https://doi.org/10.1038/s41560-024-01663-y

The path to accurate reporting

Inconsistent reporting on energy materials and devices in research papers underscores the need for standardized protocols and greater transparency. Collaborative benchmarking initiatives are paving the way for more reliable and reproducible results.

Cutting-edge materials and devices are poised to make a substantial impact on the development of next-generation clean energy technologies. However, a noticeable gap often exists between optimistic claims made in academic publications about new materials and devices, and the actual advancements. This discrepancy, particularly in light of the diverse literature, underscores the critical need for research reproducibility to verify findings.

In recent years, there has been considerable emphasis on the importance of reproducible research in energy materials and devices. In the realm of publishing, a fundamental principle is that others should be able to replicate and extend the authors' findings. To support this principle, journals have implemented various checklists for reporting energy technologies. According to the guidelines of Nature Portfolio journals, authors are required to make information about materials, data, code, and associated protocols promptly available to readers without undue qualifications.

Despite these efforts, reproducibility in energy research continues to be a challenge. A key issue is the inherent complexity of many energy materials and devices employed in energy technologies; for example, they are often composed of multiple components and numerous variables are involved in material processing, assembly, and testing protocols. Such complexities greatly hinder the consistent reproduction of results.

For instance, rechargeable batteries typically consist of positive and negative electroactive materials, electrolytes, and packaging materials. Depending on the design, they may also include current collectors, separators, conductive agents, and other additives. The properties of each component are influenced by many factors such as uniformity, thickness, and dryness for electrode materials, as well as chemical purity, moisture, and container type for electrolytes. Additionally, variables like temperature and pressure significantly impact battery cell assembly, while factors such as current density and depth of charge/discharge cycles affect performance. The interplay of these numerous interdependent variables poses a considerable challenge to reproducing battery performance results.

In an ideal world, everything would be disclosed; however, in practice, it is useful to focus on the factors that most significantly affect performance of energy devices. For example, *Nature Energy* has published the pioneering efforts of the US Battery500 Consortium in advancing high-energy, long-cycling lithium metal batteries, one of the most promising next-generation battery technologies[1]. This work identified key factors such as cathode loading, electrolyte amount, and lithium foil thickness, which impact the cell-level cycle life. These factors have since become standard components in research papers, underscoring their critical role in battery performance and the importance of detailed reporting.

Another approach to fostering reproducibility in the performance of energy devices is interlaboratory studies. In an Analysis in this issue of *Nature Energy*, twenty-one international research groups active in the field of solid-state batteries — another promising next-generation technology — report on benchmarking the reproducibility of these battery cells through interlaboratory coordination. Their approach involved supplying all groups with the same battery materials but allowing each to employ their own cell assembly protocols while adhering to the same specific electrochemical testing protocols. By gathering a large dataset from the study, they were able to evaluate correlations between cell performance metrics and various parameters in electrode/electrolyte preparation, cell assembly, and electrochemical experiments.

Unsurprisingly, this interlaboratory effort coordinated by Nella Vargas-Barbosa found huge differences in cycling behaviours, capacities, and voltages between cells prepared by different groups. More importantly, the study provides valuable insights and establishes general guidelines for researchers in the field. For instance, the initial open circuit voltage was identified as a strong indicator of high-performing cells. Additionally, the compression times used during cell assembly, along with the rate of pressure application and release at each cycling step, were found to influence the microstructure, thus affecting cell performance. These factors, largely overlooked in the literature, are now recommended for disclosure in research papers on solid-state batteries.

Vargas-Barbosa and colleagues also make two additional recommendations. First, publications that report cycling data should include not only the data from a single battery but also the average and error range from triplicate measurements. Second, cells that did not perform as expected should also be documented to allow for better classification of the reported cell data.

In terms of reporting statistics, three samples are generally insufficient to achieve statistical significance, although at *Nature Energy* we appreciate the challenges involved in preparing and measuring larger sample sizes. We therefore discourage deriving statistics from such a small sample unless there is a clear scientific justification for doing so. In line with practices in *Nature Energy* and to ensure consistency[2], we advocate plotting the independent data points when the sample size is small, allowing readers to interpret the data themselves rather than presenting possibly misleading error bars and drawing statistical inferences.

Regarding the reporting of failures, Vargas-Barbosa and colleagues found that over 40% of the cells attempted in their study did not function properly during preparation or cycling, highlighting the challenges in battery assembly. Understandably, it is common practice for authors to report only their best results in publications, making it difficult to implement the recommendation to report failures. However, documenting failures can be valuable, particularly when reporting the best available results, because it facilitates the detection of mistakes and helps advance the field by providing a more comprehensive understanding of various issues encountered.

This endeavour led by Vargas-Barbosa and colleagues marks one of the latest efforts to establish new standards for transparency and reproducibility in battery research. Similar, interlaboratory studies aimed at benchmarking other energy technologies have been conducted; for instance, Konrad Ehelebe and colleagues[3] compared fuel cell catalysts using gas diffusion electrodes prepared in different setups and laboratories before proposing protocols for comparable half-cell testing. We hope that initiatives like these will lead to more standardized reporting of energy materials and devices in publications. Meanwhile, we welcome participation from relevant communities to engage in discussions that foster greater consensus. By collectively addressing these challenges and establishing benchmark assessments, we can significantly enhance the reliability and reproducibility of performance metrics for next-generation energy technologies.

Published online: 22 October 2024

References

1. Liu, J. *Nat Energy* **4**, 180–186 (2019).
2. Vaux, D. *Nature* **492**, 180–181 (2012).
3. Ehelebe, K. et al. *ACS Energy Lett.* **7**, 816–826 (2022).