Clinical Research Publication by Spanish Dermatologists Over Time and in Comparison With Other Research Groups in 2008

B. Aranegui,* I. García-Doval, and A. García-Cruz

Servicio de Dermatología, Hospital Provincial, Complexo Hospitalario de Pontevedra, Pontevedra, Spain

Manuscript received November 11, 2009; accepted for publication May 18, 2010

KEYWORDS
Biomedical research; Evidence-based medicine; Research projects; Epidemiology; Dermatology; Spain

Abstract
Introduction: Clinical research is the form of research nearest to clinical practice. Material and methods: For the years 1992, 1996, 2000, 2004, and 2008, we identified all indexed articles published by Spanish dermatologists and calculated the percentages corresponding to clinical research according to a previously validated definition; we then calculated the proportion of clinical research articles offering higher levels of evidence. For 2008, we compared these percentages to those of French and British dermatologists and Spanish rheumatologists. We also compared these groups' rates of productivity in 2008 in relation to articles providing higher levels of evidence. Results: In 2008, 36% of Spanish dermatologists' publications reported clinical research; 7% were studies offering higher levels of evidence. The proportions did not change significantly over the period studied. Clinical research publications accounted for 35% and 43% of the articles by French and British dermatologists in 2008 and 54% of articles by Spanish rheumatologists in that year. The proportion of publications reporting clinical research was significantly higher for Spanish rheumatologists than for Spanish dermatologists. The proportions of publications offering higher levels of evidence were significantly different in 2008 only for the comparison between Spanish dermatologists and rheumatologists. Other differences were not statistically significant. In the comparison of rates of productivity in clinical research offering higher levels of evidence, British dermatologists were significantly more productive than Spanish dermatologists. Discussion: Differences were observed in relation to specialty (Spanish dermatologists vs rheumatologists) and nationality (Spanish vs British dermatologists). Conclusions: The reasons for the differences identified need to be studied in order to improve this situation.

© 2009 Elsevier España, S.L. and AEDV. All rights reserved.

*Corresponding author.
E-mail address: baranegui@gmail.com (B. Aranegui).
Clinical Research Publication by Spanish Dermatologists Over Time and in Comparison With Other Research Groups in 2008

Introduction

Clinical research seeks to answer questions that arise during medical practice concerning the etiology of disease, risk factors, diagnosis, treatment, and prognosis. The ultimate goal is to improve patient care. The objects of study might be patients or other subjects, samples, or health care systems; previously published articles might also come under scrutiny if the research is a systematic review. For clinicians, clinical research is easier to comprehend than basic research and the findings are more readily applicable to routine practice.¹

Various scales are available for evaluating the internal validity of a study before deciding the results can provide reliable guidance in a clinical setting. One of the most widely applied scales is the 5 levels of evidence set out by the Centre for Evidence-Based Medicine in Oxford²: the first (highest) level encompasses research that provides the most reliable evidence and the fifth (lowest) level those that are most susceptible to random error, confounders, and bias.

Dermatology journals contain a variety of other types of articles in addition to reports of original clinical research. The main types are basic-science studies, case reports, opinion articles, and letters commenting on previous publications.

Our aim was to determine the percentage of publications by Spanish dermatologists that fall into the category of clinical research, assess the levels of evidence the articles contribute, and detect changes over time. To provide background context for the results of this study, we compared publications by Spanish dermatologists in 2008 to the output of other groups that we thought could provide good models: French and British dermatologists and Spanish rheumatologists.

Material and Methods

We conducted a cross-sectional study of articles published in 2008 by Spanish, French, and British dermatologists and Spanish rheumatologists. In addition to the data for Spanish dermatologists for 2008, we also investigated the publication history of this group by analyzing their articles published in 1992, 1996, 2000, and 2004.

