Biostatistics: The Science Linking Data and Evidence-Based Practice

Tinku Thomas
Dept of Biostatistics
St. John's Medical College
Bengaluru

St. John's
National Academy of Health Sciences

Statistics and Biology

- **Karl Pearson –** application & development of mathematical statistics, encompassed the fields of **biology**, **anthropometry**, **medicine, psychology** and social history

- **Sir Ronald Aylmer Fisher –** English statistician, evolutionary **biologist,** mathematician, **geneticist**, and eugenicist.

- **Prasanta Chandra Mahalanobis –** Mahalanobis D was developed in his studies analysing anthropometric measurements of Anglo-Indians

Diagnostic Testing

- New tests need Good diagnostic accuracy
- Many measures such as sensitivity and specificity, positive and negative predictive values
- Accompanied by 95% confidence limits as precision of these measures
- Sometimes Gold Standard reference methods are un-available

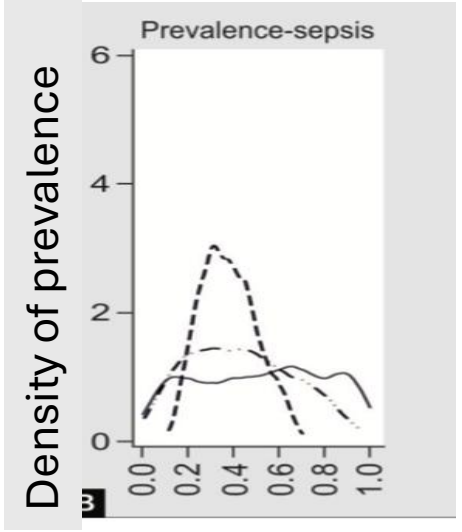
Blood Cultures and Molecular Diagnostics in Intensive Care Units to Diagnose Sepsis: A Bayesian Latent Class Model Analysis

Sriram Sampath[1], Jeswin Baby[2], Bhuvana Krishna[3], Nandini Dendukuri[4], Tinku Thomas[5]

Table 1: Cross-tabulation of results of both tests

Bacterial isolate results	*STD–positive*	*STD–negative*	*Difference in proportions (McNemar's exact–p)*
MOL–Positive	49 (13%)	141 (37%)	32%, p <0.01
MOL–Negative	18 (5%)	171 (45%)	

STD, standard blood culture methods; MOL, molecular methods



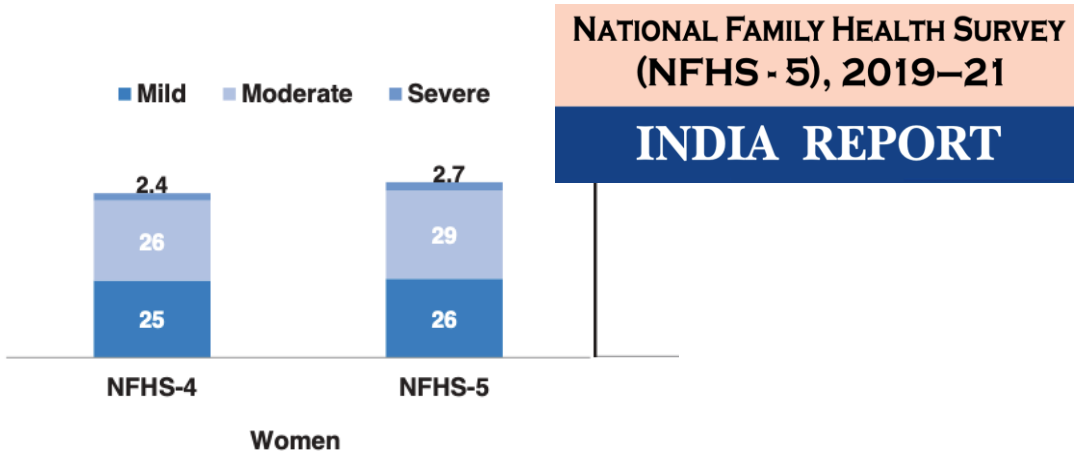
Priors: —, posterior independence assumptions: — —, posterior conditional dependence : — · · —

