

ALLERGY

# FARMI

28.06.2019

I will add now a special collection of farming studies here as many of them are [just candidates for the Ig nobel prize](#).

[The most recent study introduces FaRMI](#), a “bacterial relative abundance farm home microbiota index”, probably introduced as the authors couldn’t find anything else. It reminds me very much to the polygenic risk score that rescues your study if you could not find the gene.

Asthma prevalence has increased in epidemic proportions with urbanization

Already the first sentence is wrong if we look at the following plot where asthma is clearly levelling off.

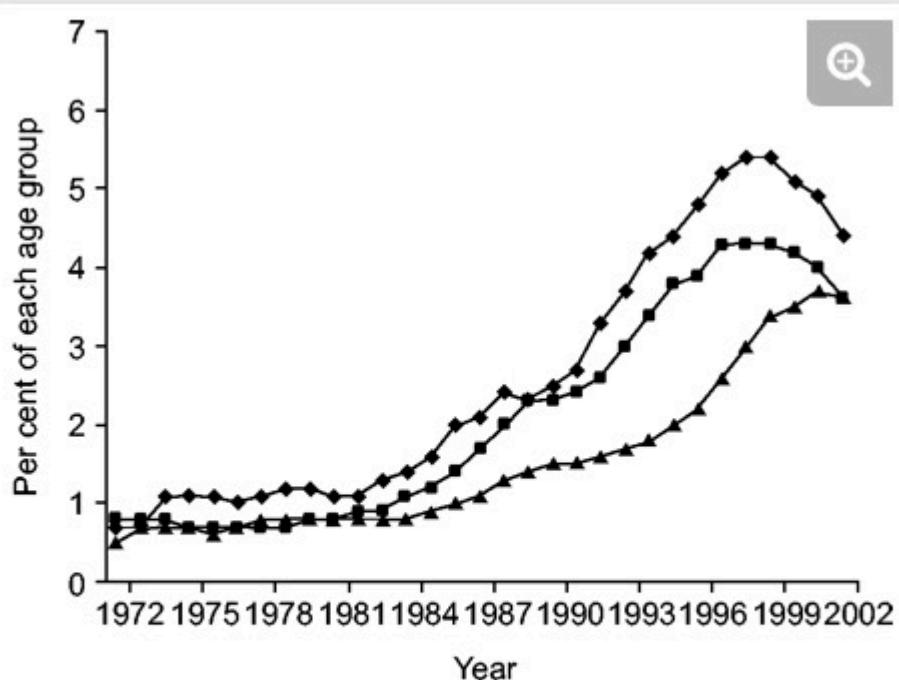


Fig. 1 –

[Download figure](#) | [Open in new tab](#) | [Download powerpoint](#)

Asthma prevalence during 1972–2002 (3-yr progressive mean) for subjects aged <45 yrs. ♦: aged 0–14 yrs; ■: aged 15–24 yrs; ▲: aged 25–44 yrs. Standardised for the population of the Netherlands in 1990 [17, 18](#).

