Limitations are not properly acknowledged in the scientific literature

John P.A. Ioannidisa,b,c

a Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece
b Biomedical Research Institute, Foundation for Research and Technology Hellas, Ioannina 45110, Greece
c Tufts University School of Medicine, Boston, MA 02111, USA

Accepted 8 September 2006

Abstract

Limitations are important to understand for placing research findings in context, interpreting the validity of the scientific work, and ascribing a credibility level to the conclusions of published research. This goes beyond listing the magnitude and direction of random and systematic errors and validity problems. Acknowledgment of limitations requires an interpretation of the meaning and influence of errors and validity problems on the published findings. An examination of the full-text files of the first 50 articles published in 2005 in the six most-cited research journals and in two recently launched leading open-access journals showed that only 67 articles (17%) used at least one word denoting limitations in the context of the presented scientific work. Only four articles (1%) used the word limitation in their abstract; none referred to limitations of the present work that materially affected conclusions. Only five articles had a separate section on limitations. Conversely, 243 articles (61%) used words detected by the roots error, valid, bias, reproducib, or false and 289 articles (72%) used words with the root important. Among the 25 top-cited journals’ instructions to the authors and editorial policies, only one encourages discussion of limitations; importance, novelty, and lack of error are typically encouraged. Limitations should be better covered and discussed in research articles. To facilitate this, journals should give better guidance and promote the discussion of limitations. Otherwise, we are facing an important loss of context for the scientific literature. © 2007 Elsevier Inc. All rights reserved.

Keywords: Limitations; Research; Journals; Bias; Authors; Error; Discussion; Instructions to authors

1. Introduction

All research work unavoidably has some limitations. Even the most important breakthroughs are unlikely to be devoid of them. Knowledge and discussion of limitations are essential for genuine scientific progress: they are useful for understanding a research finding, translating the importance of the potential errors involved, placing the current work in context, and ascribing a credibility level to it [1]. Limitations are also likely to reveal how the current research work may be improved in future experiments and what caveats should be considered in trying to incorporate this new information in the evolving body of scientific evidence. However, are limitations properly acknowledged in the scientific literature?

The question is difficult to answer. Researchers may have different ways to discuss and address caveats in their work. For example, a perceived drawback may lead to further experiments that clarify the remaining issues. In particular, empirical research papers with statistical hypothesis testing, may indeed present some measure of the errors and the uncertainty surrounding the measurements and inferences thereof. However, this does not necessarily reflect a specific acknowledgment of limitations by any means.

Claiming limitations is unavoidably a subjective process to a considerable extent. It is not simply an issue of listing the magnitude and direction of various errors, random or systematic, that have been introduced in the measurements; or the problems in the theoretical concepts and methods used in the research; or the generalizability of the findings. Limitations go beyond a proper list of methods, errors, and validity. They require an overall view, appraisal, and interpretation: do these problems with errors, methods, and validity eventually matter, and, if so, to what extent? At the end of the day, this is a crucial link to interpret the credibility of the published research and make sense of it. For measurement errors, larger errors may be more likely to be perceived as limitations, but what is a large enough error that would affect the conclusions? For external validity, such subjective appraisal is even more necessary: if a research project is limited to a very narrow focus, is this a limitation or strength? Someone needs to interpret this dilemma, ideally both sides of it. Finally, for research work
that has no measurements, for example, mathematical or theoretical constructs, it is still important to know what the perceived weaknesses are.

One may argue that scientific work should be objective; subjective statements have no place in scientific writing. I could strongly sympathize with this view, which would mean that research articles should have no Discussion section, just transparent Methods and full Results. But then, not only subjective interpretations of limitations would be eliminated from scientific writing, but also subjective claims of importance, relevance and (nonstatistical) significance would be dispensed. All these “positive” aspects of research are eventually subjective interpretations based on—hopefully—objective data and constructs. The same applies to limitations. If we accept subjective claims of importance in scientific writing, we should accept and encourage also statements of limitations.