A model estimate of 43% prevalence by latent class method

Cut-Off for Diagnostics: Statistical Basis

***Figure 10.8* Trends in Anaemia Status**

Percentage of women and men age 15-49



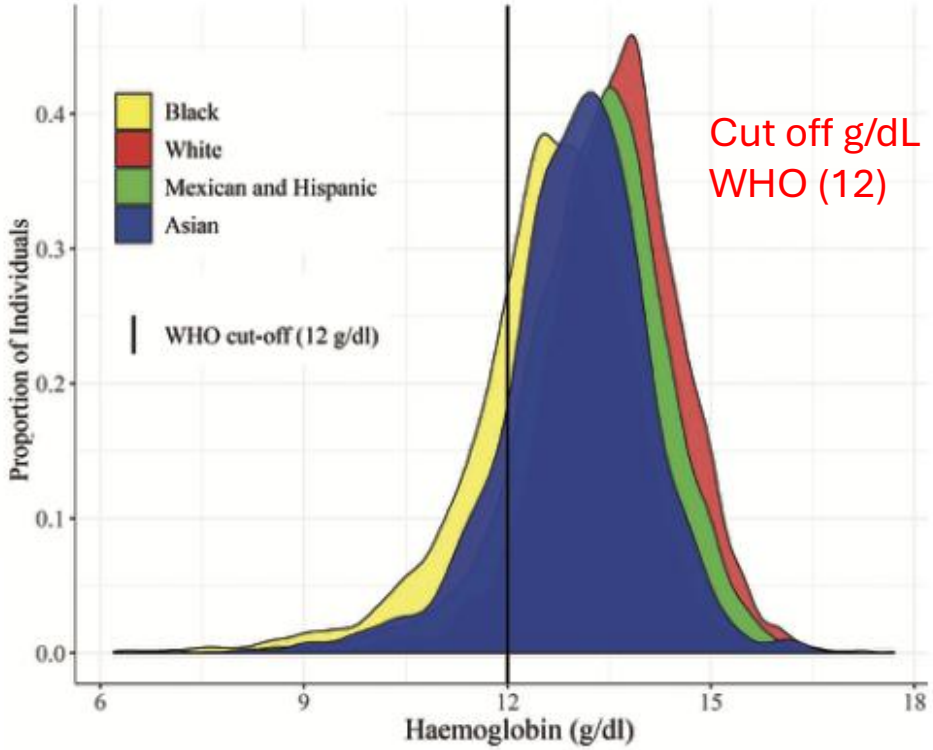
Overall-58 %

Is this prevalence true? Are we applying the correct cut off for anemia?

Indian J Med Res 150, October 2019, pp 385-389
DOI: 10.4103/ijmr.IJMR_334_18

Evaluation of haemoglobin cut-off for mild anaemia in Asians - analysis of multiple rounds of two national nutrition surveys

Jithin Sam Varghese[1], Tinku Thomas[2] & Anura V. Kurpad[3]



Cut off g/dL
WHO (12)

Fig. 2. Distribution of haemoglobin by race (unweighted sample means for each round in black).

Origin of Anemia Cut-Offs

These figures were chosen arbitrarily and it is still not possible to define normalcy precisely.

