P01-568

USING THE CES-D SCALE IN A LARGE COHORT STUDY AND DEALING WITH MISSING DATA: APPLICATION TO THE FRENCH E3N COHORT

N. Resseguier[1], H. Verdoux[2], F. Clavel-Chapelon[3], X. Paoletti[1]
[1]Biostatistics Department, Institut Curie, Paris, [2]Inserm U657, Bordeaux, [3]Inserm U1018, Villejuif, France

Introduction: The CES-D scale is commonly used to assess depressive symptoms (DS) in large population-based studies. Missing values in items of the scale may create biases.

Objectives: To explore reasons for not completing items of the CES-D scale and to perform sensitivity analysis of the prevalence of DS to assess the impact of different missing data hypotheses.

Methods: 71412 women included in the French E3N cohort returned in 2005 a questionnaire containing the CES-D scale. 45% presented at least one missing value in the scale. An interview study was carried out on a random sample of 204 participants to examine the different hypotheses for the missing value mechanism. The prevalence of DS was estimated according to different methods for handling missing values: complete cases analysis, single imputation, multiple imputation under MAR (missing at random) and MNAR (missing not at random) assumptions.

Results: The interviews showed that participants were not embarrassed to fill in questions about DS. Potential reasons of nonresponse were identified. MAR and MNAR hypotheses remained plausible and were explored.

Among complete responders, the prevalence of DS was 26.1%. After multiple imputation under MAR assumption, it was 28.6%, 29.8% and 31.7% among women presenting up to 4, to 10 and to 20 missing values, respectively. The estimates were robust after applying various scenarios of MNAR data for the sensitivity analysis.

Conclusions: The CES-D scale can easily be used to assess DS in large cohorts. Multiple imputation under MAR assumption allows to reliably handle missing values.