Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)
Demonstrations Track

Decision Platform for Pattern Discovery and Causal Effect Estimation in Contraceptive Discontinuation

Celia Cintas[1] , **Ramya Raghavendra**[2] , **Victor Akinwande**[1] , **Aisha Walcott-Bryant**[1] , **Charity Wayua**[1] and **Komminist Weldemariam**[1]
[1]IBM Research Africa, Nairobi, Kenya
[2] IBM T. J. Watson Research Center, Yorktown, NY, USA
celia.cintas@ibm.com, rraghav@us.ibm.com, victor.akinwande1@ibm.com, {awalcott, charitywayua, k.weldemariam}@ke.ibm.com

Abstract

Contraceptive use improves the health of women and children in several ways, yet data shows high rates of discontinuation which is not well understood. We introduce an AI-based decision platform capable of analyzing event data to identify patterns of contraceptive uptake that are unique to a subpopulation of interest. These discriminatory patterns provide valuable, interpretable insights to policymakers. The sequences then serve as a hypothesis for downstream causal analysis to estimate the effect of specific variables on discontinuation outcomes. Our platform presents a way to visualize, stratify, compare, and perform a causal analysis on covariates that determine contraceptive uptake behavior, and yet is general enough to be extended to a variety of applications.

1 Study of Contraceptive Discontinuation

Family Planning (FP) has emerged as a crucial component of sustainable global development [Osotimehin, 2015]. Effective use of contraceptives can significantly improve the nutritional status of both mother and children, reduce maternal mortality risks, and improve child survival [Ahmed *et al.*, 2019; Larsson and Stanfors, 2014; Azuike *et al.*, 2017; Johnson, 2017; Cahill *et al.*, 2018; Dey, 2019; Michael and Scent, 2017].

Surprisingly, complex pattern behaviours of contraceptive use and discontinuation[1] remain less explored. This work was done in partnership with the FP group at Bill & Melinda Gates Foundation (BMGF)[2]. We explore three questions specifically: (Q1) What do women transition to when they discontinue or switch between contraceptive methods? (Q2) Are there any recurrent sequences[3] of contraceptive use and discontinuation across countries? (Q3) Can we go beyond covariate analysis to establish a causal effect of particular contraceptives for a specific discontinuation reason?

Subject matter experts need the ability to look at large amounts of survey data (spanning multiple years, tens of thousands of subjects, across several countries), and be able to answer the above questions. Further, they need the ability to use this data to guide the design of policies and interventions e.g.: why do women discontinue injectables at a higher rate compared to other methods, and what if we designed better educational policies; a problem well suited for Machine Learning (ML) methods.

In contrast to previous works that look at traditional statistical methods [Michael and Scent, 2017; Johnson, 2017; Cahill *et al.*, 2018; Dey, 2019], we use ML methods to provide novel, complex insights into contraceptive discontinuation over both survey and event data. Our motivation for this work is to analyze very large data sets without predetermined theories or using only domain-established covariates, and see if new and/or unexpected patterns emerged.

2 The Decision Platform for FP

The Decision Platform for FP was designed to explore contraceptive use in countries with available DHS Surveys and Calendar data information. The platform provides the following facilities:

- **Automatically characterize subgroups of interest using patterns** that are unique to a subgroup (i.e. discriminatory), providing the ability to automatically generate hypotheses for various subgroups.
- **Estimate the causal effect** of using a particular contraceptive method on discontinuation across several countries. This technique enables us to systematically verify the hypothesis generated in the earlier step.
- **Integrated dashboard with visualization techniques** to make these ML insights accessible to a domain expert to explore discontinuation and contraceptive use.

Figure 1 depicts our prototype system and a brief description of each module is detailed below.

2.1 Pre-loaded Datasets in the Decision Platform for FP

In our particular case, we used the DHS data in consultation with our partner, team of Family Planning experts within the BMGF. The DHS calendar dataset [Croft *et al.*, 2018] is a month by month history of certain key events in the life of the respondent for the calendar period preceding the date of

[1]A discontinuation episode refers to an event that occurs in a specific month when a contraceptive method is not used for any reason.
[2]https://www.gatesfoundation.org/
[3]A sequence is an ordered list of items, in our case, a series of monthly ordered episodes from calendar data.