WORLD HEALTH ORGANIZATION
TECHNICAL REPORT SERIES
No. 405

NUTRITIONAL ANAEMIAS

Report of a
WHO Scientific Group

3. CRITERIA FOR THE DIAGNOSIS OF ANAEMIA

In detecting and evaluating an anaemia problem in a community, reference standards are necessary, even though they may be somewhat arbitrary. The report [2] of the 1958 WHO Study Group recommended haemoglobin values below which anaemia could be considered to exist. These figures were chosen arbitrarily and it is still not possible to define normality precisely.[3] However, more recent data [4] indicate that the values given previously should be modified. It is recommended that, in future studies, anaemia should be considered to exist in those whose haemoglobin levels are lower than the figures given below (the values given are in g/100 ml of venous blood of persons residing at sea level):

children aged 6 months to 6 years:	11
children aged 6–14 years:	12
adult males:	13
adult females, nonpregnant:	12
adult females, pregnant:	11

Distributional cut offs

Guideline on haemoglobin cutoffs to define anaemia in individuals and populations

2024

World Health Organization

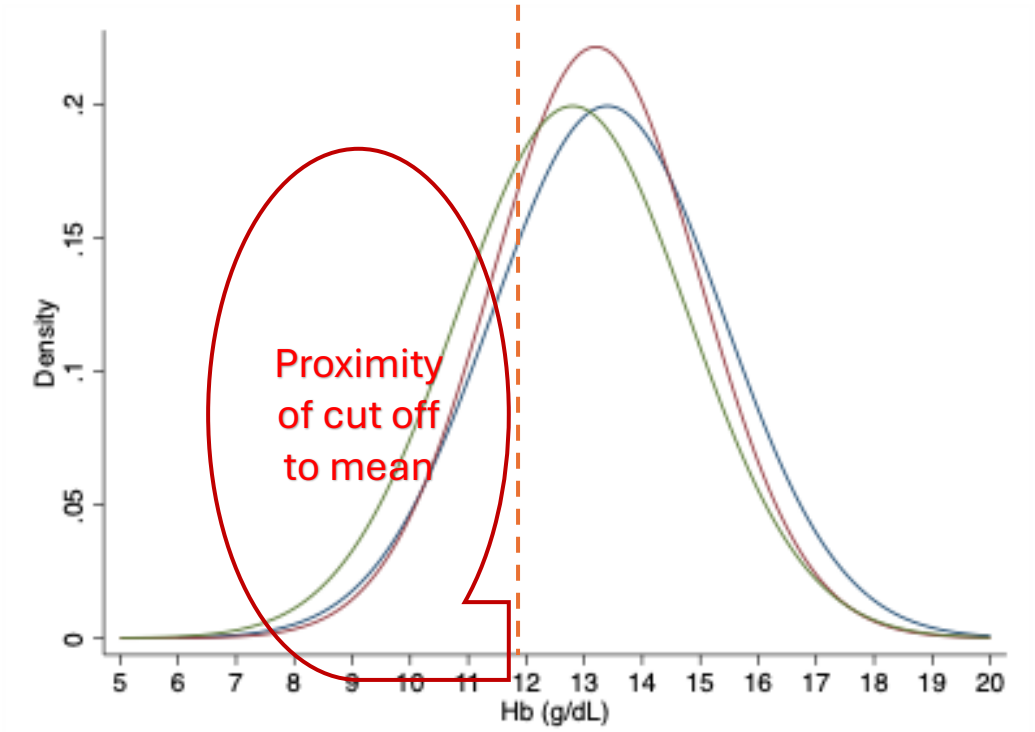
1. Haemoglobin cutoffs to define anaemia

Question 1

What should be the haemoglobin cutoffs to define anaemia in individuals and in populations?

Use of statistical and/or clinical outcomes

Lower 5th or 2.5th percentile cut off?