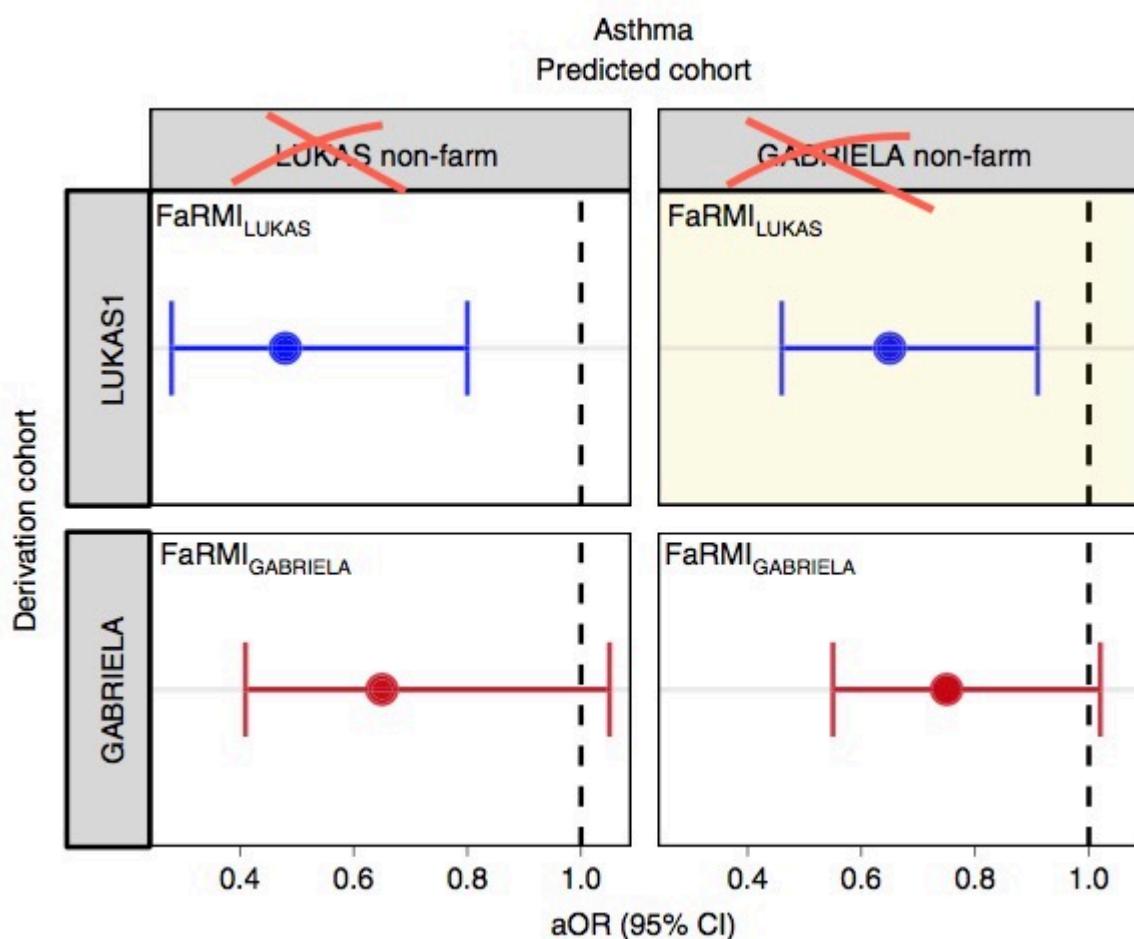
Urbanization happened in the late späten 19th century and not after 1950. Source of plot: The prevalence of asthma in children: a reversing trend ([ERJ 2005](#))

Unfortunately, the difference between farm and non farm children is never explained in the Kirjavainen et al. paper. What is the average distance of a non farm house to a farm house? [Are there any joint school or sports activities](#) of children from farms and non farms (allergens travel in the classroom)? And why is there such a strong conclusion in the title?

Farm-like indoor microbiota in non-farm homes protects children from asthma development

A lower risk score is not equivalent to protection.

And did any reviewer ever look at the plots or tables?



I do not understand Figure 3c. It even makes only sense when I cross out the top labels. But even then it means: GABRIELA does not show a significant replication.

