The role of case reports in evidence-based practice, with suggestions for improving their reporting

Joerg Albrecht, MD, a Victoria P. Werth, MD, b,c and Michael Bigby, MD d,e
Chicago, Illinois; Philadelphia, Pennsylvania; and Boston, Massachusetts

Background: Case reports are part of the evidence hierarchy in evidence-based practice and guide an important part of dermatologic practice.

Objective: This article discusses the role of case reports and series in evidence-based practice and suggests how to improve reporting.

Methods: This article was inspired by a forum on the role of case reporting in dermatologic practice presented at a meeting of the American Academy of Dermatology. It is based on an informal literature review and the authors’ experience in systematically reviewing case reports in medicine and dermatology.

Results: Case reports significantly influence medicine. Often they are the first line of evidence for new therapies but rarely are sufficient to establish treatment efficacy. Frequently they are the first and sometimes major source for detecting rare adverse events. Guidelines for improving the reporting of case reports and series are discussed.

Limitations: This paper represents the views of the authors and is neither a systematic review of the literature nor a formal consensus document.

Conclusion: Case reports are an important part of medical literature and need to be taken seriously. (J Am Acad Dermatol 2009;60:412-8.)

VALUE OF CASE REPORTS AND CASE SERIES
Case reports and case series are an important part of medical publishing and continue to be published in virtually all dermatologic journals. Often they are the first line of evidence for new therapies (Table I). On rare occasions the information provided in a case report or case series is sufficient to establish efficacy. The use of physostigmine in myasthenia gravis by Walker1 is a particularly elegant example (Fig 1). This type of all-or-none effect,2 another example is insulin for type 1 diabetes mellitus, represents the highest level of evidence in the evidence-based medicine hierarchy of evidence (Fig 2).3 It is exceedingly rare in dermatology. The use of dapsone in dermatitis herpetiformis and of steroids in pemphigus vulgaris may come close. Trials may not be necessary for these treatments, other than to optimize dosing.

Case reports and case series are often the first and sometimes a major source for detecting rare adverse events (Table II). Case reports and series identify rare side effects and are indispensable for identifying side effects that occur in less than 1 in 1000 patients. Only case-control studies, or very large cohort studies,
which are larger than the 3000 patients, are able to quantify the occurrence of these side effects. However, false alarms have destroyed well-tolerated drugs and made them unmarketable. For example, debendox/bendectin, a combination drug including pyridoxine and doxylamine succinate which was marketed for morning sickness in pregnancy, was intermittently thought to induce malformations based on case reports. Careful evaluation could not confirm this risk but the drug was withdrawn from the US market because of threat of litigation. It continues to be available in Canada with no detectable evidence of human teratogenicity.4

Case reports and case series are vivid accounts with high readability and impact. Case reports published in The Lancet have had considerable influence on the literature5; they were quoted on average 17 times and seemed to stimulate more related case reports and series.5 However, a systematic review of the relative citation impact of 2646 articles of different study designs demonstrated that case reports were cited less frequently than other study designs. Meta-analyses were cited most often, and case reports received the least number of citations.6 This observation makes editors understandably hesitant to increase the number of case reports that they publish, and more inclined to reduce their number. For example, The Lancet reduced space for brief reports, because these reports, although influential, still reduced its impact factor.5 However, the Journal of the American Academy of Dermatology (JAAD) has gone the opposite direction and tried to use the format of case letter section with a limit of 500 words to assure that the case reports that merit publication can get published in spite of space restrictions. This decision was made based on the observation that some readers shun the case report section whereas others believe they can learn more from this section than from any other section in the journal.7 The journal thus reflects clinical teaching in dermatology, which is heavily based on case presentations and discussions. Retaining these characteristics in 500 words is not easy but with editorial help this length is sufficient for most clinical communications.

Recently the JAAD decision seems to be echoed by an increased interest in the publication of case reports online through several publishers including: BioMed Central, with its journal of medical case reports8; the University of Washington Radiology Case Reports9; and Karger’s journal Case Reports in Gastroenterology.10 BioMed Central has now added the cases journal11 to its journal of medical case reports. It is headed by Richard Smith, the former editor of the British Medical Journal. This journal is a radical medical publishing experiment in that it is

**Table I. Examples of therapy in dermatology based on case reports**

<table>
<thead>
<tr>
<th>Disease and treatment</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potassium iodide in erythema nodosum</td>
<td>35-37</td>
</tr>
<tr>
<td>Dapsone in Behçet disease</td>
<td>38</td>
</tr>
<tr>
<td>TNF-α antibody treatment in psoriasis</td>
<td>39, 40</td>
</tr>
<tr>
<td>Anti-CD20 in pemphigus</td>
<td>41</td>
</tr>
<tr>
<td>Steroids in pemphigus vulgaris</td>
<td>42</td>
</tr>
</tbody>
</table>

*TNF, Tumor necrosis factor.*

Before injection the patient cannot raise her left eyelid. Thirty minutes after it the eye is fully open. (The photographs are reproduced from a cinematograph film and are reversed left for right.)