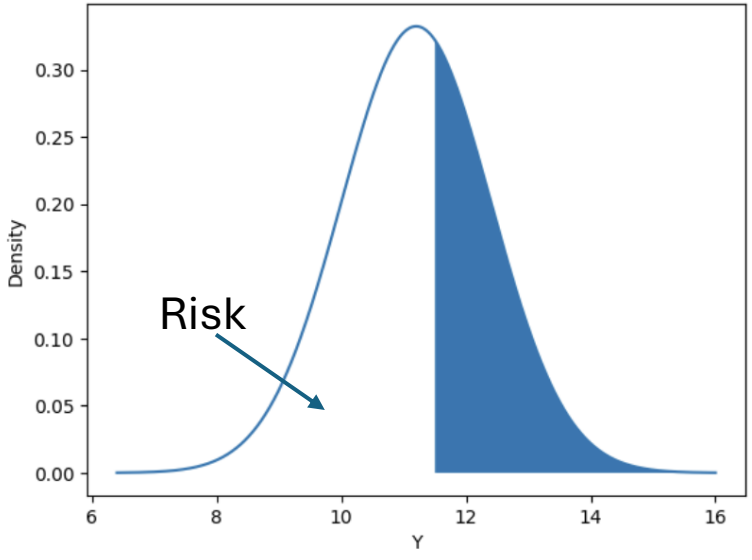
- If the distributional cut off changes to 11.2g/dL the prevalence of Anemia in Indian women reduces to **36%**
- The influence of cut off is just an example and would be true for many biomarkers
- Important for public health

Risk of Deficiency Rather than Binary Classification

- Let Y be a random variable denoting the Hb in a healthy population and has a probability distribution (Say, Normal)
- Let X be a random variable of measured Hb
- Then the risk of low Hb is

$$r(Hb) = P(Hb_{ref} > Hb_{obs})$$

The beauty of this method is that it can be applied to any probability distribution



> Indian Pediatr. 2023 Oct 15;60(10):804-810. Epub 2023 Jun 21.

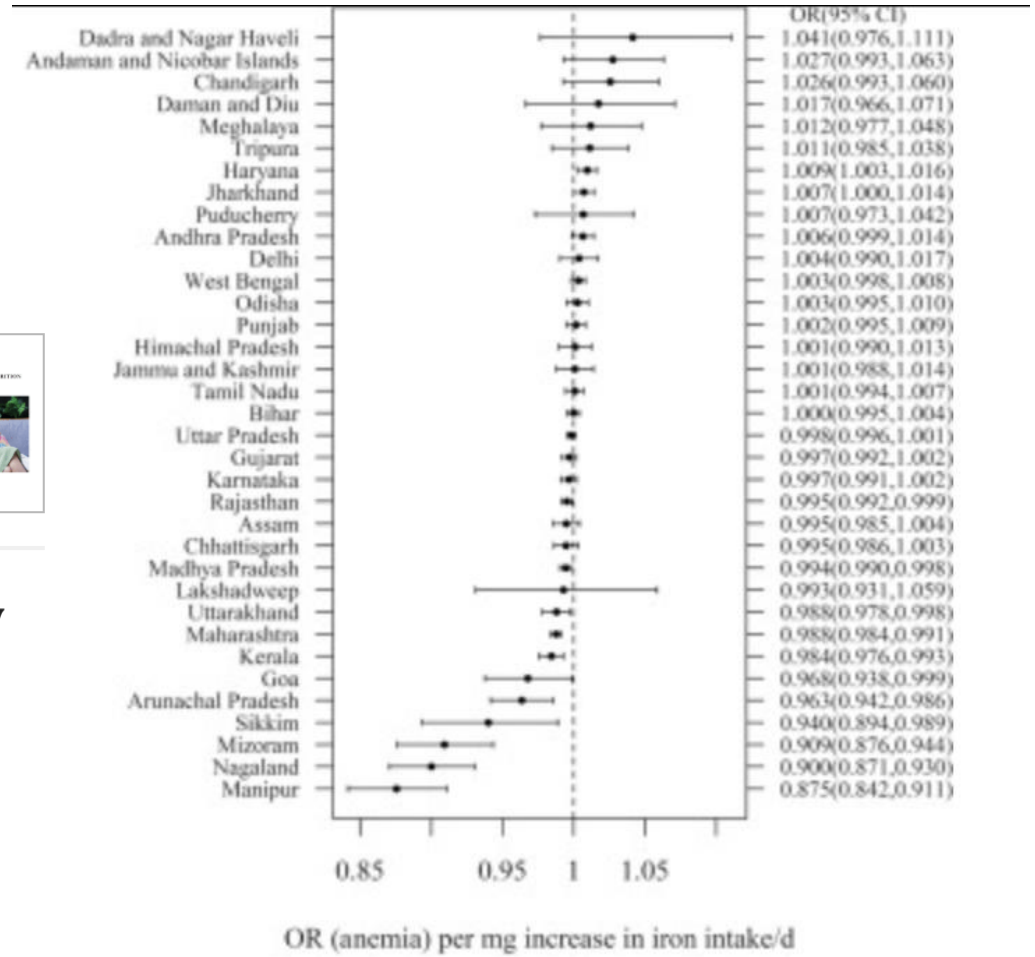
Customization of WHO Under-five Growth Standards for an Appropriate Quantification of Public Health Burden of Growth Faltering in India

