

*Supporting Information for:*  
**The Non-Targeted Analysis Study Reporting Tool (SRT): A Framework to Improve  
Research Transparency and Reproducibility**

Katherine T. Peter<sup>1,+,\*</sup>, Allison L. Phillips<sup>2,+,\*</sup>, Ann M. Knolhoff<sup>3</sup>, Piero Gardinali<sup>4</sup>, Carlos Manzano<sup>5,6</sup>, Kelsey E. Miller<sup>2</sup>, Manuel Pristner<sup>7</sup>, Lyne Sabourin<sup>8</sup>, Mark W. Sumarah<sup>8</sup>, Benedikt Warth<sup>7</sup>, Jon R. Sobus<sup>2</sup>

<sup>1</sup> U.S. National Institute of Standards and Technology, Charleston SC, USA 29412

<sup>2</sup> U.S. Environmental Protection Agency, RTP, NC USA 27709

<sup>3</sup> Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, College Park, MD USA 20740

<sup>4</sup> Institute of Environment and Department of Chemistry & Biochemistry, Florida International University, Miami FL, USA 33199

<sup>5</sup> Faculty of Science, University of Chile, Nunoa RM, Chile 7750000

<sup>6</sup> School of Public Health, San Diego State University, San Diego CA, USA 92182

<sup>7</sup> Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria

<sup>8</sup> Agriculture and Agri-Food Canada, London ON, Canada N5V4T3

September 15, 2021

*For submission to:*  
**Analytical Chemistry**

9 pages, 5 tables, and 6 figures

<sup>+</sup>These authors contributed equally

<sup>\*</sup>Co-corresponding author contact information

Katherine T. Peter  
Email: katherine.peter@nist.gov  
Mail: 331 Fort Johnson Road  
Charleston, SC 29412

Allison L. Phillips  
Email: phillips.allison@epa.gov  
Mail: 109 T.W. Alexander Drive E205-09  
RTP, NC 27711

**Supporting Information File 2 (Excel)** contains the following SI tables:

**Table S1.** Detailed information about the eight manuscripts reviewed during the evaluation effort.

**Table S2.** The original NTA Study Reporting Tool (SRT), as provided to the reviewers during the validation effort, prior to any edits or updates based on reviewer feedback and evaluation results.

**Table S3.** Complete (blinded) paper reviews by (a-f) category reviewers and (g-h) full SRT reviewers, with self-evaluation results included alongside external reviews.

**Table S4.** Follow-up questions asked during the validation effort, and reviewer responses. Questions about the entirety of the tool (#5-6) were only asked of full SRT reviewers.

**Supporting information File 3 (Excel)** is the interactive spreadsheet version of the SRT that is also available for download on the BP4NTA website ([www.nontargetedanalysis.org/SRT](http://www.nontargetedanalysis.org/SRT)), and contains the SRT, a plotting functionality that enables quick visual comparison of multiple reviewer scores, and information about the scoring system.

**Supporting Information File 4 (PDF)** is the interactive one-page PDF version of the SRT that is also available for download on the BP4NTA website ([www.nontargetedanalysis.org/SRT](http://www.nontargetedanalysis.org/SRT)).