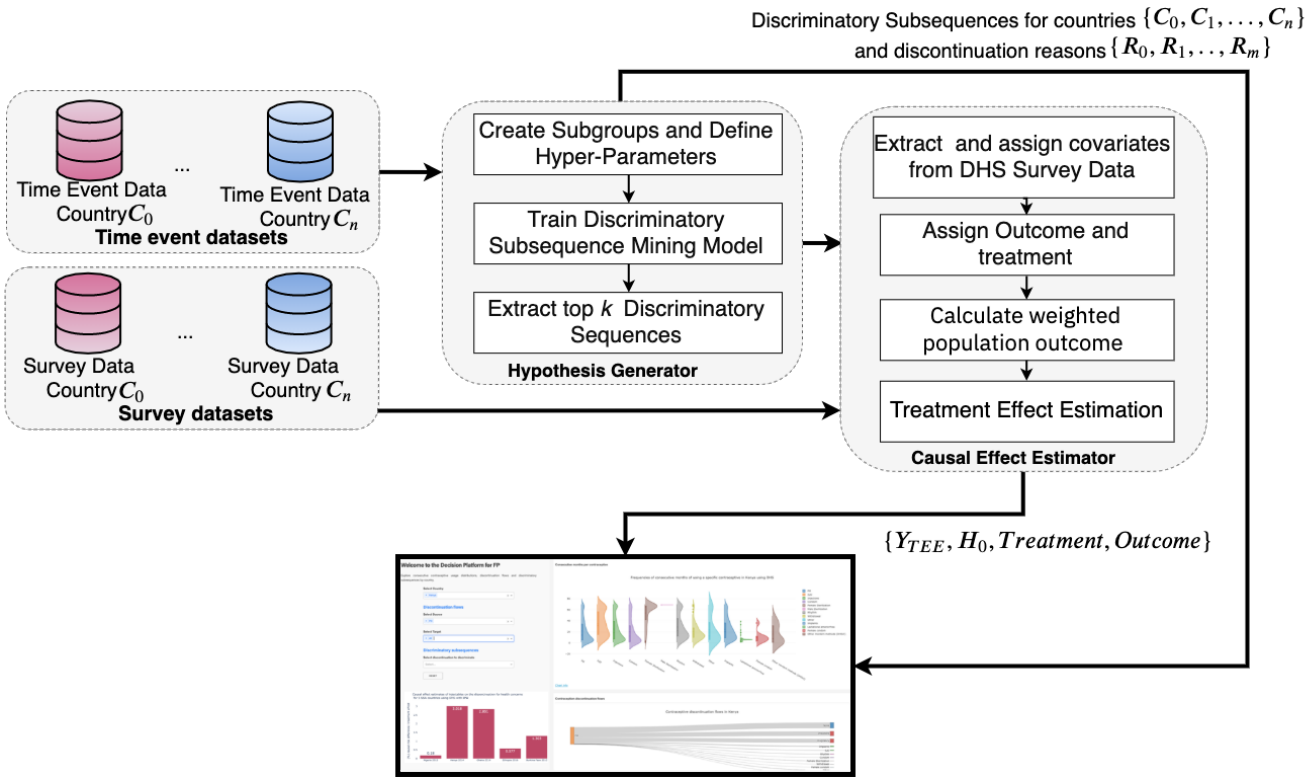

Figure 1. Method Overview. The Hypothesis Generator takes time event data from DHS Calendar Datasets across several countries to extract the most discriminative sub-sequences for a particular discontinuation reason. These sequences define the null hypothesis, treatment, and outcome for our Causal Effect Estimator. The dashboard is updated with the calculated treatment effect for a particular method under a specific outcome and all the subsequence patterns found by the Hypothesis Generator.

interview. In each month, DHS provides the type of contraception being used, the events that occurred during that month such as pregnancy, birth, termination, and the reason for discontinuation, if it occurs.

2.2 Hypothesis Generator

This module performs pattern mining on calendar data to check for common patterns in subjects that discontinue their contraceptive uptake. These patterns serve as hypothesis for downstream causal analysis to determine if and to what extent there exists causal relationship to discontinuation.

To analyze the sequential calendar data, we leverage sequence mining techniques, namely PrefixSpan [Jian Pei *et al.*, 2004]. However, given the size of the data, the unique journey of women in their fertility and contraceptive choices, out-of-the-box PrefixSpan mines very large numbers of frequent sequences that are often not unique to the subgroup of interest and hence largely unusable for manual inspection by an analyst. Thus, we extend the PrefixSpan algorithm to analyze large-scale and noisy data. We then compare two subgroups and mine patterns that would only appear predominantly in one of the subgroups. We collectively call these extensions the Discriminatory Subsequence Mining (DSM) technique. After the DSM model is trained, we extract the n discriminatory subsequences for each country C_j, we keep the top k more discriminatory sequences, based on *lift* metrics. An example of a discriminatory sequence found in DHS Nigeria 2013 was $[Injectables \rightarrow Injectables \rightarrow Non-use \rightarrow Non-use]$ under the sub-population that discontinue due to *health concerns*, this pattern appears in 34% of that subgroup and is 663 times more likely to appear in it than the rest of the population.