**Table 1 | FaRMI was associated with reduced proinflammatory responsiveness**

Cytokine	Capacity to produce		Response to bacterial exposure			In vivo activation <sup>a</sup>
	PI		LPS		PPG	
	1 year <sup>b</sup>	6 years <sup>c</sup>	1 year <sup>d</sup>	6 years <sup>c</sup>	6 years <sup>e</sup>	
IL-1 $\beta$	7.7	22.7	-27.5 <sup>f</sup> (0.021)	-25.0 <sup>f</sup>	-33.8 <sup>f</sup> (0.056)	-
IL-4	19.7	-16.5	-	-	-	-
IL-5	9.7	11.4	-	-	-	-
IL-6	-15.9	10.5	-16.6 <sup>f</sup>	-25.5 <sup>f</sup> (0.073)	-30.7 <sup>f</sup> (0.076)	24
IL-10	-28.9 <sup>f</sup> (0.022)	-13.1	-12.1	1.6	-27.8 <sup>f</sup>	13.1
IL-12p70	-	-	5.5	-43.0 <sup>f</sup> (0.050)	-33.2 <sup>f</sup> (0.084)	-
IL-13	18.1 <sup>f</sup>	6.8	-	-	-	-
IL-17A	-3.8	-9.1 <sup>f</sup>	-	-	-	-
IFN- $\gamma$	-7.5	-0.1 <sup>f</sup>	-5.2	-46.2 <sup>f</sup> (0.081)	-44.8 <sup>f</sup> (0.017)	-
TNF- $\alpha$	-4.7	-16.1	-22.9 <sup>f</sup> (0.095)	-31.8 <sup>f</sup> (0.073)	-8.3 <sup>f</sup>	-
CRP	ND	ND	ND	ND	ND	-69.2 <sup>f</sup> (0.016)

The table shows the association between FaRMI and serum CRP levels and cytokine responses to ex vivo microbial stimulants in LUKAS non-farm children. Adjusted estimates of quantile regression analysis for cytokine/CRP concentrations at the 75th percentile per IQR change in FaRMI are shown. Estimates were adjusted for sex, maternal and paternal allergic disease, maternal smoking during pregnancy, number of older siblings, maternal education and cohort. The adjusted estimates are presented as a percentage of IQR of the respective cytokine/CRP concentration to allow comparison between the different analytes. P values (two-sided, uncorrected for multiple testing) are given in parentheses where suggestive association to FaRMI was observed at the 75th percentile ( $P < 0.1$ ). ND, not determined; <sup>f</sup>, data not analyzed because over 25% of observations below the detection limit; PI, phorbol 12-myristate 13-acetate and ionomycin; LPS, lipopolysaccharide; PPG, peptidoglycan; CRP, C-reactive protein. <sup>a</sup>Cytokines measured in non-stimulated cell culture media and CRP from serum. <sup>b</sup> $n = 231$  for IL-1 $\beta$  and IL-4;  $n = 230$  for IL-5, IL-10, IL-12p70, IL-13 and IL-17A;  $n = 208$  for IFN- $\gamma$  and tumor-necrosis factor (TNF)- $\alpha$ ; and  $n = 204$  for IL-6. <sup>c</sup> $n = 155$ . <sup>d</sup> $n = 234$ . <sup>e</sup> $n = 154$  for cytokines and  $n = 179$  for CRP. Quantile regression process plots are shown in Extended Data Fig. 6 for all analytes with suggestive associations to FaRMI ( $P < 0.1$ ) anywhere between the 25th and 80th percentile.

What should these values tell us? 5 times  $p < 0.05$  in a grid of  $6 \times 11 = 66$  tests? Quantile regression that has been “adjusted” by the data?

The FaRMI calculation is overly complicated. Probably the calculation can never be replicated by changing software, changing samples, unclear wording, suboptimal dimension reduction and data snooping.

Farm home microbiota-like community composition was modeled in LUKAS1 with logistic regression analysis (PROC LOGISTIC statement, SAS version 9.3). The home location on a farm or non-farm rural environment was the dependent variable and the main components of PCoA axis scores of  $\beta$ -diversity matrices were the predictor variables. Bacterial and fungal microbiota were investigated separately. For both bacteria and fungi, separate models were built using axis scores from PCoA of abundance-unweighted and -weighted  $\beta$ -diversity matrices. The PCoA axes were selected based on the scree plot method including axes above the point at which the variance explained by the additional axes levels off (Supplementary Fig. 3). The models give an estimate of the probability that the sample is from a farm home. The farm home likeness of the microbial composition in the LUKAS2 non-farm homes was then estimated by applying the regression coefficients obtained from the LUKAS1-based models to the corresponding microbial data from LUKAS2 samples. Some analyses were performed in non-farm homes of both LUKAS2 and LUKAS1 to obtain increased sample size and power if results remain comparable as was observed. Due to the discovered association with asthma, the probability that was modeled based on the relative abundance-weighted bacterial/archaeal  $\beta$ -diversity was named FaRMI and was studied further in greater detail.