### **SI File 1 Table of Contents**

**Figure S1.** Reviewer demographics and expertise.

**Figure S2.** Summary of reviewer responses to post-evaluation questionnaire.

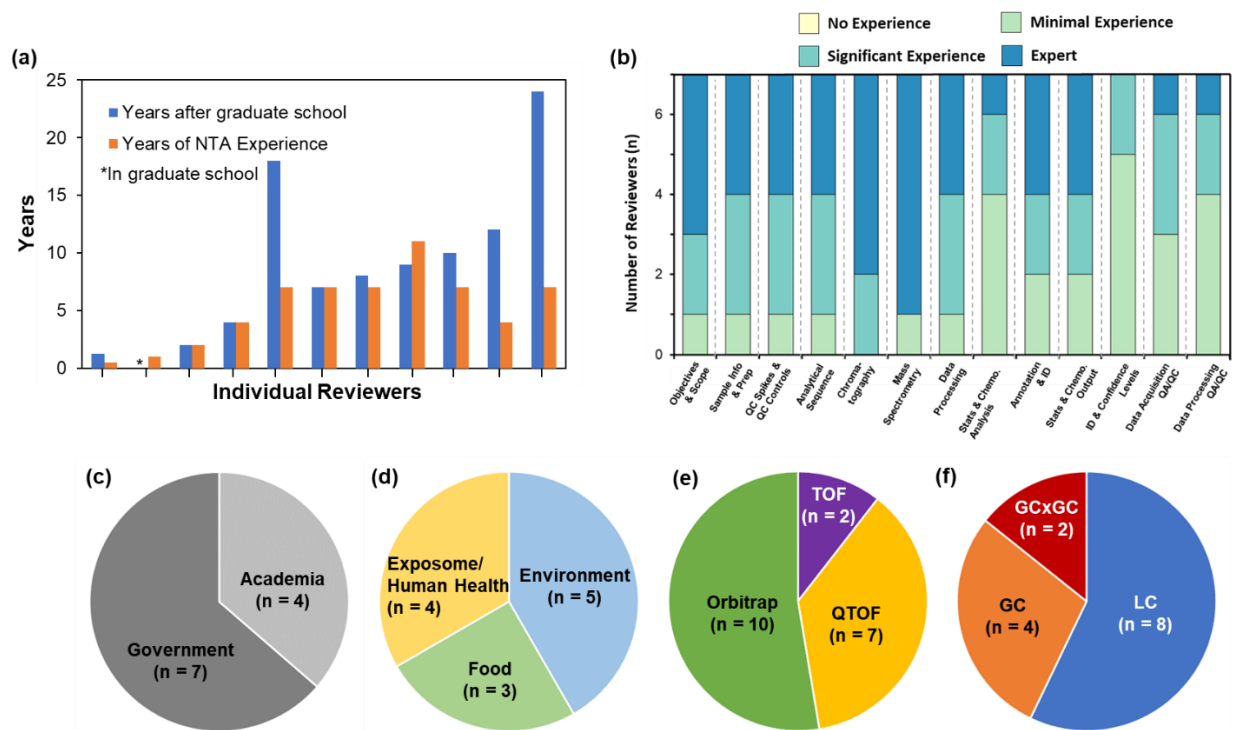
**Figure S3.** Compiled external and self-reviews for each paper.

**Figure S4.** Compiled V1, V2, and V3 scores across all papers.

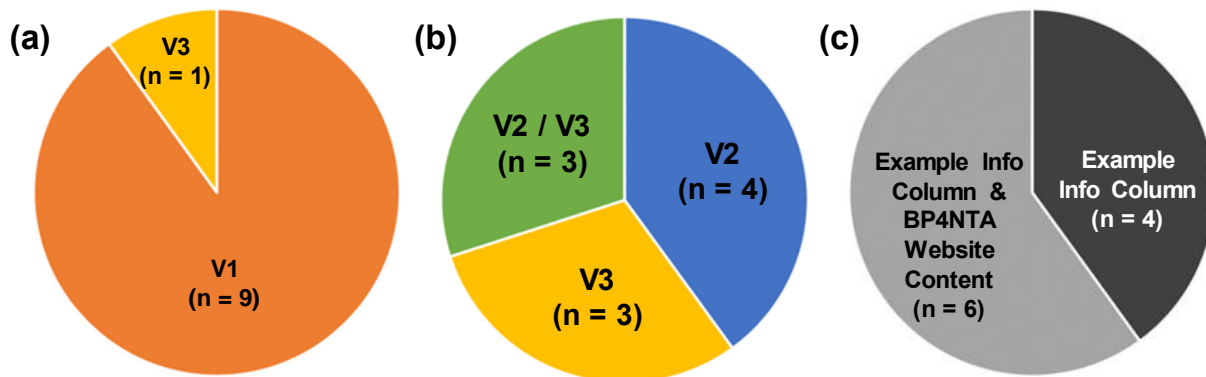
**Figure S5.** Summarized relationship between reviewer assignments of numeric (scoring system V3) and color-based (scoring system V2) scores.

**Figure S6.** Results of external reviewer reporting quality evaluations for all manuscripts, grouped by SRT sub-category, with scores translated to the final SRT scoring system.