**Fig 1.** Myasthenia gravis treated with physostigmine by Walker.1 Reprinted from The Lancet, with permission from Elsevier.

**Fig 2.** Evidence pyramid. Adapted with permission from State University of New York Downstate Medical Center Medical Research Library of Brooklyn, New York. Available at: http://library.downstate.edu/ebmdos/2100.htm. Accessed March 7, 2008.
Table II. Examples of side effects of drugs relevant for the dermatologist described in case reports or series

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Adverse event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-TNF antibody</td>
<td>Vasculitis/vitiligo</td>
</tr>
<tr>
<td>Interferon alfa</td>
<td>Psoriasis exacerbation</td>
</tr>
<tr>
<td>Terbinafine</td>
<td>Subacute cutaneous lupus erythematosus</td>
</tr>
<tr>
<td>IV immunoglobulin</td>
<td>Stroke</td>
</tr>
<tr>
<td>Chinese herb treatment of skin disease</td>
<td>Nephropathy</td>
</tr>
<tr>
<td>Naproxen</td>
<td>Pseudoporphyria</td>
</tr>
<tr>
<td>Immunostimulatory herbal supplements</td>
<td>Activation of autoimmune diseases</td>
</tr>
</tbody>
</table>

IV, Intravenous; TNF, tumor necrosis factor.

... studies of rare diseases because they have little potential to generate revenue.

**DRAWBACKS OF CASE REPORTS AND CASE SERIES**

In general, case reports and case series are relatively low in the evidence hierarchy (Fig 1) and there are many methodological arguments against the use of case reports and series to advocate new therapies (Table III). The most serious threat to the validity of case reports and series in the medical literature, however, is publication bias. In one survey, the number of published case reports and case series reporting successes was 90% versus 10% reporting failures.

The British National Institute for Clinical Excellence (NICE) found that 30% of their reviews on the effectiveness of health care technology included information from case series and commissioned a recent review on the subject. In this review, the authors extensively assessed the question of whether case series systematically overestimate or underestimate treatment response, but found no evidence that they do either, particularly if they are done well. Dalziel et al evaluated 47 reports completed between 2000 and 2002 and found that 14 reports included case series, 12 of which also reported two to 159 trials, whereas two reports relied solely on case series. One was a surgical intervention, namely the use of autologous chondrocyte transplantation for hyaline cartilage defects in knees, whereas the other concerned rituximab for refractory or recurrent stage III or IV follicular Hodgkin lymphoma. Neither of these therapies poses particular hurdles to evaluation in clinical trials nor are they concerned with rare diseases, but randomized trials were not available. It was also necessary to include information from 27 case series to evaluate safety in larger numbers of patients even for a NICE review on the effectiveness of health care technology concerning the use of atypical antipsychotics in schizophrenia. Case series were used in spite of the availability of 70 clinical RCTs. Another NICE review on the effectiveness of health care technology concerning growth hormone in children that had been evaluated in 21 RCTs used an additional 11 case series to gain information about final height, which had not been evaluated in RCTs, although it is arguably the most relevant treatment outcome. However, the timeline for measuring this end point is beyond the requirements of any licensing authority and thus there is little incentive for pharmaceutical companies to provide such data from controlled trials. Some reports defined stringent inclusion criteria for case series (e.g., >2000 patients and >2 years...
of follow-up for the safety data on atypical antipsychotics). However, for other reports the inclusion criteria (eg, n > 10 for the evaluation of rituximab in Hodgkin lymphoma, which was mentioned above) were not stringent probably because of lack of other sources of information. This lack of stringent inclusion criteria for case series is often the case in published case series in dermatology.18

Case reports and case series are responsible for many discarded therapies. Nearly every discarded once-popular therapy was probably supported by a series of favorable cases.2 To assess the influence of case reports and series on medical practice some authors have looked at off-label use of medications. Picard et al19 reviewed off-label use in a university dermatology clinic and found that 86% of the prescriptions were used on label and 14% were off label; 70% of these off-label prescriptions were not based on strong evidence. Most off-label prescriptions were made for inflammatory or hypersensitivity dermatoses. Confirming this observation, a German study delineated the problems of off-label use in Germany and pointed to the complete lack of labeled treatments for scleroderma and the observation that no topical treatment was licensed for lupus erythematosus.20 Li et al21 evaluated off-label use in dermatology. A total of 55 dermatologists from Boston, MA, were surveyed. All of these physicians used off-label therapy and only 15% (n = 8) who still used off-label therapy thought it was important to use a drug for Food and Drug Administration (FDA)-approved indications. These dermatologists were given a list of 22 off-label indications. Five of these 22 indications were thought to be FDA approved by more than 50% of the dermatologists. Fifteen of the 22 treatments were systemic and 7 were topical.