Search strategy: Articles were located in April 2010 through the PubMed portal to the MEDLINE database. Publications for each group were found using analogous, reproducible strategies (Table 1). Articles published in Actas Dermo-Sifiliográficas were excluded. This journal was not indexed in MEDLINE at the start of the study period; moreover it is likely that the percentage of clinical
Table 1: Search Strategies Used to Locate Articles Through PubMed

<table>
<thead>
<tr>
<th>Category</th>
<th>Search Strategy</th>
</tr>
</thead>
</table>

research published there would be different from the percentages published in non-Spanish journals. Data from Actas Dermo-Sifiliográficas would therefore distort the study of trend over time.

Criteria for including articles: We included MEDLINE-indexed articles for which the affiliation field for the corresponding author contained the roots Dermat* or Rheumat* (or Rheumat*) and which were given page numbers that fell in volumes for each study year (a criterion not needed for online-only journals).

Criteria for excluding articles: We excluded articles on topics that were obviously unrelated to dermatology (or rheumatology, with respect to the study of Spanish rheumatologists’ output). This exclusion was in response to the presence of publications from large departments that encompass different specialties but whose name includes the root Dermat* (or Reumat* or Rheumat*); an example is a department that groups together medicine, psychiatry and dermatology at the university hospital in Saragossa, then called Hospital Lozano Blesa.

Classification of documents: We considered an article to be related to clinical research if it met the following 3 criteria:

1. The study was performed in patients, other persons or health care systems or was based on patients. Included would be research on samples of patients or healthy individuals, biopsies, dermoscopic images, laboratory findings, etc.
2. The study set out to answer a question about clinical practice in order to solve problems of patient management. Included would be research on the etiology, diagnosis, prognosis, treatment, prevention, and prevalence of disease as well as studies on economic aspects of disease or health care systems. Systematic reviews on these aspects were also included.
3. The study had at least a level 4 evidence grade according to the Oxford Centre for Evidence-Based Medicine. This criterion meant that we did not count as clinical research any case reports or publications based on “expert opinion without explicit critical appraisal, or ... on physiology, bench research or ‘first principles.’”

A research article was classified as contributing a high level of evidence if it received a classification of at least 3 under this system.

Document processing: The results of searches for each year and group were stored in EndNote libraries (Thomson, 2006) and a spreadsheet (Excel, Microsoft Office 2007). (The database may be obtained from the authors.) The classification process began with a reading of the abstract; if the abstract gave insufficient information or there was no abstract, the entire article was read. The names of authors were not hidden from the person classifying the article, as blinding would have been difficult and previous studies indicate that knowledge of the identity of authors is not a source of bias when articles are categorized. Classification required 2 consecutive steps. First, a dermatology resident with no specific training in epidemiology (B.A.A.) assigned articles to categories according to the aforementioned criteria and then separated them by levels of evidence (5, 4, and 3). All articles with a level of 3 or better were then reviewed by a dermatologist with Masters-level training in epidemiology (I.G.D.) to confirm that the evidence level had been assigned correctly. The assessment of evidence levels by an evaluator with this level of training has been reported to be highly reproducible.

The definition of clinical research was previously validated (data pending publication). The present study is part of a series of pilot studies related to that validation. The interobserver reliability of the classification system is excellent as confirmed by a $k$ coefficient of concordance of 0.95 (95% confidence interval, 0.91–0.99) for 2 observers without specific training in epidemiology. This reliability study was based on the classification of 212 articles published in 3 high-impact dermatology journals. When the observers’ classifications were reviewed by an expert in methodology, only 3 were found to be incorrect.

Statistical analysis: The STATA 10 (StatCorp LP, 2009) software package was used to compile descriptive statistics, perform $\chi^2$ tests, and study trends.
Results

Study population: A total of 1111 articles were included according to the stipulated inclusion and exclusion criteria.

Spanish dermatologists' publications in 2008: Thirty-six percent (59/165) of Spanish dermatologists' publications in 2008 were clinical research articles; 7% (11/165) provided a high level of evidence. Tables 2 and 3 show the Spanish dermatologists' output for this year by levels of evidence.