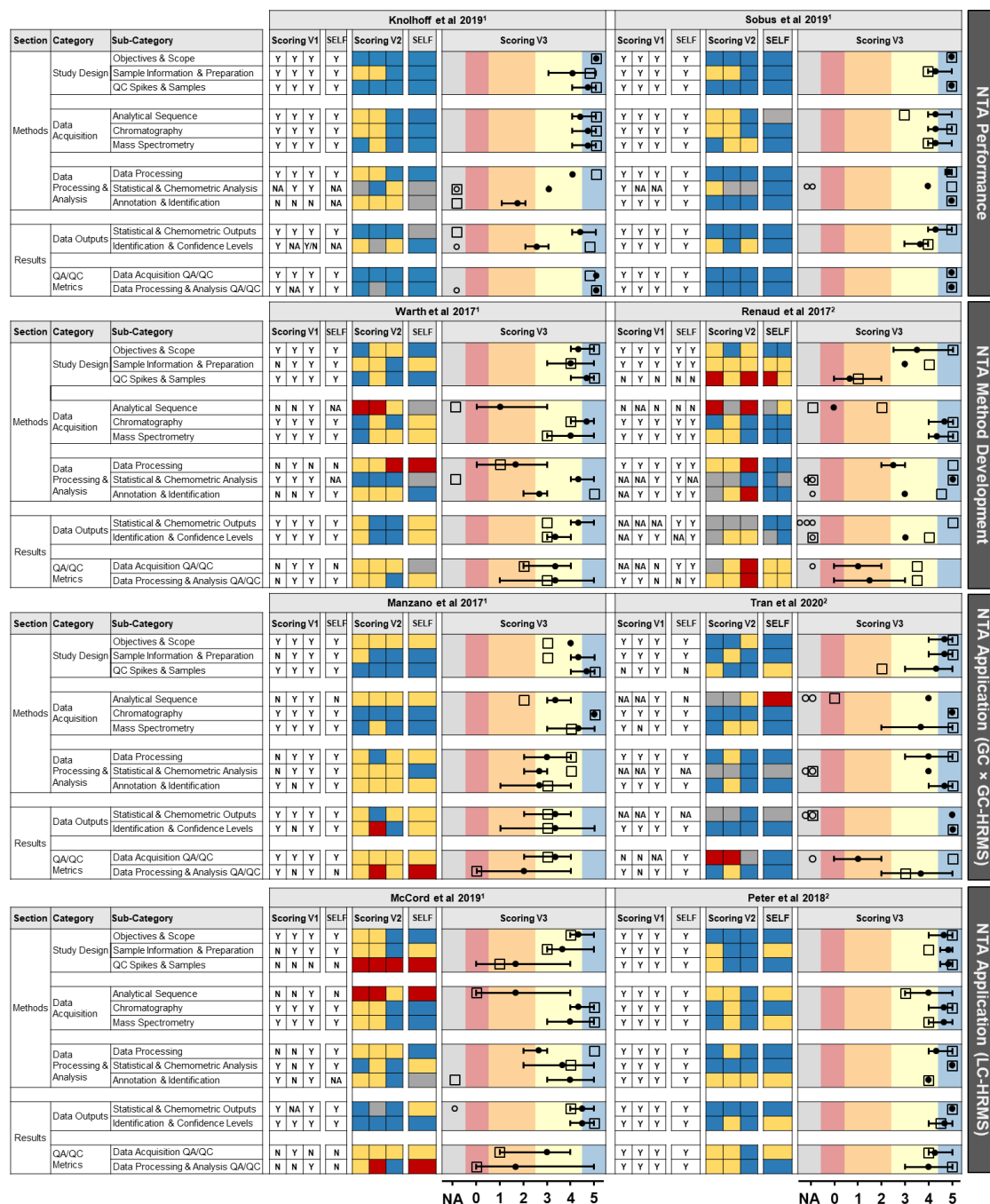
**Table S5.** Final NTA SRT scoring system, with a description of each score level and representative examples for selected sub-categories.



