

Years of genomic research botched with errors thanks to Excel quirk

A [new paper](#) has revealed the vast extent of errors in published genomics research,...[thanks to] an unfortunate quirk of Microsoft Excel...[Scientists] scanned 7,500 Excel files with gene lists accompanying 3,600 papers in 18 journals over a 10-year period. One-fifth of the files had easily identified errors, which is “quite striking and a little bit embarrassing,” says Mark Ziemann....

What happened? By default, Excel...convert some gene symbols to dates and numbers. For example, instead of writing out “Membrane-Associated Ring Finger (C3HC4) 1, E3 Ubiquitin Protein Ligase,” researchers have dubbed the gene MARCH1.

...

The thing is, there’s no excuse for data experts to be missing these errors with such regularity in the first place...It raises uncomfortable questions about the peer-review process at prestigious scientific journals, Ziemann says.

...

Still, although a one-fifth error rate may sound like a lot, it’s not nearly as bad as in the business world, where some studies suggest that as many as 90% of spreadsheets in use at companies [contain errors of one sort or another](#)....

The GLP aggregated and excerpted this blog/article to reflect the diversity of news, opinion and analysis. Read full, original post: [Years of genomics research is riddled with errors thanks to a bunch of botched Excel spreadsheets](#)