Time course of Spanish dermatologists' publications (Table 2): Overall around 31% of Spanish dermatologists' publications have been clinical research articles; about 6% of articles have offered a high level of evidence. No significant changes in the percentage of clinical research articles have occurred over the years ($\chi^2$ test, $P=.1$). Nor was there a linear trend (test for trend, $P=1$). Likewise, we observed no significant changes in the levels of evidence ($\chi^2$ test, $P=.17$) or a linear trend (test for trend comparing percentage of articles with a level of evidence ≤3 over time, $P=.11$).

Comparison of Spanish dermatologists to other groups: Table 3 shows descriptive statistics on publications in 2008 for all the study groups. Thirty-six percent (59/165) of the Spanish dermatologists' publications were clinical research articles in that year, whereas 35% (92/267) of the output of French dermatologists was in this category; the percentage was 43% (82/193) for British dermatologists and 54% (35/65) for Spanish rheumatologists. The differences observed in these figures reflect the total number of publications by these groups (not a population sample). If we consider these figures to be a representative sample of what these groups do over longer periods of time, the differences were not significant between Spanish and French dermatologists ($\chi^2$ test, $P=.78$) or Spanish and British dermatologists ($\chi^2$ test, $P=.19$). Only the difference between Spanish dermatologists and rheumatologists was significant ($\chi^2$ test, $P=.01$). Thus, we cannot rule out that the differences between dermatologist groups in 2008 were not due to random variation. The actual levels of evidence provided by the publications also differed (Table 3): in 2008 clinical studies providing a high level of evidence accounted for 7%, 9% and 12% of publications by Spanish, French and British dermatologists, respectively, and 32% of publications by Spanish rheumatologists. Once again, if we consider these data to represent what these groups do over a longer period of time, we cannot conclude that the differences in evidence levels are significant between Spanish and French ($\chi^2$ test, $P=.29$) or Spanish and British ($\chi^2$ test, $P=.2$) dermatologists. Only the difference between Spanish dermatologists and rheumatologists is significant ($\chi^2$ test, $P<.001$).

We calculated the number of physicians for each article giving a high level of evidence as a measure of scientific productivity for each group (Table 4). For this calculation, the number of dermatologists was considered to be 1804 in Spain,
The number of Spanish rheumatologists was 1500 according to the Spanish Society of Rheumatology (1431 members plus approximately 5% of practicing rheumatologists who were not members). Spanish dermatologists had the lowest productivity rate, at 1 original clinical research article providing high-level evidence for every 164 dermatologists in 2008. Considering the figures for 2008 to be representative, we found no significant difference between the Spanish and French dermatologists (exact test, \( P = .76 \)), but the rates for British dermatologists and Spanish rheumatologists were significantly higher (exact test, \( P < .001 \) and \( P = .01 \), respectively).

**Discussion**

**Study findings:** Thirty-six percent of Spanish dermatologists’ publications in 2008 were clinical research articles; 7% provided high-level evidence. No changes were observed over the period under study: overall, around 31% of articles corresponded to clinical research and 6% provided high-level evidence.

Differences in absolute percentages of clinical research articles and articles giving high levels of evidence were observed in 2008; differences in rates adjusted for numbers of physicians were also seen that year. The marked differences between Spanish dermatologists and rheumatologists were all significant. Factors peculiar to the 2 specialties may account for the variation. There may be differences in hours devoted to academic medicine or private practice in each specialty, or one specialty may favor case reports more than the other. The differences between groups of dermatologists were smaller. Only slight differences were found between Spanish and French dermatologists in 2008 and we cannot rule out random variation. The comparison between Spanish and British dermatologists for 2008, however, showed that the latter produced more clinical research and that nearly twice the proportion provided high-level evidence (Spanish dermatologists, 7%, vs 12% for British dermatologists). With the size of the sample we studied, we could not rule out that the differences were due to random variation, but the recurrence of the same pattern across several comparisons suggests that the differences were real. The scientific productivity of the British dermatologists was also higher in terms of rate of publications providing high-level evidence, on adjustment for numbers of physicians, and this difference was clearly not due to random variation in 2008.

