Appendix 3 Description of methods of statistical analysis

Statistical analyses were performed using the R software (version 3.2.3). Continuous variables are presented as means with standard deviations (SDs), or medians with interquartile ranges (IQRs), due to the skewness of the latter. Categorical data is described as numbers (frequencies) and percentages. The assumption of normal distribution was tested using Shapiro-Wilk test. Bivariate comparisons were made using t-test or Mann-Whitney test for the quantitative variables, whereas chi-square or Fisher’s exact tests were applied to qualitative data. Univariate and multivariate models were based on logistic regression. Results from logistic regressions were expressed as odds ratios (ORs) with associated 95% confidence intervals (CIs). Multivariate models were built using the following strategies: only significant variables with univariate analysis (model 1) or forward/backward selection (model 2). The goodness of fit in multiple regression was assessed using Bayesian information criterion (BIC). The 2-tailed tests were carried out at a significance level of p ≤ 0.05.