Thalidomide is an extreme example of a drug that is used primarily off label. It was licensed in 1998 for the treatment and prevention of erythema nodosum leprosum. This disease is extremely rare with 96 new cases reported in the United States in 2002.22 However, certainly until multiple myeloma was added in 2006 as an approved indication, this drug was used off label by far more than 90% of the physicians who prescribed it. It is now used for a number of difficult-to-treat diseases, even though its use is only informally endorsed (Table III).23 Thalidomide is teratogenic and the FDA has put strict and burdensome limitations on its use. Thus, it is not surprising that many indications for which it is a second- or third-line therapy have not been evaluated. For one prominent use, namely HIV cachexia, the randomized trial providing evidence for this indication was only published in 2006. This was 11 years after the FDA approved its use for this indication through an expanded access program.24

Pharmaceutical companies in the United States are not allowed to advertise off-label indications of their drugs; however, they do. That they do was demonstrated in detail for gabapentin.25 One means of promotion is educational supplements. The JAAD has published educational supplements such as the one on imiquimod financed by 3M (St Paul, MN) in October 2002. It is clear that the goal of the company was to disseminate information on the off-label use of the drug via publication of case reports and small series, which all strove to demonstrate therapeutic success. Although there are RCTs for some of the indications available now, this sponsored publication is problematic in the sense that it seems to violate the principle that off-label use of a drug should not be advertised.

**SUGGESTIONS FOR IMPROVING THE REPORTING OF CASE REPORTS AND CASE SERIES**

Although there is little, there is some literature on how to improve reporting of cases and case series. Articles by Moses2 and Abel,26 chapters in the dermatologic literature,27 and, most comprehensively, *Clinical Case Reporting in Evidence-based Medicine* by Jenicek28 provide guidance. Table IV is a compilation of suggestions from these sources, general ethical considerations, and of the authors’

### Table III. Thalidomide off-label indications and supporting evidence based on PubMed search (November 28, 2007, search terms: disease name, randomized controlled clinical trial)

<table>
<thead>
<tr>
<th>Observation evidence</th>
<th>Randomized controlled clinical trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain neoplasm</td>
<td>Behçet syndrome</td>
</tr>
<tr>
<td>Crohn disease (lenalidomide failed trial)</td>
<td>Cachexia caused by HIV</td>
</tr>
<tr>
<td>Discoid lupus erythematosus</td>
<td>Graft-versus-host disease</td>
</tr>
<tr>
<td>Erosive lichen planus</td>
<td>Mycobacteriosis</td>
</tr>
<tr>
<td>Erythema multiforme</td>
<td>Prevention of recurrent aphthous stomatitis (negative)</td>
</tr>
<tr>
<td>Prevention of graft-versus-host disease</td>
<td>Recurrent aphthous stomatitis</td>
</tr>
<tr>
<td>Subacute cutaneous lupus erythematosus</td>
<td></td>
</tr>
</tbody>
</table>

---

**Table III.** Thalidomide off-label indications and supporting evidence based on PubMed search (November 28, 2007, search terms: disease name, randomized controlled clinical trial)
experience based on systematic reviews of case reports and series in dermatology and general medical journals.

The patient in a case report should be described fully and succinctly. The description should include the patient’s age, sex, clinical history up to the intervention of interest, comorbidities, and outcome. The intervention should be described in detail. The description should include the dose, schedule of administration, and duration of medications received. If a diagnosis is rendered, it should be clearly defined and met by the patient.

Case reports and series are often the result of compassionate use of treatment rather than planned clinical trials. For the latter, informed consent and institutional review board approval is required. Institutional review board approval is not mandatory for compassionate use but informed consent needs to be sought from the patient and reported for any novel intervention.

