**Response: To statistical significance and beyond**

Sizheng Steven Zhao1,2, David M Hughes3

1 Musculoskeletal biology I, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK

2 Department of Academic Rheumatology, Aintree University Hospital, Liverpool, UK

3 Department of Biostatistics, Institute of Translational Medicine, University of Liverpool, Liverpool, UK

Correspondence to:

Dr Sizheng S Zhao

Institute of Ageing and Chronic Disease,

University of Liverpool,

Aintree University Hospital

Liverpool

L9 7AL

UK

Email: stevenzhao25@gmail.com

We are grateful for Dr Kardes’ correspondence and the opportunity to engage in scientific discourse. We would like to highlight that our letter really was not about gout (or bruxism), but instead the effect of seasonality on patient reported outcome measures, particularly in rheumatic diseases [1]. Our intention, by citing Dr Kardes, was so that readers can access a body of works related to GoogleTrends – not to criticise. We hope this resolves any unintended misunderstanding or offence.

We reiterate the conclusion of our letter: patients may report swelling during assessment of their rheumatic disease that could be influenced by other factors, namely seasonality; therefore, the critique that “it is not possible to attribute all search queries […] to rheumatic diseases” is not in disagreement.

Dr Kardes further suggested that 1) the northern and southern hemispheres should be analysed separately to demonstrate trends in opposite seasons and 2) seasonal trends should undergo statistical testing for significance.

First, we agree that analyses should be performed for each hemisphere, but did not wish to labour the same point already made by Dr Kardes’ paper within the constraints of a letter. We refer readers to Figure 1, which shows differences in search volumes between the USA and Australia.

Second, we encourage readers to move away from over-reliance on “statistical significance” [2]. A consistent doubling in search volume is meaningful whether or not it is *statistically* significant. Imagine there is a two-fold increase in the incidence of measles: do we wait for a statistical test to interpret its significance? Let us note the size of the effect, not only the P-value. Furthermore, since both Dr Kardes’ and our aims were to *describe* data, and not to interrogate a *causal* hypothesis, focusing on statistical significance has limited meaning. Traditional hypothesis testing depends on sample size which, again, has limited meaning since GoogleTrends provide *relative* search volumes.

We would like to leave readers with a quote from Amrhein and colleagues’ recent article in Nature: “How do statistics so often lead scientists to deny differences that those not educated in statistics can plainly see? For several generations, researchers have been warned that a statistically non-significant result does not ‘prove’ the null hypothesis […] Nor do statistically significant results ‘prove’ some other hypothesis. Such misconceptions have famously warped the literature with overstated claims and, less famously, led to claims of conflicts between studies where none exists. [2]”

**Disclosures**: The authors declare no conflicts of interest.



Figure 1. Seasonal variations in Google searches for “feet swelling” differ in northern and southern hemispheres.

**References**

1. Zhao S, Duffield SJ, Hughes DM. Seasonal variations of Google searches for joint swelling: implications for patient-reported outcomes. Clin Rheumatol. 2019 [cited 2019 May 22]; Available from: http://link.springer.com/10.1007/s10067-019-04534-0

2. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature 2019 [cited 2019 May 22]. Available from: https://www.nature.com/articles/d41586-019-00857-9