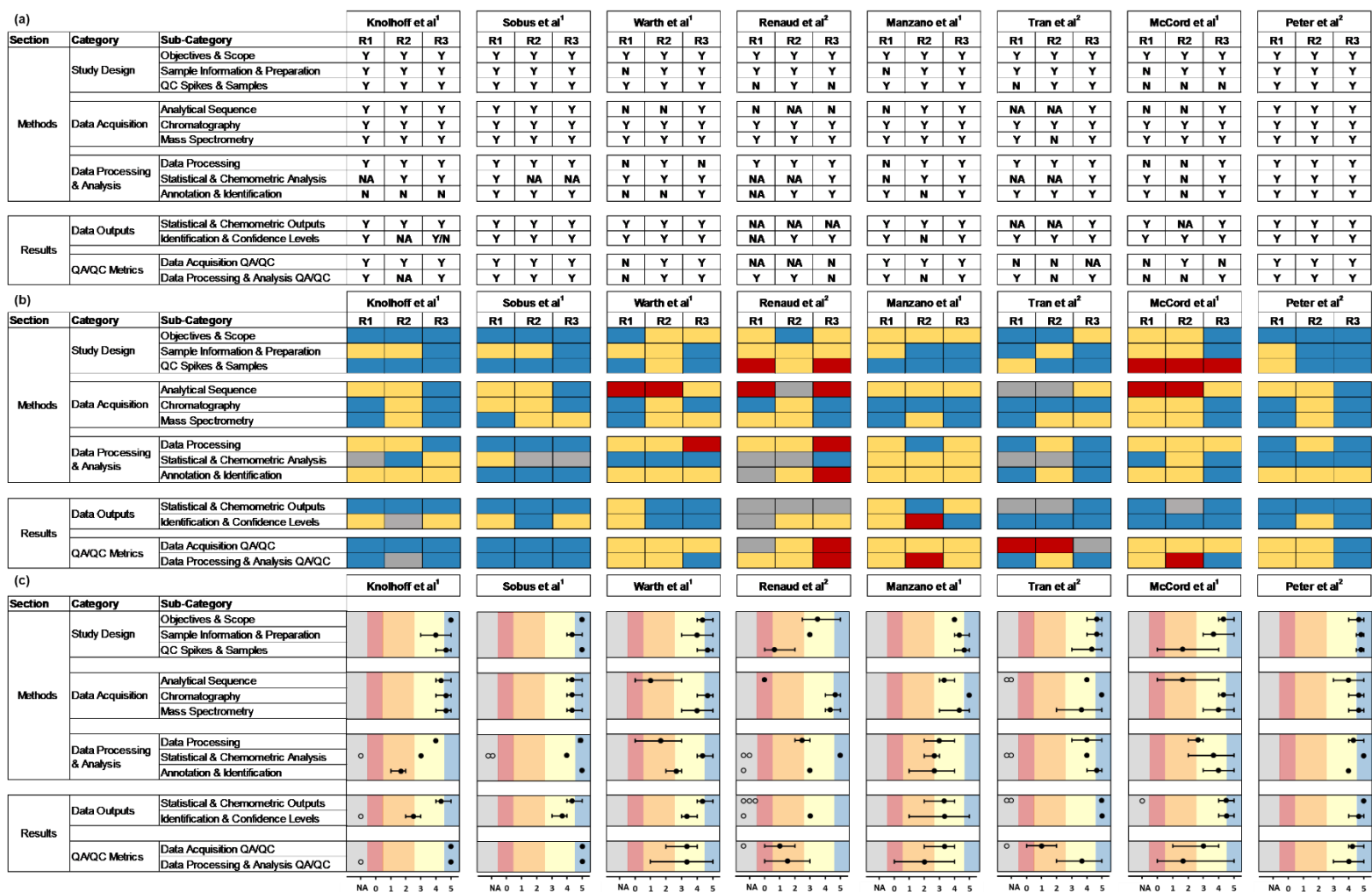
**Figure S1.** Reviewer (n=11) demographics, including (a) years since completing graduate school (if applicable; one reviewer was in graduate school) and years of experience performing NTA research; (b) reviewer (n=7, with two entries reflecting an aggregation of a graduate student/advisor pair) experience level for each SRT category; (c) reviewer employment sector (according to their employment at the time of the review); (d) reviewer research area (note that one reviewer identified with two of the three categories); and (e)-(f) reviewer experience with HRMS and chromatography platforms, respectively.