2. Limitations in the high-profile research literature

I searched electronically the full-text (except references) of the first 50 original research articles published in 2005 in each of the six scientific journals that receive the highest total number of citations [2]. These journals include multidisciplinary publications (Science, Nature, Proceedings of the National Academy of Sciences [PNAS]) and leading journals in biochemistry (Journal of Biological Chemistry), physics (Physical Review Letters), and chemistry (Journal of the American Chemical Society). I also evaluated the first 50 full research papers published in the same year (2005) in two recently-launched open-access journals, PLoS Biology and PLoS Medicine. Open-access journals champion free dissemination of scientific information, they are mainly electronically published and they have no space constraints [3,4].

Ninety-four articles (24%) used at least one of the words identified by the roots limitation, caveat, or cautio in any context (Fig. 1). However, several appearances of these words did not pertain to limitations of the presented scientific work. After excluding technical terms with entirely different meaning and appearances where the authors discussed strictly the limitations of previous work, only 67 articles (17%) used at least one of these words in the context of their own presented scientific work; the rates were somewhat higher for the two open-access electronic journals (PLoS Biology 26%, PLoS Medicine 42%).

Use of these words did not necessarily mean that the authors accepted the presence of weaknesses in their work. In fact, in several instances, the authors outright denied or downplayed the presence of limitations or used this wording only to reinforce the importance of their work. A typical example is PNAS: in all four articles where the word limitation was used referring to the presented work, the authors used it in denial, that is, they claimed that their work and methods “overcome” (twice), “avoid”, or “circumvent” limitations.

The words shortcoming(s) or drawback(s) appeared in only 11 articles. In nine of the 11 cases they focused only on shortcomings and drawbacks of previous work that were supposedly addressed and solved by the current research. The word weakness was also used almost ubiquitously with a different or technical meaning (e.g., “motor weakness”), whenever it appeared: only one article used it to convey a sense of caveats in the presented work.

3. Limitations in abstracts, section headings, first sentence in paragraph

Only four of the 400 articles used a word on limitations in their abstract. One of them was a technical term (“growth limitation”) in a Journal of Biological Chemistry article. Another was an opening sentence in a Nature article that set the setting about the superiority of the presented nano-device methodology: “devices have been investigated… that could overcome the physical and economic limitations of current semiconductor devices”. A PNAS article claimed that there is “an important existing