The implicit comparison for case reports and case series is either the natural course of the disease or the course under established treatment. The natural history of the disease or expected course after intervention should, therefore, be described. Failed treatments per se do not establish incurability. Many dermatologic diseases have a waxing and waning natural courses and many may spontaneously remit. A bad outcome after an intervention does not necessarily imply causality. Establishing causality is enhanced by having multiple cases, positive results from withdrawal and rechallenge, and more formal clinical studies (eg, case-control or cohort studies).

The outcomes of the therapy should be well defined and clinically relevant. Objective outcomes such as the complete disappearance of skin lesions are well defined and easy to report and understand. However, not all patients are cured and demonstrating clinical improvement can be difficult. Ideally, validated outcome measures such as SCORAD for atopic dermatitis or CLASI for cutaneous lupus erythematosus are used. Even they have limitations. For many diseases there are no available outcome instruments, the measures of improvement are subjective, and improvement is often described in terms of percentage of change toward being free of disease. This approach has its pitfalls, particularly if the disease, like most, has more than one dimension (eg, pruritus, erythema, and scaling). Therefore, what constitutes a 50% improvement has to be clearly defined. Even the assessment of area skin surface involved is notoriously difficult and inexact; even though it gives the impression of precision, accuracy as well as intraobserver and interobserver reliability are low.

The patients’ perception of any intervention needs to be reported. This admonition is particularly so because of the difficulties in assessing patient outcomes noted above. Dermatologic treatments can be tedious and expensive, thus we need to know about the patient’s perception of the intervention and their clinical response to it. If patients discontinue treatment the reasons need to be documented, if they are known.

Inclusion and exclusion criteria should be explicitly stated in case series. Explicit criteria help readers to apply the series to their patients and help to define the patients who have received the intervention and those who should not receive it.

### Table IV. Checklist for quality assessment of case reports or case series describing innovative therapies

<table>
<thead>
<tr>
<th>Checklist for reporting cases and case series</th>
<th>Are diagnostic criteria clearly identified, and met by patient(s)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informed consent</td>
<td>Has patient consent been documented?</td>
</tr>
<tr>
<td>Natural disease course</td>
<td>Is there any reference to natural course of disease, or, if applicable, course on standard treatment?</td>
</tr>
<tr>
<td>Dosages</td>
<td>Are treatments’ dosage, duration, and titration described adequately?</td>
</tr>
<tr>
<td>Outcome measures</td>
<td>Are outcomes well defined and clinically relevant? Are they subjective, standardized, or subjective?</td>
</tr>
<tr>
<td>Patient perception</td>
<td>Is there any documentation of the patients’ perception of outcome and of the intervention?</td>
</tr>
<tr>
<td>Safety</td>
<td>Do the authors describe known risks associated with intervention?</td>
</tr>
<tr>
<td>Authors’ conclusions</td>
<td>Do the authors abstain from unfounded claims about safety and efficacy?</td>
</tr>
<tr>
<td>Inclusion and exclusion criteria*</td>
<td>Are inclusion and exclusion criteria clearly stated?</td>
</tr>
<tr>
<td>Consecutive cases*</td>
<td>Are all consecutive patients treated by one physician or at one institution included?</td>
</tr>
</tbody>
</table>

IRB, Institutional review board.

*For case series.
For example, patients with renal failure, children, or pregnant women may have been excluded. In addition, the definition of treatments, which may be less toxic, that need to have failed before the experimental treatment is tried should be included in the list of inclusion and exclusion criteria.

The greatest threat to the credibility of case series is selection bias that is often introduced by reporting selected cases. Therefore, all consenting, eligible patients at an institution or under the care of the physician or physicians reporting the case series should be included in the series. If eligible patients refused to be part of the study or for other reasons did not receive the treatment, it should be noted and the reader should learn about their outcome. If patients were lost to follow-up, it should be documented. The patients who cannot be followed up may differ significantly from those who come back. The former may be better, worse, or unchanged than the patients reported.

The conclusions that can be drawn from case reports and case series are generally quite limited. They can demonstrate efficacy of treatment only under extremely rare circumstances (eg, when the effect is dramatic and no other effective therapy is available). Because common and uncommon side effects may not occur in case reports or case series, it is impossible to demonstrate the safety of any intervention. It is, therefore, incumbent on the authors to review what is known about the risks of the intervention from other available sources. A case report or case series that concludes that the intervention is “safe and effective” should be viewed with extreme scepticism.

CONCLUSION
Case reports have influenced and continue to influence dermatology, and in spite of all the caveats associated with their use they should continue to be published. The quality of reporting is particularly important. Guidelines to improve their reporting are discussed.

We thank Dr Jeffrey Bernhard for his participation in the original workshop from which this article stems as well as insightful comments and recommendations during the preparation of this manuscript.

REFERENCES


