World-first program uncovers errors in biomedical research results

March 4 2019

Credit: University of Sydney

Just like the wrong ingredients can spoil a cake, so too can the wrong ingredients spoil the results in biomedical research. The difference is that the latter involves years of work, financial and personal investment

and promise.

Cancer researcher Professor Jennifer Byrne from the University of Sydney is hoping to change this with the creation of a world-first fact checking program that is tackling the problem of incorrectly published biomedical research results, whether intentional or otherwise.

In a paper, published in *PLOS ONE*, Professor Byrne and colleague Dr. Cyril Labbé of the University of Grenoble Alpes (France) detail 'Seek & Blastn', the fact-checking computer program they have developed and made freely available to researchers.

The program verifies the identities of published nucleotide sequence reagents (DNA and RNA constructs used to target genes) by seeking out sequences within papers and running them through a database holding the wealth of knowledge on genes to date.

"Biomedical reagents are like ingredients in cooking. You use them to discover your experimental results. Doing an experiment with wrong reagents either means that you cook something different from what you thought you were cooking, or what you cook is a failure," said Byrne, Professor of Medical Oncology in the Sydney Medical School.

In a cohort of 155 research papers the new fact-checker combined with manual analysis identified 25 per cent of papers as having sequence errors. The researchers were testing on a suspected group of the papers so while the figure doesn't reflect a baseline error rate, the numbers are still startling.

"That's quite a lot of wrong sequences in a small group of papers and there will be many more out there, unfortunately, given that nucleotide sequence reagents have been described in literally hundreds of thousands of biomedical publications," said Professor Byrne.

The researchers found that errors represented both identity errors (sequences which were completely incorrect) and typographic errors (sequences that contained the equivalent of spelling mistakes). The authors propose that sequence identity errors could represent a particular hallmark of research fraud, and could be applied to identify fraudulent papers and manuscripts.

"Our hope is that tools like Seek & Blastn will prospectively deter publications that describe incorrect nucleotide sequence reagents and may flag existing publications so that their conclusions can be re-evaluated," said Professor Byrne.

Errors uncovered included:

- Sequence reagents that are supposed to target a particular gene, but are in fact predicted to target a different gene from that stated in the publication, resulting in acquired data having nothing to do with system under study.
- Sequence reagents that are not supposed to target any gene (as a negative control) but instead are predicted to target a human gene , meaning researchers aren't comparing experimental data to a proper negative control.
- Sequence reagents that are supposed to target a human gene that in fact don't seem to target any gene, which could result in experiments not working but researchers being unaware.

Professor Byrne, named in *Nature* journal's Top 10 researcher for 2017, is well known for her detective work uncovering fraudulent results published in scientific journals.

Her work so far has resulted in seventeen retractions, but she says the process is slow and arduous, with the lack of responses from journals disheartening.

In an editorial in *Nature* last month she wrote: "Such papers claim to uncover mechanisms behind a swathe of cancers and rare diseases. They could derail efforts to identify easily measurable biomarkers for use in predicting disease outcomes or whether a drug will work.

"We create the literature that we deserve. We must act against this under-recognized threat to valid science."

More information: Cyril Labbé et al. Semi-automated fact-checking of nucleotide sequence reagents in biomedical research publications: The Seek & Blastn tool, *PLOS ONE* (2019). DOI: 10.1371/journal.pone.0213266

Jennifer Byrne. We need to talk about systematic fraud, *Nature* (2019). DOI: 10.1038/d41586-019-00439-9

Provided by University of Sydney