

SUPPLEMENTAL DIGITAL CONTENT 1. Additional information on statistical methods

In the calculation of incidence rates, to adjust for sample collection, we divided the crude incidence of IPD by the proportion of all admitted children meeting blood sampling criteria who had a blood culture performed. Adjustments were stratified by month and age group. Bootstrapping with biased-corrected confidence intervals was used to account for uncertainty in adjusted incidence rates.¹

In the calculation of time at risk for incidence rates, the beginning of the time at risk was defined for each child as the first day of the study period (January 1st, 2003), date of birth or date of immigration, whatever occurred later. The end of follow-up was defined for each child as the last day of the study period (December 31st, 2018), the date they turned less than 10 weeks of age or the date of death, whatever occurred first.

Negative binomial regression models were estimated to determine incidence rate ratios (IRR) between post- and pre-PCV periods. These models used random intercepts to account for repeated measures since children could change age or exposure during follow-up.

¹ Mackinnon DP, Dwyer JH. Estimating Mediated Effects in Prevention Studies. *Eval Rev.* 1993;17(2):144–58.