The possible causes of these differences (especially productivity), other than those that were possibly random, should be studied further.

**Strengths of the study:** This study is the first to look at the publication of clinical research by Spanish dermatologists overall and publications providing high-level evidence in particular. Likewise, it is the first to study these aspects over time. We have also investigated the same publication patterns among similar groups in 2008 for purposes of comparison.

The PubMed search portal provides exhaustive access to indexed articles, and we applied search strategies and inclusion and exclusion criteria that ensured that large-
scale omissions would be avoided and that most articles by Spanish dermatologists would be located. The definition of clinical research we used has been validated with respect to content, feasibility, and reproducibility. It allows clinical research publications to be distinguished from other types without large measurement error.

Limitations: Articles published by Spanish dermatologists in *Actas Dermo-Sifiliográficas* were not included, as this journal began to be indexed in MEDLINE after the start of the study period. This exclusion was necessary in order to detect changes over time in the percentage of Spanish authors’ publications that corresponded to clinical research in a homogeneous journal sample and to guarantee that the groups could be compared. We think it is quite likely that *Actas Dermo-Sifiliográficas* does publish a larger percentage of clinical research or research providing high levels of evidence than can be found in international journals, as that has been shown to be the case in rheumatology.

Therefore if our exclusion of *Actas Dermo-Sifiliográficas* in fact was a source of bias in the present study, its probable effect was to have made the Spanish dermatologists’ output seem more positive than it actually was. Excluding this journal makes the comparison of results with those of rheumatologists more useful, given that the journal of the Spanish rheumatology association (*Reumatología Clínica*) is not indexed by MEDLINE and was not included in the study.

PubMed searches have their limitations. First, assignment of an article to a group is determined by means of the affiliation field of the database. Searching this field reduces subjectivity and avoids bias given that the same strategy was applied to all the groups equally. However, this approach is imperfect. Some articles, such as letters, do not have an affiliation field. Second, errors are possible. For example, the word *British* might be part of the phrase *British Columbia* and lead to the incorrect inclusion of non-British articles. Another example is the term *Dermat* in the institutional name of large multidisciplinary departments, leading to the incorrect inclusion of non-British articles. Another example is the term *Dermat* in the institutional name of large multidisciplinary departments, leading to the incorrect inclusion of non-British articles. Another example is the term *Dermat* in the institutional name of large multidisciplinary departments, leading to the incorrect inclusion of non-British articles. Another example is the term *Dermat* in the institutional name of large multidisciplinary departments, leading to the incorrect inclusion of non-British articles. Another example is the term *Dermat* in the institutional name of large multidisciplinary departments, leading to the incorrect inclusion of non-British articles.

### Table 3  Distribution of Clinical Research Published in 2008 by Level of Evidence and Study Group: Clinical Research (Evidence Levels 1-4) and Other Publications by Spanish, French, and British Dermatologists and by Spanish Rheumatologists During the Study Period

