Statistical Methods: Need for a Rethink

We read with interest the recent article [1] in Indian Pediatrics on reporting statistical results, and would like to commend the authors on lucidly summarizing such an important topic. Indeed, as the British economist Ronald Coase had said, *"if you torture the data long enough, it will confess."* Arguably, nowhere is the sentence more appropriate than describing the biomedical fraternity's obsession with statistical significance. A recent analysis [2] found a significant increase in reporting of *P*-values in Medline abstracts over the past twenty-five years, and unsurprisingly most of them reported significant results. Moreover, *P*-values between 0.041 and 0.049 have increased manifold in the last couple of decades [3]. While this might imply an increase in 'P-hacking', the consequences of which are being debated worldwide, a more worrisome trend is the over-reliance on frequentist inference and the widespread misunderstanding of the *P*-value. Sterne, *et al* [4], in their authoritative piece in the British Medical Journal, stated *"an arbitrary division of results, into "significant" or "non-significant" according to the P value, was not the intention of the founders of statistical inference."* The American Statistical Association (ASA) has acknowledged common misuses of the *p*-value and categorically asserted *"P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone"* [5]. At best, *P*-value can inform our decision regarding whether the data under consideration is compatible with a particular null hypothesis, but by itself it is not sufficient to comment on neither the truth of the null hypothesis, nor the biological or clinical significance of the results. While statisticians have voiced such concerns for years, we continue to teach, learn, use, over-use and misinterpret *p*-values in our literature. These concerns may sound only semantic, but are by no means irrelevant when we consider the larger landscape of publication bias. The use of Bayesian inference can potentially circumvent some of the problems [4], for example by generating a reliable 'credible interval' for an estimated parameter even for small sample sizes. Relatively modern statistical techniques like bootstrap and confidence distribution are worth exploring as well. Although the biomedical community has traditionally been slow in taking up and implementing new approaches, we hope a greater awareness of proper statistical methods and the willingness to adopt new techniques can change the scenario. This can be made possible through closer collaboration between biomedical scientists and statisticians.

*TAMOGHNA BISWAS AND #SUBHABRATA MAJUMDAR
From *Department of Pediatric Medicine
Medical College Kolkata, India;
and #School of Statistics,
University of Minnesota, Twin Cities Minneapolis, USA.
*tamoghnab@gmail.com

REFERENCES

1. Khan AM, Ramji S. Reporting statistics in biomedical research literature: The numbers say it all. Indian Pediatr. 2016;53:811-4.
2. Chavalarias D, Wallach JD, Li AH, Ioannidis JP. Evolution of reporting P values in the biomedical literature, 1990-2015. JAMA. 2016;315:1141-8.
3. de Winter JC, Dodou D. A surge of p-values between 0.041 and 0.049 in recent decades (but negative results are increasing rapidly too). Peer J. 2015;3:e733.
4. Sterne JA, Davey Smith G. Sifting the evidence-what's wrong with significance tests? BMJ. 2001;322:226-31.
5. Wasserstein RL, Lazar NA. The ASA's statement on p-values: Context, process, and purpose. The American Statistician. 2016:70:129-33.

Utilization of Anganwadi Services in a Rural Population of Kerala

Integrated Child Developmental Services (ICDS) has played an important role in improving the health and education of children of India [1]. The state of Kerala has better health and socio-economic indicators compared to most other parts of the country. However, there are not many studies assessing the utilization of Anganwadi services in Kerala [2-4]. This study was done in a rural area near Kochi. In the first stage of sampling, we selected four wards randomly. In the selected wards, the investigators visited consecutive houses in all directions from the entry point to the ward. We collected data from 352 households regarding utilization of Anganwadi services in the past 1 year. The information was obtained from an adult member present at the time of visit, preferably a lady.

In our sampling frame, 136 families had children below 6 years, 66 families had adolescent girls (11-19), and 193 women were in reproductive age group out of which 26