**LETTER TO THE EDITOR CONCERNING STEVELINK ET AL. (2021):
“CHRONIC FATIGUE SYNDROME AND OCCUPATIONAL STATUS:
A RETROSPECTIVE LONGITUDINAL STUDY”**

Dear Editor,

Occupational Medicine recently published a paper from Stevelink et al. (2021) called “Chronic fatigue syndrome and occupational status: a retrospective longitudinal study.” Unfortunately, the paper features major technical and methodological errors that warrant urgent editorial attention.

To recap: The study started with 508 participants and had follow-up data for 316 of them. The primary outcome was occupational status. Many participants had dropped out by follow-up—only 316, or 62%, provided follow-up data. Of those 316, 88% who reported no change in employment status. As a group, the participants experienced either no changes or only insignificant ones in a range of secondary outcomes, including fatigue and physical function. The poor follow-up scores on fatigue and physical function alone indicate that the group remained, collectively, severely disabled after treatment.

In several sections of the paper, the authors’ description of their own statistical findings is incorrect. They make a recurring elementary error in their presentation of percentages. The authors repeatedly use the construction “X% of *patients who did Y at baseline*” when they should have used the construction “X% of all 316 patients (*i.e., those who provided follow-up data*)”.

This recurring error involving the core findings undermines the merit and integrity of the entire paper.

For example, in the Abstract, the authors state that “53% of patients who were working [at baseline] remained in employment [at follow-up].” This is not accurate. Their own data (Table 2) show that 185 patients (i.e., 167 + 18) were working at baseline, and that 167 patients were working at both time points. In other words, the proportion working continuously was in fact 90% (i.e., 167 out of 185). The “53%” that the authors refer to is the percentage of the sample who were employed at both time points (i.e., 167 out of 316), which is an entirely different subset. They have either misunderstood the percentage they were writing about, or they have misstated their own finding by linking it to the wrong percentage.

This error is carried over into the section on “Key Learning Lessons”, where the authors state that “Over half of the patients who were working at baseline were able to remain in work over the follow-up period...” While 90% is certainly “over half”, it seems clear that this phrasing is again incorrectly referring to the 53% subset.

The same error is made with the other key findings. For example, the Abstract states that “Of the patients who were not working at baseline, 9% had returned to work at follow-up”. But as above, this is incorrect. A total of 131 patients (i.e., 104 + 27) were recorded as “not employed” at baseline and 27 were recorded as not working at baseline but as working at follow-up. This is 21%, not 9%. Once again, the authors appear to misunderstand their own findings. The “9%” they refer to is a percentage of *the sample of 316*; it is not, as they have it, a percentage of *that*

subset of the sample who were initially unemployed. This erroneous “9%” conclusion appears as well in the “Key Learning Lessons” and in the Discussion.

And again, the authors state in the Abstract that “*of those working at baseline, 6% were unable to continue to work at follow-up*”, a claim they repeat in the section on “Key Learning Lessons” and in the Discussion. This statement too is wrong. Once more, the authors mistakenly interpret a percentage of the sample of 316 as if it were a percentage of a targeted subset. In this case, they *think* they are referring to a percentage of patients working at baseline, but they are *actually* referring to a percentage of the full group that provided follow-up data.

The authors present the raw frequency data in Table 2. Readers can see for themselves how their sample of 316 patients is cross-tabulated into four subsets of interest (i.e., “working at baseline and follow-up”; “not working at baseline and follow-up”; “dropped out of work at follow-up”; “returned to work at follow-up”). From Table 2, it is clear that the prose provided in the body of the paper is at odds with the actual data.

It is undeniable that the text of this paper is replete with elementary technical errors, as described. Inevitably, the narrative is distorted by the authors’ failure to understand and correctly explain their own findings. It is unclear to us how these basic and self-evident errors were not picked up during peer-review. Although we don’t know the identities of the peer-reviewers, we speculate that groupthink and confirmation bias will have played their part. After all, it is generally reasonable for peer-reviewers to presume that authors have understood their own computations.

There are several other features of this paper that cause concern. These include the following:

- The authors state that they evaluated participants using guidance from the UK's National Institute for Health and Care Excellence (NICE). (Presumably they are referring to the 2007 NICE guidance, not the revision published in October 2021.) But the reference for this statement is a 1991 paper that outlines the so-called "Oxford criteria", a case definition that differs significantly from the 2007 NICE guidance. Moreover, in a paper about the same participant cohort previously published by Occupational Medicine — "Factors associated with work status in chronic fatigue syndrome"-- the authors state explicitly that these patients were diagnosed using the Oxford criteria. This inconsistency is non-trivial, because the differences between these two diagnostic approaches have substantive implications for how the findings should be interpreted. The authors' confusion over the matter is hard to comprehend and raises fundamental questions about the validity of their research.
- According to Table 1, there were either no changes or no meaningful changes in average scores for fatigue, physical function, and multiple other secondary outcomes between the preliminary sample of 508 and the final follow-up sample of 316. The authors themselves acknowledge that the patients who dropped out before follow-up were likely to have had poorer health than those who remained. Therefore, the fact that Table 1 presents combined averages for the entire preliminary sample -- i.e., combined averages for patients who dropped out and those who did not -- muddies the waters. Presenting

combined baseline scores for all patients will mask any declines that occurred for these variables in the subset who were followed up. It would have been far more appropriate to have isolated and presented the baseline data for the 316 followed-up patients alone. Doing so would have reflected the authors' research question more correctly, as well as enabling readers to make their own like-with-like comparisons.

- Finally, the authors state that “*Studies into CFS have placed little emphasis on occupational outcomes, including return to work after illness.*” However, they conspicuously fail to mention the PACE trial, a high-profile large-scale British study of interventions for CFS. The PACE trial included employment status as one of four objective outcomes, with the data showing that the interventions used -- the same ones as in the *Occupational Medicine* study -- have no effect on occupational outcomes. This previous finding is so salient to the present paper that it is especially curious the authors have chosen to omit it. The omission is all the more disquieting given that the corresponding author of the paper was a lead investigator on the PACE trial itself.

Authors of research papers have an obligation to cite seminal findings from prior studies that have direct implications for the target research question. Not doing so -- especially where there is overlapping authorship -- falls far short of the common standards expected in scientific reporting.

Even putting these additional matters aside, the technical errors that undermine this paper's reporting of percentages render its key conclusions meaningless. The sentences used to describe the findings are simply incorrect, and the entire thrust of the paper's narrative is thereby

contaminated. We believe that allowing the authors to publish a correction to these sentences would create only further confusion.

We therefore call on the journal to retract the paper.

Yours, etc.

Brian M. Hughes

School of Psychology

National University of Ireland, Galway

e-mail: brian.hughes@nuigalway.ie

David Tuller

University of California, Berkeley

e-mail: davetuller@berkeley.edu

References

Stevelink, S. A. M., Mark, K. M., Fear, N. T., Hotopf, M., & Chalder, T. (2021). Chronic fatigue syndrome and occupational status: a retrospective longitudinal study. *Occupational Medicine*. Online ahead of print. doi: 10.1093/occmed/kqab170.

Conflicts of Interest

David Tuller is a senior fellow in public health and journalism at the Center for Global Public Health at the University of California, Berkeley; members of the ME/CFS patient and advocacy community have donated to crowdfunding campaigns in support of his position at Berkeley.