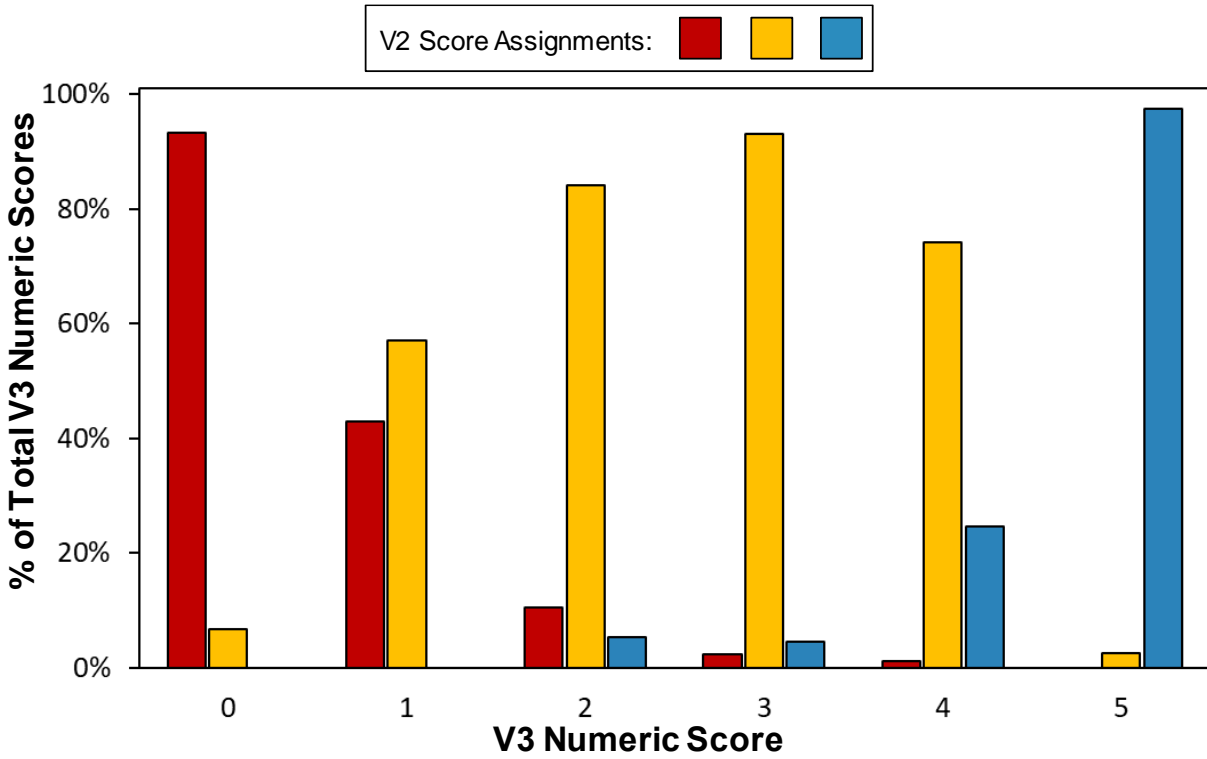
**Figure S2.** Summary of reviewer responses to questions about (a) least preferred scoring system (Question #2), (b) most preferred scoring system (Question #1), and (c) use of “Example Information to Report” column and/or BP4NTA website reference content during the review (Question #3). Scoring systems were: V1 – *Yes/No/NA*, V2 – Color-based scores (*Gray/Blue/Yellow/Red*), and V3 – Numeric scores (*NA, 0-5*). See **Table S4** for written responses to questions.