---

Fig. 1. Proportion of articles with at least one appearance of at least one word denoting limitations. The 50 eligible articles for each journal were selected in consecutive order from the online journal sites, focusing on full research papers, and excluding news, reviews, editorials, comments, and brief communications. Words were identified in automated searches of pdf files using the roots limitation, caveat, cautio. Each root may capture several words (e.g., cautio would capture caution, cautious, cautionary). The bars also show separately the proportion of articles with mention of limitations on the present work (excluding technical terms, different meaning, alluding to other work only).
<table>
<thead>
<tr>
<th>Journal</th>
<th>Comments on publication preferences/discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal of Biological Chemistry</td>
<td>The primary criteria for judging the acceptability of a manuscript are its novelty and scientific importance</td>
</tr>
<tr>
<td>Nature</td>
<td>The criteria for publication of scientific papers are that they report original scientific research, are of outstanding scientific importance, and reach a conclusion of interest to an interdisciplinary readership</td>
</tr>
<tr>
<td>PNAS</td>
<td>Research Reports describe the results of original research of exceptional importance</td>
</tr>
<tr>
<td>Science</td>
<td>Priority is given to papers that reveal novel concepts of broad interest</td>
</tr>
<tr>
<td>Journal of the American Chemical Society</td>
<td>Articles of high scientific quality and originality that are of interest to the wide and diverse contemporary readership JACS will be given priority for publication</td>
</tr>
<tr>
<td>Physical Review Letters</td>
<td>The paper must satisfy criteria of validity, importance, and broad interest. The work must be sound, free of detectable error, and presented in reasonable detail. The results must be new...the experiments must be demonstrated to be novel and feasible/Each paper should present as a complete discussion as possible within the constraints of a short communication</td>
</tr>
<tr>
<td>Physical Reviews B</td>
<td>It is the policy of the American Physical Society that the Physical Reviews accept for publication those manuscripts that significantly advance physics and have been found to be scientifically sound, important to the field, and in satisfactory form...Papers must contain new results in physics. Confirmation of previously published results of unusual importance can be considered as new, as can significant null results. Papers advancing new theoretical views on fundamental principles or theories must contain convincing arguments that the new predictions and interpretations are distinguishable from existing knowledge, at least in principle, and do not contradict established experimental results.</td>
</tr>
<tr>
<td>NEJM</td>
<td>Original Articles are scientific reports of the results of original clinical research</td>
</tr>
<tr>
<td>Astrophysical Journal</td>
<td>Papers published in the Astrophysical Journal present the results of significant original research...the results presented must constitute significant new [sic] and research that is directly relevant to astrophysical applications</td>
</tr>
<tr>
<td>Journal of Chemical Physics</td>
<td>Articles are reports of original research of significance to the chemical physics community</td>
</tr>
<tr>
<td>Cell</td>
<td>Cell publishes reports of novel results in any area of experimental biology. The work should be not only of unusual significance within its field but also of interest to researchers outside the immediate area. The basic criterion for considering papers is whether the results contain significant novel insights into, or raise provocative questions and hypotheses regarding, an interesting biological question.</td>
</tr>
<tr>
<td>Lancet</td>
<td>The Lancet gives priority to reports of original research that are likely to change clinical practice or thinking about a disease</td>
</tr>
<tr>
<td>Circulation</td>
<td>Provisional or final acceptance is based on originality, scientific content, and topical balance of the journal.</td>
</tr>
<tr>
<td>Applied Physics Letters</td>
<td>...a weekly journal featuring concise, up-to-date reports on significant new findings in applied physics. Emphasizing rapid dissemination of key data and new physical insights...</td>
</tr>
<tr>
<td>Journal of Immunology</td>
<td>Major criteria for acceptance are quality, originality, clarity, and conciseness</td>
</tr>
<tr>
<td>Journal of Geophysical Research</td>
<td>The decision to accept a contribution for publication is made by the journal editor solely on the basis of suitability of subject matter to the focus of the journal, originality of the contribution, and scientific merit.</td>
</tr>
<tr>
<td>Cancer Research</td>
<td>...only those that report results of novel, timely, and significant research and meet high standards of scientific merit are accepted for publication/Discussion. The data should be interpreted concisely without repeating material already presented in the Results section. Speculation is permissible, but it must be well founded, and discussion of the wider implications of the findings is encouraged.</td>
</tr>
<tr>
<td>Blood</td>
<td>Acceptance of papers is based on the originality of the observation or investigation, the quality of the work described, the clarity of presentation, and the relevance to our readership... Acceptance of Regular Articles is based on the originality, definitiveness, and importance of the findings to the field of hematology.</td>
</tr>
<tr>
<td>Biochemistry-US</td>
<td>Preference will be accorded to manuscripts that develop new concepts or experimental approaches, particularly in the advancing areas of biochemical science... Publication of preliminary or inconclusive results is discouraged/The purpose of the discussion is to interpret the results and to relate them to existing knowledge in the field. In general, observe the utmost brevity consistent with clarity and avoid personal polemics.</td>
</tr>
<tr>
<td>Journal of Neuroscience</td>
<td>The Journal publishes papers on a broad range of topics of general interest to those working on the nervous system.../The discussion section (not to exceed 1,500 words) should be as concise as possible and should include a brief statement of the principal findings, a discussion of the validity of the observations, a discussion of the findings in light of other published work dealing with the same or closely related subjects, and a statement of the possible significance of the work. Extensive discussion of the literature is discouraged.</td>
</tr>
<tr>
<td>JAMA</td>
<td>Manuscripts should meet the following criteria: material is original; writing is clear; study methods are appropriate; data are valid; conclusions are reasonable and supported by the data; information is important; and topic has general medical interest. From these basic criteria, the editors assess a paper’s eligibility for publication/...a comment section placing the results in context with the published literature and addressing study limitations</td>
</tr>
<tr>
<td>Journal of Applied Physics</td>
<td>Journal of Applied Physics is the American Institute of Physics’ (AIP) archival journal for significant new results in applied physics...We seek to publish papers that contain substantial advancement of established knowledge or that report significant new developments in the field...Papers must contain new results to be published.</td>
</tr>
</tbody>
</table>
limitation’’ that was resolved with the presented work. Only a single PLoS Medicine article acknowledged limitations in the presented work, although these were not considered to affect the main conclusions (‘‘despite some limitations’’).

