The Use of “Trend” Statements to Describe Statistically Nonsignificant Results in the Oncology Literature

The use and interpretation of P values in the biomedical literature is problematic. The importance of a study is often inappropriately defined by the P value. This problem is highlighted by the use of “trend” to refer to statistically nonsignificant results. There is no definition of a trend toward statistical significance and, therefore, describing “almost significant” results as a trend introduces substantial subjectivity and the opportunity for biased reporting language that could mislead a reader (eg, assuming $P < .10$). To deemphasize $P$ values, some journals prohibit the use of statements about a trend toward significance.\(^1\) Instead, presentation and discussion of observed differences and their uncertainty (eg, CIs) are encouraged. The degree of overreliance on $P$ values, and how this overreliance results in unclear reporting practices, is not characterized in the oncology literature, to our knowledge. We examined recent original research articles in oncology journals with high impact factors to evaluate the use of statements about a trend toward significance to describe statistically nonsignificant results.

**Methods** | Two of us (K.T.N. and M.R.W.) reviewed all original research articles published from November 2016 to October 2017 in the *Journal of the National Cancer Institute* (JNCI), *Journal of Clinical Oncology* (JCO), *JAMA Oncology*, and *Lancet Oncology* to identify instances where describing a trend toward significance was used to describe a statistically nonsignificant result. We focused on this period to describe current practices. We compared the proportion of original research articles with at least 1 statement describing a trend toward significance across journals by comparing percentages and 95% CIs. This project was exempt from approval by the University of Pennsylvania Institutional Review Board because it was not human participants research.

**Results** | We found that 8.7% (63 of 722) of original research articles in major oncology journals used at least 1 trend statement to describe statistically nonsignificant results. There were

---

**Table. Examples of Problematic Uses of the Word “Trend” to Describe Statistically Nonsignificant Results**

<table>
<thead>
<tr>
<th>Text</th>
<th>Trend Statement</th>
<th>Comparator Statistics Not Included With Trend Statement (Location in Article)</th>
<th>Point Illustrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>“There was a trend toward long-term survival in favor of GemErlo (estimated survival after 1, 2, and 5 y for GemErlo was 77%, 53%, and 25% vs 79%, 54%, and 20% for Gem, respectively).”(^2)</td>
<td>Abstract</td>
<td>$P = .61$ (Figure 2B)</td>
<td>Large $P$ value and trend statement in abstract without supporting comparator statistics</td>
</tr>
<tr>
<td>“BCL2 expression indicated a trend toward inferior outcome within GCB-like DLBCL, but not within ABC-like DLBCL (Appendix Figure A2, online only).”(^4)</td>
<td>Results, main text</td>
<td>$P = .547$ (EFS), $P = .351$ (PFS), $P = .125$ (OS) (Appendix)</td>
<td>Large $P$ value and comparator statistics to support trend statement found only in appendix</td>
</tr>
<tr>
<td>“Women with stage I–II PBL had overall survival superior to women with stage I–II systemic presentations of the same lymphoma subtype (Figure 4), except for ALCL-PBL where the same trend was seen, although not statistically significant at the 5% level.”(^5)</td>
<td>Results, main text</td>
<td>$P = .69$ (Figure 4)</td>
<td>Large $P$ value</td>
</tr>
<tr>
<td>“In unselected patients, a trend for ramucirumab survival benefit was observed in patients with HCC in the Child-Pugh 5 disease subgroup.”(^3)</td>
<td>Key points section; abstract conclusion</td>
<td>HR = 0.80; 95% CI = 0.63–1.02; $P = .06$ (results text and Figure 1A)</td>
<td>Trend statement highlighted in Key Points section and as primary study conclusion (abstract) without mention of potentially clinically relevant effect size</td>
</tr>
</tbody>
</table>