<table>
<thead>
<tr>
<th>Publication Type</th>
<th>Group</th>
<th>Spanish Derm, n (%)</th>
<th>French Derm, n (%)</th>
<th>British Derm, n (%)</th>
<th>Spanish Rheum, n (%)</th>
<th>Total, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>[Cum %]</td>
<td>[Cum %]</td>
<td>[Cum %]</td>
<td>[Cum %]</td>
<td>[Cum %]</td>
</tr>
<tr>
<td>Clinical research, evidence level 1</td>
<td>3 (1.8)</td>
<td>9 (3.4)</td>
<td>5 (2.6)</td>
<td>8 (12.3)</td>
<td>25 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Clinical research, evidence level 2</td>
<td>2 (1.2)</td>
<td>8 (3.0)</td>
<td>9 (4.7)</td>
<td>10 (15.4)</td>
<td>29 (4.2)</td>
<td></td>
</tr>
<tr>
<td>Clinical research, evidence level 3</td>
<td>6 (3.6)</td>
<td>7 (2.6)</td>
<td>9 (4.7)</td>
<td>3 (4.6)</td>
<td>25 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Clinical research, evidence level 4</td>
<td>48 (29.1)</td>
<td>68 (25.5)</td>
<td>59 (30.6)</td>
<td>14 (21.5)</td>
<td>189 (27.4)</td>
<td></td>
</tr>
<tr>
<td>Not clinical research</td>
<td>106 (64.2)</td>
<td>175 (65.5)</td>
<td>111 (57.5)</td>
<td>30 (46.2)</td>
<td>422 (61.2)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>165 (100.00)</td>
<td>267 (100.00)</td>
<td>193 (100.00)</td>
<td>65 (100.00)</td>
<td>690 (100.00)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Cum, cumulative; Derm, dermatologists; Rheum, rheumatologists.

### Table 4  Productivity of Each of the Studied Groups in 2008

<table>
<thead>
<tr>
<th>Publication Type</th>
<th>Publications, Evidence Level ≤3, n</th>
<th>No. of Physicians in 2008</th>
<th>Ratio of High-Evidence Level Publication to Physicians in 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish dermatology</td>
<td>11</td>
<td>1804</td>
<td>1 article: 164 physicians</td>
</tr>
<tr>
<td>French dermatology</td>
<td>24</td>
<td>3500</td>
<td>1 article: 146 physicians</td>
</tr>
<tr>
<td>British dermatology</td>
<td>23</td>
<td>451</td>
<td>1 article: 20 physicians</td>
</tr>
<tr>
<td>Spanish rheumatology</td>
<td>21</td>
<td>1500</td>
<td>1 article: 71 physicians</td>
</tr>
</tbody>
</table>
A second problem is that MEDLINE has changed over time, mainly in recent years. For searches prior to 2008, no year-related changes were relevant, but in 2008, PubMed began to index both printed articles and those posted ahead of print. If a search is repeated after this year, some of the articles found in 2008 because they were posted ahead of print will be counted again when they are printed, whether in 2008, 2009, or (more rarely) 2010. The most recent search we performed, in April 2010, took into consideration only the date of print publication by including “NOT ‘2009’[PDAT] NOT ‘2010’[PDAT]” in the search term string to exclude articles published ahead of print in 2008 and printed in 2009 or 2010 (Table 1). Only 1 article was found to have been posted ahead of print in 2008 and those published in 2010; thus, improper inclusion would take place only occasionally; it would affect the percentage only slightly and not introduce bias. Our search strategy included both articles posted ahead of print in 2007 and published in 2008 in online journals without a printed version (eg, Dermatology Online Journal).

Categorizing by levels of evidence also presents some difficulty. A problem first emerges because the categories are oriented to assessing the level of evidence in results, but we were more interested in using them to classify research design. For this reason we were occasionally faced with anomalies, such as articles reporting clinical trials which did not provide level 1 evidence because the results were unclear. Another problem is that the categories are not exhaustive: some designs are not included and there are gaps between the levels, such that a certain degree of subjectivity was needed to classify some articles. Prevalence studies, which are common in dermatology, were an example of a type that gave problems if there was no explicit mention that they were cross-sectional studies. Also difficult were clinical guidelines that fail to describe the search methodology and could therefore not be considered systematic reviews. Reviews of systematic reviews were similarly problematic, as were clinical assessment tools. All of these were included as level 3 studies. Systematic reviews often answer several questions and the level of evidence available for answering each may be different. In these cases we assigned the highest level attained. These difficulties affected a small proportion of articles and did not affect the classification of levels 4 and 5 at all. Therefore, they did not jeopardize the identification of clinical research articles of interest for this study. Furthermore, we do not believe that the percentage of incorrectly classified articles would differ from group to group and therefore this would not cause confounding. Case series were also problematic. Given the difficulty of assigning an arbitrary patient number as a cutoff so that a report might be considered a case series, we decided to assign a level of 4 to any report of 2 or more cases.

