# Appendix: Development of non-reporting weights to address incomplete reporting to the National Healthcare Safety Network during the COVID-19 pandemic

For the first half of 2020, requirements for hospital HAI reporting to the National Healthcare Safety Network (NHSN) were suspended,4 although many hospitals continued reporting some HAI to NHSN via automated processes during this time period, we observed increases in incomplete reporting in California.5 In addition, there is evidence that reporting completeness did not return to pre-pandemic levels when reporting requirements were reinstated in July 2020.5 Increased incomplete reporting in 2020 has the potential to bias estimates of the effect of the COVID-19 pandemic on HAI incidence if non-reporting is associated with HAI risk or incidence. Therefore, we addressed missing data in our regression models by using non-response weights. Weights were created for each month in the reporting period via logistic regression models. Predictor variables for these logistic models were: number of licensed beds (<125 beds, 125-250 beds, and >250 beds); teaching vs non-teaching status; facility ownership (public, private not-for-profit, and for-profit/physician owned); the number of infection control (IC) staff; ratio of IC staff to licensed beds; and the number of full-time equivalent hospital epidemiologists. We also considered using the total hours of staff time spent on surveillance and total hours spent on other IC activities but did not include these in the model due to high collinearity with the number of IC staff. Predictor variables for the non-response weight model were obtained from the HAI reporting files and the NHSN Annual Facility Survey for 2020, 2019, and 2018 [cite annual survey.16 If variables from the annual facility survey were missing for a particular year, data from the nearest non-missing year was substituted if available, with 2018 data prioritized over 2020 data for 2019. The list of facilities with assigned NHSN numbers for each year was used as the universe of facilities. Non-response weights were only created for facility months where less than 98% of facilities had reported complete data for the month.