Devoting a section to study limitations was an extremely rare occurrence: only three PLoS Medicine articles, one PLoS Biology article, and one Science article had such a clearly identifiable section in their Discussion. Another two PLoS Medicine articles had a similar title in a distinct section of the accompanying ‘‘Patient Summary’’ (in the same pdf as the article).

Seventeen articles used limitation(s) or a similar word in the first sentence of a paragraph. Sixteen of these pertained to the two open-access journals (PLoS Medicine n = 11, PLoS Biology n = 5) and one appeared in Science.

4. Errors and importance vs. limitations

The sparse allusion to limitations does not mean that these articles did not discuss errors in their scientific work. In fact, 243 of 400 articles (61%) used at least once one or more words detected by the electronic roots error, valid, bias, reproducib, or false. However, the vast majority of these appearances had a neutral context, for example, measures of errors or error bars were given without further comment, or estimates were given of validity or validation procedures. However, it was not mentioned whether these problems were large or small and thus whether they would be limitations or strengths of the research work. Most (although not all) scientists were likely to provide some estimates of errors, both random and systematic, in their work. However, they were unlikely to take the extra step to focus and discuss these potential errors or to focus on broader study design and context issues as specifically posing constraints to their inferences.

There was an imbalance of emphasis between limitations and importance. In automated searches for the root important (e.g., important, importance, importantly) showed that 72% of these articles (289/400) had used specifically such words. Only two of the 400 articles used the word ‘‘unimportant’’. One appearance was in a footnote claiming that an effect was unimportant and thus not worth discussing there (Journal of the American Chemical Society) and another appearance was made also for technical reasons in the Methods section to convey that selection of different cut-offs for the analyses did not make a difference (PLoS Biology). Limitations and importance are not mutually exclusive. However, these data suggest that authors, peer reviewers, or editors may believe that importance should be highlighted, while limitations are a nuisance.

5. Weaknesses of the empirical survey

This survey has limitations. Capturing all the language that touches on limitations is impractical. Thus, I relied
on automated searches and simple context analysis. Additional research using more sophisticated qualitative linguistic methods might be useful to expand our understanding of this problem. Secondly, I focused on top-cited, highly competitive journals. The pressure to highlight importance and hide weaknesses may be stronger in these journals. However, if the most prestigious journals give the wrong example, it would be weird to expect lesser ones to resist the dominating trend. While some empirical evaluations have suggested that major journals may sometimes foster exaggerated results [5], significance chasing is probably characteristic of the whole scientific literature [6], down to local non-English language journals [7]. Third, it is impossible to tell whether lack of acknowledgment of limitations is the authors’ choice, or is compounded by peer-reviewers and editors. Do peer-reviewers and editors have a share in downplaying limitations and strengthening the perceived importance of the published articles? Conversely, perhaps they added some limitations that were not mentioned in the original submitted versions.