Abbreviations: ABC, activated B cell; ALCL, anaplastic large cell lymphoma; DLBCL, diffuse large B-cell lymphoma; EFS, event-free survival; GCB, germinal center B cell; Gem, gemcitabine; GemErlo, gemcitabine and erlotinib; HCC, hepatocellular carcinoma; HR, hazard ratio; PBL, primary breast lymphoma; PFS, progression-free survival; OS, overall survival.
notable differences across journals, with 14.4% (47 of 326) of articles in the JCO, 6.5% (8 of 124) of articles in the JNCI, 4.0% (6 of 149) of articles in JAMA Oncology, and 1.6% (2 of 123) of articles in Lancet Oncology using at least 1 statement describing a trend toward significance (Figure). Problematic uses of such statements are described in the Table. 2-5

There were 125 instances where a trend statement was used to describe a statistically nonsignificant result across 63 articles. Eleven statements (8.8%) provided no data by which to evaluate the magnitude of difference between groups (eg, hazard ratio). Of the 86 instances where a P value was presented, 35 (40.7%) were P ≥ .10 (15 [17.4%] were P = .10 to P < .20, 13 [15.1%] were P ≥ .20 to P < .50, and 7 [8.1%] were P ≥ .50).

Discussion | We found that trend statements are frequently used to describe statistically nonsignificant results, commonly with large P values and minimal supporting data. In addition, when P values approached statistical significance, promising clinical significance was often deemphasized to highlight the proximity of the P value to .05. This finding highlights an overemphasis on P values in the reporting of data in the oncology literature.

The biomedical literature currently has a problem with P values. In response to this problem, the American Statistical Association released a statement outlining primary P value principles.6 Among these principles were the following: (1) scientific conclusions should not be based solely on a P value threshold, (2) a P value does not measure the importance of a result, and (3) a P value does not provide a good measure of evidence regarding a hypothesis. Trend statements violate these principles. We must de-emphasize P values and shift our focus to the clinical relevance of the finding (eg, the magnitude of the result along with CI), the power of the study to address a clinically meaningful difference, and the appropriateness of the study design. This shift is additionally important to address the overinterpretation of statistically significant but clinically meaningless findings. As others have proposed, increased methodological and statistical training of scientists and clinicians may enhance the quality of data analysis, reporting, and interpretation. The oncology research community—in particular, leading oncology research journals—should take the lead in implementing the highest standards for reporting of results.

Kevin T. Nead, MD, MPhil
Mackenzie R. Wehner, MD, MPhil
Nandita Mitra, PhD

Author Affiliations: Department of Radiation Oncology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia (Nead); Department of Dermatology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia (Wehner); Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia (Mitra).

Accepted for Publication: August 7, 2018.

Corresponding Author: Nandita Mitra, PhD, Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104 (nanditam@mail.med.upenn.edu).

Published Online: October 18, 2018. doi:10.1001/jamaoncol.2018.4524

Author Contributions: Dr Mitra had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Drs Nead and Wehner are co-first authors.

Concept and design: All authors.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Nead.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: All authors.
Supervision: Mitra.

Conflict of Interest Disclosures: None reported.

Funding/Support: Dr Wehner was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases/National Institutes of Health Dermatology Research Training grant T32 AR7465 (principal investigator, Sarah E. Miller).

Role of the Funder/Sponsor: The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.


Early Adoption of Biosimilar Growth Factors in Supportive Cancer Care

In 2010, the US Food and Drug Administration established an approval pathway for biosimilar agents, biological agents with an active ingredient highly similar to the reference biological agent, to facilitate drug competition and lower costs. It is unclear whether biosimilar drugs could yield large cost savings in the United States given the previously documented market and regulatory barriers.1,3 Moreover, few studies have investigated the real-world safety and effectiveness profile of biosimilar agents compared with that of brand-name (reference) biological agents.

The colony-stimulating growth factors (CSF) filgrastim-sndz and tbo-filgrastim reduce risk of chemotherapy-induced neutropenia and were the first biosimilar agents approved in the United States. We examined the incidence of febrile neutropenia (FN), CSF-related adverse events (AEs), and drug cost among commercially insured US patients with cancer treated with chemotherapy.

Methods | This retrospective observational study analyzed administrative health claims data from a large commercially insured population. Only deidentified data were used, and the study was exempt from review by an institutional review board.