Santu Ghosh [1], Rajesh Majumder [2], Harshpal Singh Sachdev [3], Anura V Kurpad [4], Tinku Thomas [5]

Risk Factor Assessment of Diseases

- Can use regression models to examine the independent and joint risk of disease with respect to some exposures
- How can multiple survey data be utilized?

Statistical matching of NSS and NFHS



ELSEVIER

The Journal of Nutrition

Volume 149, Issue 5, May 2019, Pages 831-839

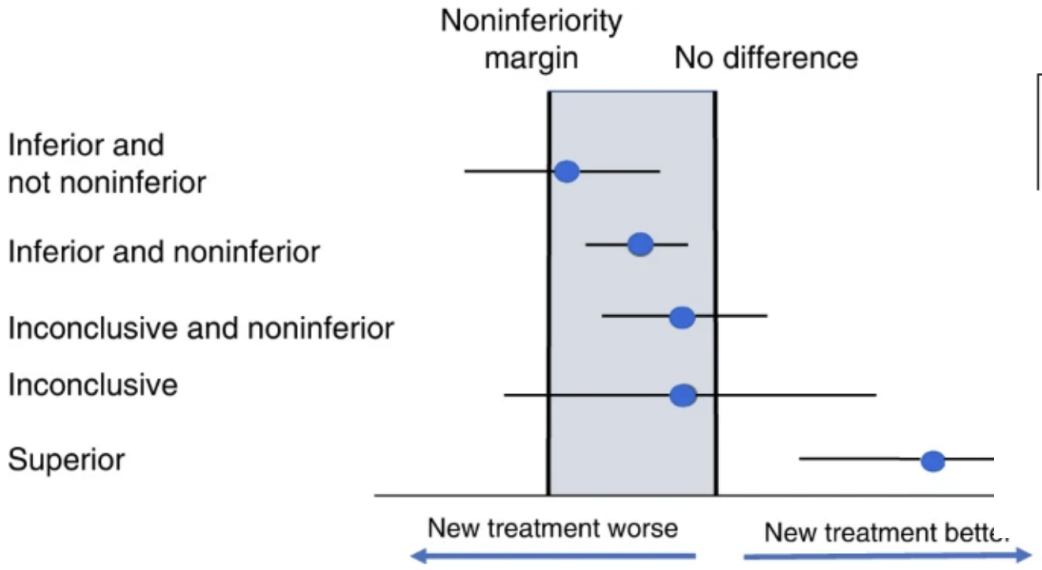
Dietary Iron Intake and Anemia Are Weakly Associated, Limiting Effective Iron Fortification Strategies in India [1], [2]

Swaminathan Sumathi [1], Ghosh Santu [2], Varghese Jithin Sam [1], Sachdev Harshpal S [3], Kurpad Anura V [4], Thomas Tinku [2]

Clinical and Community Trials: A Design for Ascertaining Causal Effects

- Noninferiority trials test if a new experimental treatment is NOT unacceptably less efficacious than an active control treatment
- Determination of non-inferiority margin is challenging: absolute vs relative risk difference

Fig. 1: Treatment effect and 95% confidence interval.



