. Cancer survivors' adherence to lifestyle behavior recommendations and associations with health-related quality of life: results from the american cancer society's scs-ii. *Journal of Clinical Oncology*, 26(13):2198–2204, 2008.

Tiffany Bullard, Mengmeng Ji, Ruopeng An, Linda Trinh, Michael Mackenzie, and Sean P Mullen. A systematic review and meta-analysis of adherence to physical activity interventions among three chronic conditions: cancer, cardiovascular disease, and diabetes. *BMC public health*, 19(1):1–11, 2019.

Nir Friedman. The Bayesian structural EM algorithm. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, pages 129–138, 1998.

Nir Friedman, Moises Goldszmidt, and Abraham Wyner. Data analysis with Bayesian networks: a bootstrap approach. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pages 196–205, 1999.

Adilson Marques, Hugo Sarmento, João Martins, and Luís Saboga Nunes. Prevalence of physical activity in european adults—compliance with the world health organization's physical activity guidelines. *Preventive medicine*, 81:333–338, 2015.

World Health Organization et al. *Global recommendations on physical activity for health*. World Health Organization, 2010.

Simone CMW Tummers, Arjen Hommersom, Lilian Lechner, Catherine Bolman, and Roger Bemelmans. Gaining insight into determinants of physical activity using Bayesian network learning. In *Benelux Conference on Artificial Intelligence*, pages 172–187. Springer, 2020.

Jing Xia, Shengyu Zhang, Guolong Cai, Li Li, Qing Pan, Jing Yan, and Gangmin Ning. Adjusted weight voting algorithm for random forests in handling missing values. *Pattern Recognition*, 69:52–60, 2017.