Comparison with the literature: The low level of evidence available in the medical literature, and its scarce relationship to the practical needs of physicians and patients is currently of great concern. More clinical research is needed, it should be relevant to practice, and it should provide a high level of evidence.

We are not aware of previous studies that evaluated the percentage of publications that fall into the category of clinical research. Studies of research productivity are not usually conducted as ours was. Some focus on 1 or several journals and are more interested in evaluating the editorial quality of a particular journal than assessing the scientific output of a specific population of physicians. Most classify articles according to research design. Some distinguish between weak designs and strong ones on the basis of the “statistical inference” (sic) that these designs theoretically support. Weak designs (case reports and case series) were found in 70% to 80% of articles published in Spanish journals with the highest impact (Medicina Clínica and Revista Clínica Española) in 1975 and 1984 whereas such designs accounted for half the articles in The New England Journal of Medicine and The Lancet. The gap had widened in 1997 when this aspect was studied again. It appears evident that studies that give higher-level evidence are published in journals with higher impact factors.

The publications of Spanish dermatologists have been studied by hospital, institution, autonomous community, and province, so that figures are available for different authorship groups; however, no analysis has looked at whether publications corresponded to clinical research or not. Nor has the level of evidence been studied. Between 1997 and 2006, case reports accounted for 69.3% of the publications of Spanish dermatologists in international journals and a downward trend was noted. Descriptive studies accounted for 81.9% of the publications in Anais Brasileiros de Dermatologia between 2003 and 2007. Of French dermatologists’ publications in 1998, 21% were original articles (with no differentiation by type), 9% were case series, 37% were case reports, and 31% were tutorials (articles for continuing professional development).

Spanish rheumatologists publish more clinical research than Spanish dermatologists and the level of evidence provided by the rheumatologists’ work is higher. One group compared the publications of 8 French dermatology departments to the output of the 8 rheumatology departments of the same hospitals, finding that the rheumatologists published fewer articles but their work was more often found in English language journals with high impact factors or in basic science journals. When the articles by Spanish rheumatologists in 7 specialty journals with the highest impact factors and visibility and in Revista Española de Reumatología were compared, “inferential studies” (sic) and studies relying on descriptive statistics each accounted for 20% of their output; designs providing higher levels of evidence were more often found in international journals than in the Spanish journal.

Those findings seem to be consistent with ours. The level of evidence of publications of in dermatology would rise if there were more interaction between clinical practice and epidemiology.

Conclusions

Clinical research accounts for a relatively low proportion of Spanish dermatologists’ publications and the proportion did not tend to rise during the study period. Nor did we observe a rising trend in the proportion of clinical research
affording a high level of evidence. In comparison with publications by other groups in 2008, Spanish dermatologists published proportionally less clinical research than British dermatologists or Spanish rheumatologists. Moreover, the level of clinical evidence provided by Spanish dermatologists’ clinical research articles tended to be lower than that of the other 3 groups. The differences between Spanish dermatologists and others in the same specialty were small and random variation cannot be ruled out. The differences between dermatologists and rheumatologists in Spain were greater and it is not likely that they are random. Characteristics related to the respective specialties may be responsible. Research productivity, in terms of the rate of publication of articles providing high-level evidence adjusted for numbers of physicians, was similar for French and Spanish dermatologists, but there were marked differences with respect to both British dermatologists and Spanish rheumatologists, whose productivity rates were higher. The reasons for the differences identified need to be studied if this situation is to improve.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank Professor Jaime Toribio for his comments and collaboration when this work was presented for the Diploma of Advanced Studies. We also thank Javier Osorio for his assistance in collecting the articles.

References