

## THE ASA PRESIDENT’S TASK FORCE STATEMENT ON STATISTICAL SIGNIFICANCE AND REPLICABILITY

BY YOAV BENJAMINI<sup>1</sup>, RICHARD D. DE VEAUX<sup>2</sup>, BRADLEY EFRON<sup>3</sup>, SCOTT EVANS<sup>4</sup>, MARK GLICKMAN<sup>5,\*</sup>, BARRY I. GRAUBARD<sup>6</sup>, XUMING HE<sup>7</sup>, XIAO-LI MENG<sup>5,†</sup>, NANCY REID<sup>8</sup>, STEPHEN M. STIGLER<sup>9</sup>, STEPHEN B. VARDEMAN<sup>10</sup>, CHRISTOPHER K. WIKLE<sup>11</sup>, TOMMY WRIGHT<sup>12</sup>, LINDA J. YOUNG<sup>13</sup> AND KAREN KAFADAR<sup>14</sup>

<sup>1</sup>*Department of Statistics and Operations Research, Tel Aviv University, [ybenja@tauex.tau.ac.il](mailto:ybenja@tauex.tau.ac.il)*

<sup>2</sup>*Department of Mathematics and Statistics, Williams College, [deveaux@williams.edu](mailto:deveaux@williams.edu)*

<sup>3</sup>*Department of Statistics and Department of Biomedical Data Sciences, Stanford University, [brad@stat.stanford.edu](mailto:brad@stat.stanford.edu)*

<sup>4</sup>*Department of Biostatistics & Bioinformatics, George Washington University, [sevans@bsc.gwu.edu](mailto:sevans@bsc.gwu.edu)*

<sup>5</sup>*Department of Statistics, Harvard University, \* [glickman@fas.harvard.edu](mailto:glickman@fas.harvard.edu); † [meng@stat.harvard.edu](mailto:meng@stat.harvard.edu)*

<sup>6</sup>*Biostatistics Branch, National Cancer Institute, [barry.graubard@nih.gov](mailto:barry.graubard@nih.gov)*

<sup>7</sup>*(Co-chair), Department of Statistics, University of Michigan, [xmhe@umich.edu](mailto:xmhe@umich.edu)*

<sup>8</sup>*Department of Statistics, University of Toronto, [reid@utstat.utoronto.ca](mailto:reid@utstat.utoronto.ca)*

<sup>9</sup>*Department of Statistics, University of Chicago, [stigler@uchicago.edu](mailto:stigler@uchicago.edu)*

<sup>10</sup>*Department of Statistics and Department of Industrial & Manufacturing Systems Engineering, Iowa State University, [vardeman@iastate.edu](mailto:vardeman@iastate.edu)*

<sup>11</sup>*Department of Statistics, University of Missouri, [wikle@missouri.edu](mailto:wikle@missouri.edu)*

<sup>12</sup>*Center for Statistical Research and Methodology, United States Bureau of the Census, [tommy.wright@census.gov](mailto:tommy.wright@census.gov)*

<sup>13</sup>*(Co-chair), Research & Development, National Agricultural Statistics Service, [linda.j.young@usda.gov](mailto:linda.j.young@usda.gov)*

<sup>14</sup>*(Ex-officio), Department of Statistics, University of Virginia, [kkafadar@virginia.edu](mailto:kkafadar@virginia.edu)*

Over the past decade, the sciences have experienced elevated concerns about replicability of study results. An important aspect of replicability is the use of statistical methods for framing conclusions. In 2019 the President of the American Statistical Association (ASA) established a task force to address concerns that a 2019 editorial in *The American Statistician* (an ASA journal) might be mistakenly interpreted as official ASA policy. (The 2019 editorial recommended eliminating the use of “ $p < 0.05$ ” and “statistically significant” in statistical analysis.) This document is the statement of the task force, and the ASA invited us to publicize it. Its purpose is two-fold: to clarify that the use of  $P$ -values and significance testing, properly applied and interpreted, are important tools that should not be abandoned, and to briefly set out some principles of sound statistical inference that may be useful to the scientific community.

$P$ -values are valid statistical measures that provide convenient conventions for communicating the uncertainty inherent in quantitative results. Indeed,  $P$ -values and significance tests are among the most studied and best understood statistical procedures in the statistics literature. They are important tools that have advanced science through their proper application.

Much of the controversy surrounding statistical significance can be dispelled through a better appreciation of uncertainty, variability, multiplicity, and replicability. The following general principles underlie the appropriate use of  $P$ -values and the reporting of statistical significance and apply more broadly to good statistical practice.

**Capturing the uncertainty associated with statistical summaries is critical.** Different measures of uncertainty can complement one another; no single measure serves all purposes. The sources of variation that the summaries address should be described in scientific articles and reports. Where possible, those sources of variation that have not been addressed should also be identified.