6. Instructions to the authors and editorial policies

What do journals seek in the articles they publish and what do they instruct authors to do for limitations? I downloaded from the web the instructions to the authors and editorial policies for the 25 scientific journals with highest total number of citations (Journal Citation Reports, 2004 edition) (Table 1). Importance, novelty, originality, significance, and relevance figure prominently in the language of what drives acceptance for publication. Cell asks for “unusual significance”, PNAS for “exceptional importance”, and EMBO Journal for “manuscripts that merit urgent publication”.

Despite very meticulous instructions even on minute style details, instructions to the authors on how to discuss research findings are sparse. Only six of the 25 journals mention anything relevant, and only one of them definitively hints to limitations: JAMA asks for a comment section “addressing study limitations”. The Journal of Neuroscience requests “a discussion of the validity of the observations”, but this is not necessarily limitations-oriented. Of the two examined new-open-access journals, PLoS Medicine also clearly instructs the authors to “include a summary of the limitations of your study” in the Discussion.

Conversely, EMBO Journal asks that even the Discussion “should clearly draw attention to the novelty and significance of the data reported”. Some other statements in the instructions may even deter authors from acknowledging potential problems in their work. Physical Review Letters ask the work to be “free of detectable error” and Biochemistry-US states that “publication of preliminary or inconclusive results is discouraged.” Of course no one would seriously favor salami publications of least publishable units or the promotion of poorly done research. However, much research effort will unavoidably be preliminary, and even some perfectly designed studies will end up being inconclusive. Acknowledgment of these limitations should be encouraged.

Statements that further reinforce the need to show novelty and importance rather than identify limitations may lead authors to downplay the importance of even unavoidable errors and to show that their results are more conclusive than they really are. An empirical evaluation of papers published in the Lancet [8] found that authors acknowledged more limitations to their work compared with the published version when they were surveyed after their paper had been published. Researchers may feel under pressure to show that their work is important, robust, and free of limitations.

7. Final considerations: limitations, expectations and discussions

I argue that a section of limitations should be routinely considered in publications of original research. This does not have to be extensive, but should alert the reader to key weaknesses in the presented work. Identifying weaknesses is a first important step toward formulating new research questions. It may also help toward improving research methodology and the validity of extrapolation of results. Given the dearth of attention given to discussion of limitations, I suspect that the current literature is unbalanced. A few scattered articles discuss limitations in detail, while most say nothing about them even though they may have as many or even more limitations. There is a real danger of adopting dangerously flawed findings, if research is extrapolated and applied based on its lack of reported limitations.

Online publishing should alleviate any concerns about increasing the length of a manuscript, whereas space for print articles is limited. The superiority of open-access journals in acknowledging limitations may reflect space availability. Some open-access journals have also shown already a commitment to the cautious replication of research claims and the dissemination of both “positive” and “negative” results [9]. Traditional print journals have fared rather poorly in this regard to-date. Perhaps in the quest for brevity and outstanding novelty and importance, limitation sections have had very low priority.

Discussion sections in scientific articles have been shown to be notoriously weak, lengthy, and chaotic [10]. Thus, one might argue that Discussion sections should be curtailed as much as possible—perhaps with the exception of limitations. Clinical research and clinical epidemiology may be more sensitized to this problem than other scientific disciplines. For example, the CONSORT statement for the reporting of clinical trials [11] clearly refers to the need to discuss limitations. Similar guidance should be offered also to other disciplines.
The notion of infallibility is not compatible with scientific thinking. Unilateral emphasis on the importance and promises of research may make scientific effort lose credibility, when inflated promises without acknowledged limitations fail to deliver. An empirical evaluation has shown that of 101 publications in top basic science journals that made a clear promise for a major clinical application of their findings, only five were translated into licensed drugs in use 25 years later, and only one of these five had truly a major clinical impact [12,13]. Any success of science, either on a practical application or on a theoretical front, is a major cause for celebration for researchers and possibly for humankind at large, even if the success rate is one in 100 or one in 1,000. However, tempering our expectations and focusing on what can be improved is more constructive than focusing on what is already perfect.

References