2.3 Causal Effect Estimator

This module uses both information from the subsequences found by the *Hypothesis Generator* and DHS Survey Data. The H_0, treatment and outcome (R_i) are extracted from the discriminatory subsequence. We conduct a series of steps from the observational data obtained from each individual woman's records of the DHS survey data. We want to evaluate if use of a particular contraceptive method had a causal effect on discontinuation for health concern reasons. The outcome is defined as whether or not a woman surveyed in the DHS reported "discontinuation for reason R_i". Third, we create the treatment assignment based on the use of that contraceptive in the past year. With that problem setup, we use *Inverse Probability Weighting* (IPW) [Rosenbaum and Rubin, 1983] with stabilized weights to calculate weighted population outcome for each subgroup stratified by treatment assignment. Finally, an example of a finding generated by the FP platform is that the use of injectables was 5 times more likely to cause discontinuation for health concerns for women surveyed in Kenya in 2014 compared to women surveyed in Ethiopia in 2016.

2.4 Decision Dashboard

The decision dashboard is updated with information from both previous modules. For example, we use interactive violin plots to display consecutive months per contraceptive type and country (multiple selections for comparisons are available). Furthermore, to easily interact with contraceptive switching patterns we implemented Sankey plots were the source and target can be set by the domain expert. For sequence mining insights, we use a combination of directed graphs and bar plots (for support metrics). For causal analysis, we display average treatment effect and several evaluation metrics of the Estimator, such as violation of positivity, consistency, balanced covariates and accuracy of our propensity model.

3 Concluding Remarks

The Decision Platform for FP provides automatic sequence mining and causal analysis for DHS Survey and Contraceptive Calendar Data. We demonstrate in this work the potential to provide decision support systems for policymakers regarding contraceptive use and discontinuation. At its core, our work is about analyzing sequential event data, extracting patterns that might help identify associations between behavioral patterns and discontinuation, and identifying if there is causal link to such associations.

Acknowledgements

This work is funded by Bill & Melinda Gates Foundation, investment ID 52720.

References

[Ahmed *et al.*, 2019] Saifuddin Ahmed, Yoonjoung Choi, Jose G. Rimon, Souleymane Alzouma, Peter Gichangi, Georges Guiella, Patrick Kayembe, Simon P. Kibira, Fredrick Makumbi, et. al, Funmilola OlaOlorun, Elizabeth Omoluabi, Easmon Otupiri, Sani Oumarou, Assefa Seme, Solomon Shiferaw, Philip Anglewicz, Scott Radloff, and Amy Tsui. Trends in contraceptive prevalence rates in sub-Saharan Africa since the 2012 London Summit on Family Planning: results from repeated cross-sectional surveys. *The Lancet Global Health*, 2019.

[Azuike *et al.*, 2017] EC Azuike, LC Ikeako, I Ezeobi, IU Ezebialu, JC Umeobika, KN Obi, JO Anene, and ED Azuike. Predictors of discontinuation of contraceptive use among nigerian women: Results of 2013 nigeria demographic and health surveys. *J Sci Res Stud*, 4(7):171–176, 2017.

[Cahill *et al.*, 2018] Niamh Cahill, Emily Sonneveldt, John Stover, Michelle Weinberger, Jessica Williamson, Chuchu Wei, Win Brown, and Leontine Alkema. Modern contraceptive use, unmet need, and demand satisfied among women of reproductive age who are married or in a union in the focus countries of the family planning 2020 initiative: a systematic analysis using the family planning estimation tool. *The Lancet*, 391(10123):870–882, 2018.

[Croft *et al.*, 2018] Trevor Croft, Sarah Bradley, and Courtney Allen. The dhs program - calendar tutorial, 2018. https://www.dhsprogram.com/data/calendar-tutorial/, last accessed on 06/02/20.

[Dey, 2019] Ajit Kumar Dey. Socio-demographic determinants and modern family planning usage pattern-an analysis of national family health survey-iv data. *International Journal Of Community Medicine And Public Health*, 6(2):738–749, 2019.

[Jian Pei *et al.*, 2004] Jian Pei, Jiawei Han, B. Mortazavi-Asl, Jianyong Wang, H. Pinto, Qiming Chen, U. Dayal, and Mei-Chun Hsu. Mining sequential patterns by pattern-growth: the prefixspan approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(11):1424–1440, Nov 2004.

[Johnson, 2017] Ofonime E Johnson. Determinants of modern contraceptive uptake among nigerian women: Evidence from the national demographic and health survey. *African journal of reproductive health*, 21(3):89–95, 2017.

[Larsson and Stanfors, 2014] Cecilia Larsson and Maria Stanfors. Women's education, empowerment, and contraceptive use in sub-saharan africa: findings from recent demographic and health surveys. *African Population Studies*, 28:1022–1034, 2014.

[Michael and Scent, 2017] Turnwait O Michael and Grace AT Scent. Correlates of contraceptive use and the desire for less children in nigeria. *The Nigerian Journal of Sociology and Anthropology*, 15(2):101–116, 2017.

[Osotimehin, 2015] Babatunde Osotimehin. Family planning as a critical component of sustainable global development. *Global Health Action*, 2015.

[Rosenbaum and Rubin, 1983] Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.