Supplemental Digital Content 1:

Statistical methods

The association between cytokine levels and incidence of LRI at age 0-3 yrs was analyzed by quasi Poisson regression. The associations were investigated in a hierarchy with three levels; a) A univariable model using each cytokine from each of the 3 bacterial stimulations (21 tests) estimating Incidence Risk Ratios (IRR) with 95% CIs of LRI, b) A trivariable model using each cytokine across all three bacterial stimulations (7 tests), and c) A multivariable model assessing the entire immune profile by combining the seven cytokine responses across all three bacterial stimulations (1 test).

a) The univariable model was calculated as simple quasi-poison regression:

$$No.\ of\ LRI = intercept + Cytokine_x(Bacteria_y) + covariates$$

b) The trivariable model was calculated by the inference between two linear nested models, likewise assuming quasi-Poison distribution):

(I) Large model$_{tri}$ (1 cytokine, 3 stimulations (HI, MC, SP):
$$No.\ of\ LRI = intercept + Cytokine_x(HI) + Cytokine_x(MC) + Cytokine_x(SP) + covariates$$

(II) Small model:
$$No.\ of\ LRI = intercept + covariates$$

Trivariate model:
$$Inference = ANOVA\ (Large\ model_{tri}(3\ variables)\ ;\ Small\ model)$$

c) The multivariate model was likewise based on the inference between two linear nested models:

(I) Large model$_{multi}$ (7 cytokines, 3 stimulations (HI, MC, SP):
$$No.\ of\ LRI = intercept + IFN\text{-}\gamma(HI) + IFN\text{-}\gamma(MC) + IFN\text{-}\gamma(SP) + \ldots + TNF\text{-}\alpha(HI) + TNF\text{-}\alpha(MC) + TNF\text{-}\alpha(SP) + covariates$$

(II) Small model:
$$No.\ of\ LRI = intercept + covariates$$

Multivariate model:
Inference = ANOVA (Large model$_{multi}$(21 variables) ; Small model)

Principal Component Analysis (PCA)

To explore findings from the multivariate model, we applied principal component analysis (PCA), decomposing the complex data set into fewer dimensions in order to capture the largest variability in the data. In the interpretation of the immune response pattern relevant for the association with LRI, we applied principal component analysis (PCA) on a block concatenated (bacteria wise) matrix of immune profiles, with a decomposition of a 873 (291 × 3) by 7 matrix (291 children included in the analysis and each child has 3 stimulations performed with 7 different cytokines measured).

In the PCA analyses, we extracted underlying orthogonal components that described the systematic part of the variation across the cytokines and the stimulations. The PCA decomposes the signal into consecutive sources of variation where principal component 1 (PC1) captures the dominating variation, PC2 the second dominating variation and so forth. Associating the principal components one-by-one with the number of LRI per child points at how dominating the relevant pattern is.