eRegQual analysis

Chris Rose, Norwegian Institute of Public Health (09:14:1512 May 2020)

Git revision: 8abf871

# Introduction

TODO: Write this

# Methods

Because outcome data were missing for about a third of participants (see results), we used Little's tests (Little 1988) of the null hypotheses that missing values of the constituent outcomes were jointly missing completely at random (MCAR) and covariate-dependent missing (CDM). We then used multiple imputation via chained equations (van Buuren 2007) to create and analyze 5 multiply-imputed datasets. Methodologists currently regard multiple imputation as a state-of-the-art technique that is expected to reduce bias and increase precision relative to other missing data techniques. We imputed each of the constituent outcomes using the auxiliary variables age, BMI, years of education, average monthly household income (transformed to the log scale due to the highly skewed distribution of income), and variables that indicated whether a laboratory or ultrasound were available at the clinics; the variables included in the analysis described below were also included. We were not able to include auxiliary variables that indicated previous pregnancy with pre-eclampsia or previous history of GDM due to collinearity. We evaluated the convergence of the imputation algorithm by inspecting trace plots and evaluated imputed data by inspecting kernel density and bar plots comparing the distributions of imputed and complete case data.

For each imputed data set, we computed the composite outcome from the imputed constituent outcome data. An adverse pregnancy outcome was defined to have occurred if at least one of the constituent outcomes occurred, and not to have occurred if none of the constituent outcomes occurred. Because we imputed values of the constituent outcomes, none of the composite outcomes could be missing in the imputed data. For each imputed data set, we estimated a risk ratio to compare treatment to control, adjusting for the stratification variable as a fixed effect, using generalized estimating equations (GEE; binomial errors and log link) to account for the cluster design. Estimates were then combined using Rubin's rules. For comparison, we also performed a complete case analysis under the MCAR assumption.

We followed the intention-to-treat principle: participants were analyzed in the arms to which they were randomized and — with the exception of the complete case analysis — all participants were included in the analyses. We computed 95% confidence intervals and used the significance criterion P<0.05 throughout. Statistical analyses were performed using Stata 16 (StataCorp LLC, College Station, Texas, USA).

# Results

Outcome data were missing for between 11.8% and 35.5% of the constituent outcomes, and 33.8% of the composite outcome. We were unable to reject the MCAR and CDM hypotheses (P=0.15 and P=0.64, respectively).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TODO: Main result | | | | | | |
| y | exp(b) | Std. Err. | t | P>|t| | [95% Conf. Interval] | |
| arm |  |  |  |  |  |  |
| D | 1.02 | 0.04 | 0.37 | 0.71 | 0.93 | 1.11 |
|  |  |  |  |  |  |  |
| strat\_var |  |  |  |  |  |  |
| 8d9c30 | 1.10 | 0.09 | 1.25 | 0.21 | 0.94 | 1.29 |
| 9d5ed6 | 0.94 | 0.09 | -0.64 | 0.52 | 0.77 | 1.14 |
| e1e1d3 | 1.09 | 0.09 | 1.07 | 0.28 | 0.93 | 1.27 |
| ff4457 | 1.15 | 0.09 | 1.83 | 0.07 | 0.99 | 1.34 |
|  |  |  |  |  |  |  |
| \_cons | 0.23 | 0.02 | -21.19 | 0.00 | 0.20 | 0.27 |

# Appendix

TODO: Write this

# References

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American statistical Association, 83(404), 1198-1202.

van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. Statistical methods in medical research, 16(3), 219-242.