

- 8 Nakagawa I, Nakamura K, Oyama M, *et al.* Long-term effects of the Niigata-Chuetsu earthquake in Japan on acute myocardial infarction mortality: an analysis of death certificate data. *Heart* 2009; 95: 2009–2013.
- 9 Bertazzi PA, Bernucci I, Brambilla G, *et al.* The Seveso studies on early and long-term effects of dioxin exposure: a review. *Environ Health Perspect* 1998; 106: Suppl. 2, 625–633.

- 10 Poole WK, Chi JS, Walton JD, *et al.* Increased cardiovascular mortality associated with the turn of the millennium in Los Angeles County, California. *J Epidemiol Community Health* 2005; 59: 205–206.

DOI: 10.1183/09031936.00078910

From Paul's predictions in the World Cup to the publication bias in genetic studies on complex traits

To the Editors:

Recently, the octopus Paul has attracted much attention through the general media by his accurate predictions in the FIFA 2010 World Cup. During the World Cup, Paul predicted the outcomes of eight matches with 100% success. Assuming that Paul has 50% accurate choices as independent coin flips, the probability for a continual eight success prediction is 0.0039. However, we may think that Octopus Paul is more likely to select the box with the German flag, for some reason. For example, when predicting the outcome of the Germany games Paul selected the box with the German flag five times out of seven, giving a percentage rating of 71.4% for selection of the German flag. Assuming Paul is more likely to select the German flag with a probability rate of 0.714, the probability for the correct predictions Paul achieved in FIFA 2010 World Cup is 0.0076 ($0.714 \times 0.714 \times 0.714 \times 0.714 \times 0.714 \times 0.286 \times 0.286 \times 0.5$). In the academic world, we usually select a probability boundary of 0.05 or 0.01 to reject the null hypothesis [1]. Therefore, we may conclude that Paul has the ability to correctly predict the outcomes of the world cup based on traditional statistical theory. Let us say the psychic octopus Paul has supernatural powers of predictions.

However, Paul's psychic powers became the focus of media attention just after the quarter-finals. At the group stage and the quarter-finals, Paul had predicted five accurate outcomes, with a probability value of 0.031 and 0.074, respectively, for Paul's predictions similar to a coin flip or Paul's likelihood in selecting the German flag. It is obvious that there were many people around the world who were using animals or other ways to predict outcomes of the World Cup. Before the semi-finals, we can assume there were more than 100 attempts (experiments) to guess the results of the World Cup. The p-values of 0.031 or 0.074 should be adjusted for multiple tests. With the Bonferroni correction, the adjusted p-value will be close to 1. Therefore, the significance of Paul's prediction before the semi-finals is unremarkable. After the quarter-finals, Paul predicted three accurate outcomes with p-values of 0.125 and 0.102, respectively for a guess with a coin flip or Paul's likelihood of selecting the German flag. With the eyes of the whole world now watching Paul's predictions after the quarter-finals, we can assume that it was only one experiment, therefore, we do not need to adjust for multiple tests. However, the p-values of 0.125 or 0.102 are not significant in the view of statistics. Therefore, we conclude that the octopus, Paul, is a normal octopus without psychic powers of predictions.

Comparing the experience of Paul's prediction with the challenges of publication bias in genetic studies, we find they are extremely similar. This phenomenon may partly explain the inconsistencies in candidate gene association and genome-wide association (GWA) studies on complex traits. Before the era of GWA studies, thousands of candidate gene association studies had been conducted. Due to publication bias and authors' desire and hidden intentions to publish positive results, many negative results were not submitted or accepted for publication. The published positive findings included significant false-positive results due to poorly accounting for multiple tests in general. Eventually, many of the candidate gene association studies on complex diseases (traits) were regarded as fruitless because many findings could not be consistently replicated in subsequent studies [2]. Now we are in the era of GWA studies. Hundreds of GWA studies have been published with many novel genes and new variants recently identified to be associated with disease phenotypes. However, among the reported genetic variants, how many are falsely positive? The published GWA studies usually included replications for their reported genetic variants in independent populations. However, are these replication attempts similar to Paul's prediction in the World Cup? How many GWA studies have been conducted and failed replication? As more and more GWA studies have been conducted, the issue of bias of publication will become more and more significant [3]. For example, GWA studies on asthma have shown inconsistency [2, 4, 5]. Will the GWA studies end up similar to candidate gene association studies? What can we do in the academic community to avoid the publication bias and include all negative findings?

Addressing this at the start, the National Institutes of Health has undertaken efforts to catalogue all published GWA studies. GWA studies are important and productive as they are identifying more and more genes related to complex diseases. However, it may be more important to amass all genetic information for all GWA investigations. It is timely to register all GWA studies regardless of their sample size and their positive or negative findings, and consequently have a collective dataset to hold all codes and genotype–phenotype association information for the human genome. This collective data will eventually help interpret our millions of genetic codes regarding our health. Although not a problem with advancements in computer storage technology, problems may arise in terms of ethics and practicality. However, if we still

ignore these issues, in the near future publication bias may teach us another lesson in our endeavour to understand the codes of the human genome.

G. Zhang and P.N. LeSouëf

School of Paediatrics and Child Health, Faculty of Medicine and Dentistry, The University of Western Australia, Perth, Australia.

Correspondence: G. Zhang, School of Paediatrics and Child Health, University Of Western Australia, GPO Box D184, Perth WA 6840, Australia. E-mail: gczhang@meddent.uwa.edu.au

Support Statement: G. Zhang is supported by the National Health and Medical Research Council, Australia.

Statement of Interest: None declared.

REFERENCES

- 1 Neyman J, Pearson ES. On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference, Part I. *Joint Statistical Papers*. Cambridge, Cambridge University Press, 1928; pp. 1–66.
- 2 Zhang G, Goldblatt J, LeSouef P. The era of genome-wide association studies: opportunities and challenges for asthma genetics. *J Hum Genet* 2009; 54: 624–628.
- 3 Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA* 2008; 299: 1335–1344.
- 4 Hancock DB, Romieu I, Shi M, *et al*. Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in mexican children. *PLoS Genet* 2009; 5: e1000623.
- 5 Mathias RA, Grant AV, Rafaels N, *et al*. A genome-wide association study on African-ancestry populations for asthma. *J Allergy Clin Immunol* 2010; 125: 336–346.

DOI: 10.1183/09031936.00118510