And isn't that just an association that may have a rather simple explanation?

As FaRMI is weakly associated with muramic acid concentration in dust, the authors make Gram-positive bacteria responsible for the effect. The rhizosphere of soil is extremely rich of bacteria. The world's first soil atlas showed hundreds of taxa but never differentiated between water resistant, gram positive and less water resistant gram-negative taxa. Maybe Gram positive Streptococcaceae are ubiquitous and depend on where you draw your samples?

FaRMI is found in non farm / rural children by bacterial/archaeal operational taxonomic units (OTUs) of soil origin which basically confirms my initial assumption: There was the same contamination of soil both in farm and non-farm homes if we look at supplement table 6 where walking indoors with outdoor shoes results in significant higher FaRMI values... Maybe the microbiome hype is already over.

Our results warrant translational studies to confirm the causal relationship through indoor microbial exposure-modifying intervention that may also form a novel strategy for primary asthma prevention.

Good luck with your future studies.

BTW - The [scripts at Github](#) are useless references to shell and Python scripts that will never run due to “” characters. And what about that baby code?

```
outfolder=getwd()
eigenfile <- paste(outfolder, "/", prefix, "_PCoA_eigenvalues.txt",
sep="")
writerow <- paste("Eigenvalue min / max: ", min.eigen, " / ",
max.eigen, sep="")
write(writerow, file=eigenfile, append=F)
writerow <- paste("Sum of all eigenvalues: ", round(neg.eigensum,
digits=6), sep="")
write(writerow, file=eigenfile, append=T)
writerow <- paste("Sum of all eigenvalues (negatives as 0): ",
round(nonneg.eigensum, digits=6), sep="")
write(writerow, file=eigenfile, append=T)
writerow <- "Eigenvalues (pos & neg): "
write(writerow, file=eigenfile, append=T)
writerow <- paste(pcoa$value$Eigenvalues, collapse="\t")
write(writerow, file=eigenfile, append=T)
writerow <- "Percents (Negatives as negatives): "
write(writerow, file=eigenfile, append=T)
writerow <- paste(paste(neg.percent, " %", sep=""), collapse="\t")
write(writerow, file=eigenfile, append=T)
writerow <- "Percents (Negatives as 0): "
write(writerow, file=eigenfile, append=T)
writerow <- paste(paste(nonneg.percent, " %", sep=""), collapse="\t")
write(writerow, file=eigenfile, append=T)
```

Using R heredoc syntax I can rewrite 20 unreadable by 8 readable lines.

```
tmp <- 'Eigenvalue min / max: min.eigen / max.eigen
Sum of all eigenvalues: neg.eigensum
Sum of all eigenvalues (negatives as 0): nonneg.eigensum
Eigenvalues (pos & neg): pcoa
Percents (Negatives as negatives): neg.percent %
Percents (Negatives as 0): nonneg.percent %
for (i in
c("min.eigen", "max.eigen", "neg.eigensum", "nonneg.eigensum", "pcoa$value
$Eigenvalues", "neg.percent", "nonneg.percent") ) { tmp <-
gsub(i, get(i), tmp) }
write(tmp, file=paste(getwd(), out, "/"))
```

And why moving to SAS for a simple logistic regression? Is there anyone else in the academic world who pays [\\$8,700 annually](#) for a basic SAS Windows Analytics package just for that reason?