

nature medicine

Structural integrity

It is highly improbable that a chemistry journal would agree to publish the synthesis of a new compound without the structure of the reagents, intermediates or final product. Why then should it be deemed acceptable in biology to report the identification of a new gene, without its accession number or sequence, a drug without its chemical structure, or a new algorithm for analysis of 'omics' data without the code?

And yet a recent perusal of biology journals suggests that the practice of omitting key information, which impairs the critical evaluation and replication of scientific findings, may be quite commonplace.

In March 2004, a prominent cancer journal published an article on the use of a new kinase inhibitor, obtained from a pharmaceutical company, to treat solid and hematological malignancies. The structure of the molecule was not disclosed. In September, another high-profile journal published the identification of a regulator of mitosis without its nucleotide sequence or an accession number, suggesting that it had not been deposited in a public repository. That same month, the aforementioned cancer journal published results on a new drug treatment for brain cancer, once again without including the structure of the compound used. If one follows the trail of references provided, we find a prior report using the drug in a neuroscience journal, again without its structural identity, and a pharmaceutical company's patent, listing a family of related compounds, without specification of the drug used in the study.

Although oversights do occur, it is ultimately the journal's responsibility to ensure that the reader has all the necessary information to reproduce the published work. At *Nature Medicine*, our guide to authors indicates that the publication of articles including new genes, proteins or crystallographic structures is contingent on deposition of the accession number and/or structural coordinates in a publicly accessible database (http://www.nature.com/nm/about/ed_policies/#materials). Although not explicitly referred to, the reporting requirements extend to the chemical structures of drugs, as well as sequences of oligonucleotides used in antisense strategies and RNAi—a partial list, and subject to amendment as the need for disclosure expands with the advances of science.

The desire to keep unpublished data out of the public domain is evident at virtually any scientific conference these days—whereas in the past researchers spoke more freely about their work, now a litany of published results counters the proliferation of video cameras at posters and talks. The consequent limited discussion of unpublished data—prompted by a fear of being scooped—and its detrimental effect on the collaborative nature of science are

indeed regrettable. But the climate of secrecy surrounding new drug candidates is prompted by other concerns.

Corporate motivations for restricting full disclosure of a compound are numerous. They may include pending patent applications and regulatory approval, or may simply reflect the desire for a competitive edge. Even in the case of patented, but unapproved drugs, structures may be difficult to identify because of the use of alternate names, or, as in the case of the brain cancer inhibitor cited above, because of reference to a group of related compounds without explicit structural description and characterization of any single one.

Admittedly, publicizing information on a parent structure or active compound may provide valuable information to a competing company. And public disclosure of the structure of a drug in clinical trial has inherent risks. Its use by the research community can jeopardize completion of the trial or regulatory approval of an investigational new drug application as adverse events that might affect patient safety, including those in preclinical studies, must be reported.

Moreover, regulatory agencies do not necessarily require full disclosure of chemical structures in the early stages of drug development. A fact evident in the recent announcement by a pharmaceutical company of the chemical identity of a new treatment for Alzheimer disease only subsequent to completion of a phase 2 trial.

But from our standpoint, if a publication using the drug is warranted, the structure must be made available to fellow researchers and referees—at the time of submission of the manuscript. In a similar vein, we require the submission of the raw data for microarray papers, and in a standardized format. The intent is to facilitate evaluation of the results and the authors' analysis. And equally it is an extension of *Nature Medicine's* policy on ensuring the availability of data and materials to the referees prior to publication, and to the reader upon publication. Specifically, this policy requires that authors provide reagents and methods upon request. In that context, the usefulness of distributing a compound without structural information is questionable.

Both referees and editors must consider the effects of these omissions on the scientific enterprise and establish and enforce editorial guidelines that continue to uphold integrity in scientific publishing and ensure the utility of the printed work. If science cannot be reproduced by independent researchers because of a lack of adequate information, and if the quality of the results, as well as the specificity, efficacy and toxicity of a drug cannot be assessed, the importance of the findings may not be put into proper perspective. Of what value then is the publication to the research community?