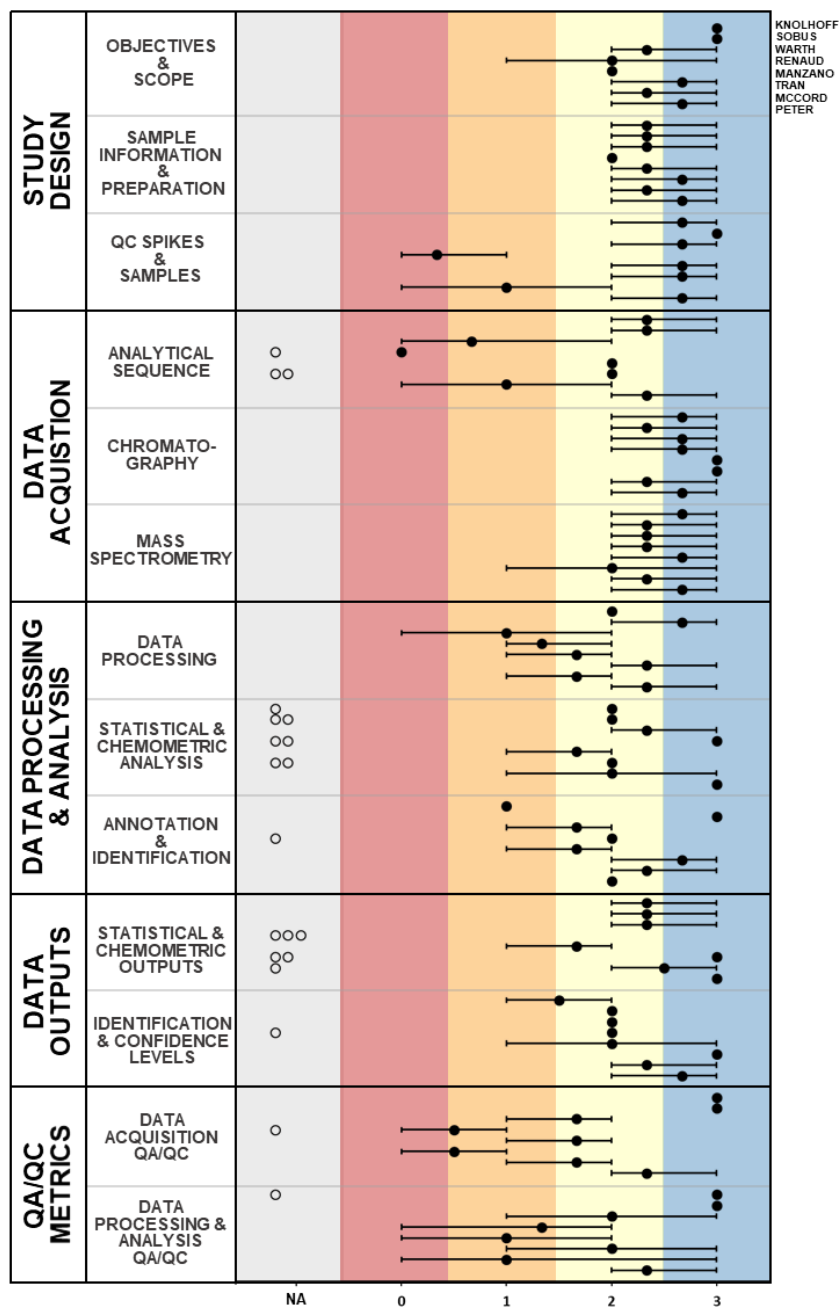
**Figure S3.** Compiled external and self-reviews for all papers. Results from all three original scoring versions are shown side-by-side, with self-evaluation results provided alongside external reviewer scores. In Scoring V3, filled circles show the average external reviewer score, error bars show the range of external reviewer scores, open circles show external NA scores, and open squares show self-evaluation scores. Coloration behind Scoring V3 was applied after the evaluation process (i.e., reviewers did not assign numeric scores with paired colors) by mapping the 6-level V3 scores (NA, 0-5) onto the final 5-level color scheme: *gray* = NA; *red* = 0 – <0.5; *orange* = 0.5 – <2.5; *yellow* = 2.5 – <4.5; *blue* = 4.5 – 5. <sup>1</sup>Manuscripts reviewed by SRT category reviewers; <sup>2</sup>Manuscripts reviewed by full SRT reviewers.



**Figure S4.** Compiled (a) V1, (b) V2, and (c) V3 scores across all papers. In (c), filled data points show the average external reviewer score, error bars show the range of reviewer scores, and open circles indicate external reviewer NA scores. Coloration in (c) was applied after the review process (i.e., reviewers did not assign numeric scores with paired colors during the review process) by mapping the 6-level V3 scores (NA, 0-5) onto the final 5-level color scheme: *gray* = NA; *red* = 0 – <0.5; *orange* = 0.5 – <2.5; *yellow* = 2.5 – <4.5; *blue* = 4.5 – 5. <sup>1</sup>Manuscripts reviewed by SRT category reviewers; <sup>2</sup>Manuscripts reviewed by full SRT reviewers.



**Figure S5.** Summarized relationship between reviewer assignments of numeric (scoring system V3) and color-based (scoring system V2) scores ( $n = 290$  total;  $0$ ,  $n = 15$ ;  $1$ ,  $n = 7$ ;  $2$ ,  $n = 19$ ;  $3$ ,  $n = 43$ ;  $4$ ,  $n = 89$ ;  $5$ ,  $n = 117$ ). Numeric scores were rounded up to the nearest whole number (e.g., scores of 2.5 are grouped with 3) and NA scores were excluded for creation of the plot.