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Two Randomized Trials of Low-Dose Calcium Supplementation in Pregnancy

Pratibha Dwarkanath, Ph.D., Alfa Muhihi, M.D., M.P.H.,
Christopher R. Sudfeld, Sc.D., Blair J. Wylie, M.D., M.P.H., Molin Wang, Ph.D.,

Time-to-Event Analysis

- Survival analysis models **time to an event**, and in doing so works with **event rates over time**, rather than just whether an event occurred.
- When an event can repeat in an individual, we call it **recurrent events**.
- Complexity in recurrent event modelling occurs when the **observation** in some individuals is **truncated**.

RESEARCH ARTICLE

Adjusting for truncated study duration in recurrent event analysis: A weighting approach for clinical trials

John Michael Raj A[1,2], Tinku Thomas[2*], Pratibha Dwarkanath[3]

1 Center for Doctoral Studies, Manipal Academy of Higher Education, Manipal, India, **2** Department of Biostatistics, St. John's Medical college, Bangalore, India, **3** Division of Nutrition, St. John's Research Institute, Bangalore, India

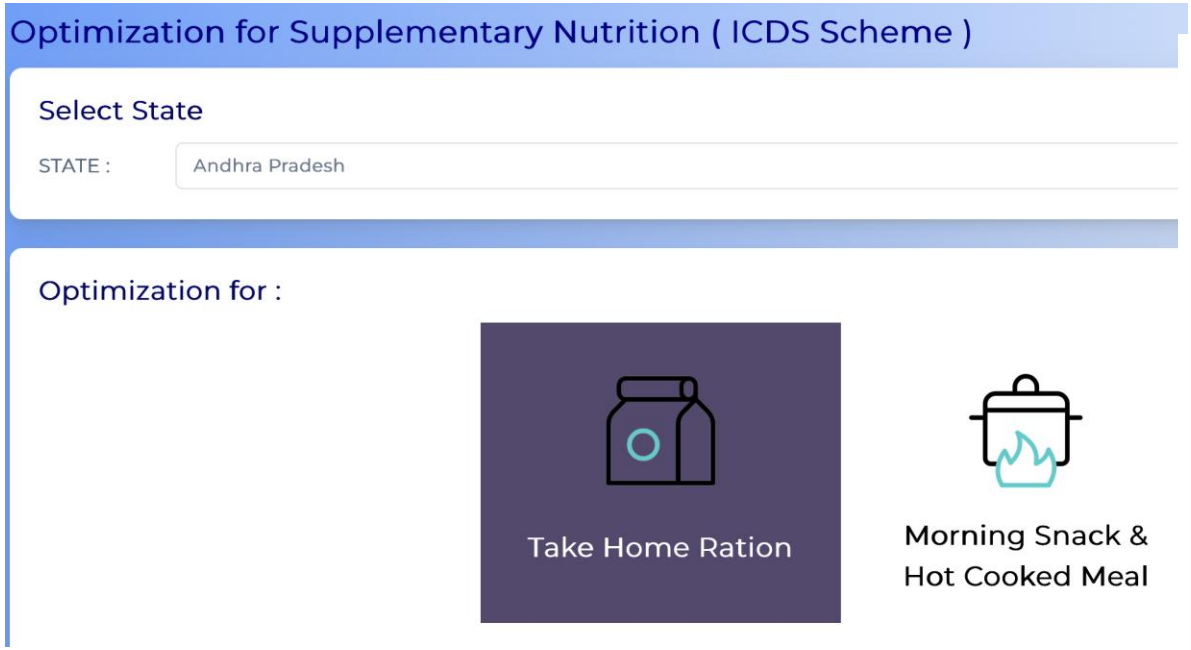
* tinku.sarah@sjri.res.in

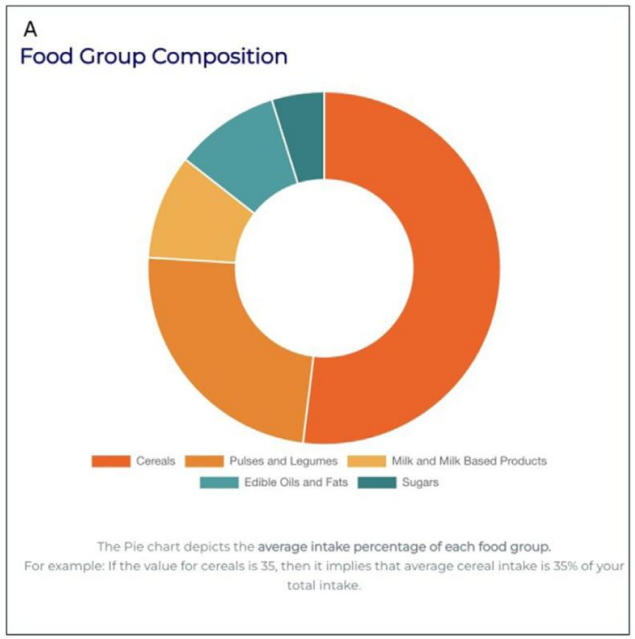
Table 2. Bias and mean beta estimates of total time weighted analysis.

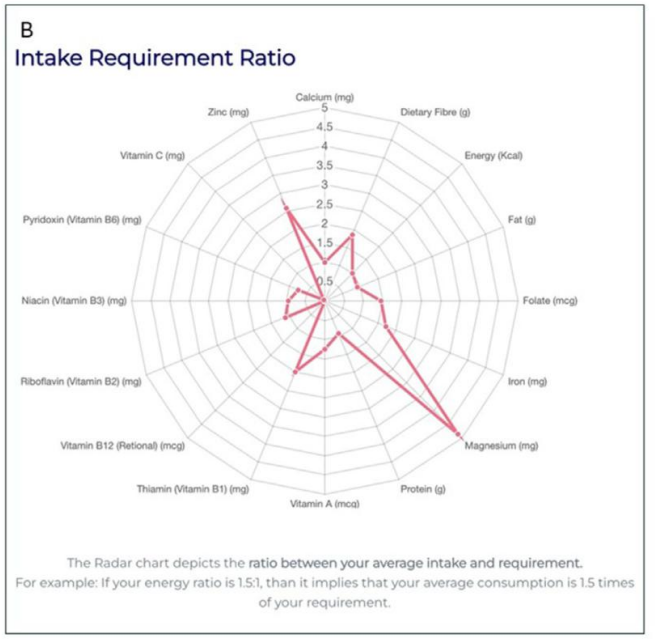
Scenarios of maximum number of visits	True beta	Bias	Estimated beta	
			Mean	Min, Max
3 to 9 Visits				
Unweighted	3	0.04	3.04	2.78, 3.38
Weighted	3	0.03	3.03	2.77, 3.37
4 to 9 visits				
Unweighted	3	0.05	3.05	2.77, 3.40
Weighted	3	0.04	3.04	2.77, 3.41

Tools for Diet Optimization

- Linear and non-linear programming as an effective tool for formulating nutritionally sound food-based recommendations in a limited budget
- ICDS Supplementary Nutrition







A Great Team: Physiology, Pediatrics, Nutrition, Statistics, Data Science & AI

Anura V Kurpad

HPS Sachdev

Tinku Thomas

Santu Ghosh

Sumithra Selvam

John Michael

Fathima Ayoob

Jawahar Manivannan

Ashikh Ahamed


Thank You