**Figure S6.** Results of external reviewer reporting quality evaluations for all manuscripts, grouped by SRT sub-category, with scores translated to the final SRT scoring system. Assigned scores were translated as follows, and average scores were recalculated after translation of individual reviewer scores: NA  $\rightarrow$  NA (gray);  $<1 \rightarrow 0$  (red);  $1 - <3 \rightarrow 1$  (orange);  $3 - <5 \rightarrow 2$  (yellow);  $5 \rightarrow 3$  (blue). All publication scores are listed in the same order (from top to bottom, as noted at the top right of the figure) in each sub-category, and are grouped according to study type as in Figure 2. Filled dots represent the average reviewer score, error bars represent the range of reviewer scores, and open circles are shown for NA scores. Coloration was applied after the evaluation process (i.e., reviewers did not assign numeric scores with paired colors).



**Table S5.** The final NTA SRT scoring system, based on feedback and results of the evaluation effort, with a description of each score level and representative examples for selected sub-categories. The hybrid color-coded/numerical hybrid system has 5 levels: 0 – red, 1 – orange, 2 – yellow, 3 – blue; NA – gray. Assessment of ‘relevant reporting elements’ for a given study requires researcher/reviewer expertise.

| Score     | Description  | Example 1 – Analytical Sequence  | Example 2 – Statistical & Chemometric Analysis  | Example 3 – Data Processing & Analysis QA/QC   |
|-----------|--|--|---|--|
| <b>0</b>  | <b>No elements of relevant reporting are present.</b>                                      | No details are provided regarding analysis order and batch information. Based on reporting, the experiment could not be replicated.  | Statistical analyses were performed, but no information was provided about software, methods, assumptions, and thresholds.  | No details are provided about the quality, boundary, accuracy, and precision of the data processing and analysis method(s). Based on reporting, the overall performance of the processing and analysis method(s) is unclear.   |
| <b>1</b>  | <b>Some elements of relevant reporting are present, but major improvements are needed.</b> | Some details are provided regarding analysis order and batch information, but major details that would assist in the interpretation of the results are lacking (e.g., the authors reported that blanks were analyzed, but no information about frequency/injection or sample order within the acquisition sequence was provided).                      | Statistical analyses were performed, but gaps in method details limit method reproducibility.   | Some details are provided about the quality, boundary, accuracy, and precision of the data processing and analysis method(s), but major gaps limit understanding of overall performance (e.g., impacts of method choices on observed chemical space are discussed, but true/false positive rates, workflow QC checks, and reproducibility of identification are not reported). |
| <b>2</b>  | <b>Most elements of relevant reporting are present, but minor improvements are needed.</b> | Most analysis order and batch information is provided, but some details are missing that could assist interpretation of the results (e.g., analysis order was reported, including sample randomization and the frequency/order of external reference standards and method blanks; however, information about multiple analytical batches was unclear). | Statistical analyses were performed, and the majority of method details are provided (e.g., the authors clearly reported the software, methods, and assumptions, but minor details regarding method thresholds are missing that would improve the reader’s ability to interpret the results). | Detailed information is provided about the quality, boundary, accuracy, and precision of the data processing and analysis method(s), but minor details that could improve the reader’s ability to interpret the impact on results are excluded (e.g., the authors do not discuss possible sources of variability in the outcomes of their identification workflow).            |
| <b>3</b>  | <b>All elements of relevant reporting are present.</b>                                     | All details regarding sample randomization, replicate injections, the inclusion of blanks and QC samples in the acquisition sequence, and information about analytical batches are provided so that the experiment could be reproduced by an outside researcher.   | Statistical analyses were performed, and all method details are reported in sufficient detail to allow a reader to reproduce the method and clearly understand the impact of method choices on the analysis results.  | All necessary detailed information about the quality, boundary, accuracy, and precision of the data processing and analysis method(s) are provided and the implications are discussed.   |
| <b>NA</b> | <b>Reporting not relevant to the study.</b>  | Analytical sequence and batch information should be reported for all studies; however, it is less critical for studies that do not rely on comparisons of measurement data across samples.   | Statistical analyses were not performed (e.g., the study is solely focused on chemical annotation & identification).  | Reporting the quality, boundary, accuracy, and precision of the data processing & analysis methods is informative for all studies.   |