Medical journals get several thousands of submissions per year and editors make every effort to select the best manuscripts for publication. Many of us have had the frustrating experience of seeing our manuscripts rejected by scientific journals. In such a situation, we usually blame the reviewers or the editors thinking that they misunderstood the relevance of our manuscript. In the European Journal of Cardio-Thoracic Surgery (EJCTS), associate editors (AEs) play the hard role of gatekeepers [1], opening or closing the way for a final approval of manuscripts by the Editor-in-Chief. We, as AEs, have to review hundreds of manuscripts a year and are exposed to making mistakes in both ways: either accepting not-so-good papers or rejecting manuscripts that are finally approved and highly cited in other journals. Agreement between reviewers facilitates very much the editorial process. Indeed, if two reviewers have found major problems and both recommend a paper to be excluded from publication we should admit that some wrong aspects of the manuscript have to be modified before a sound text can be printed. The problem arises when major discrepancies are evidenced by the reviewers’ comments. In such instances, an arbitrator can be invited or, alternatively, the AE makes a thorough assessment of the manuscript, if the topic is within his/her field of expertise.

The role of reviewers’ agreement in peer-reviewed medical journals has been studied several times in the literature. While some authors conclude that agreement is only slightly better than would be expected by chance [2, 3], others, using a much better statistical methodology [4], conclude that there is substantial agreement between referee recommendations. The purpose of this editorial is to let the readers know how the agreement between thoracic reviewers for EJCTS is.

We have retrospectively reviewed all thoracic manuscripts submitted to EJCTS from January up to August 2013. Papers rejected outright by the AE were excluded from the analysis. Of 810 papers, we randomly selected 202 (25%) as a valid sample. Manuscripts were sorted according to the date of submission and selected for the study following a random list of numbers created in Excel. Reviewers’ recommendations were coded from 1 to 4 as follows: 1, accept outright; 2, revise or accept in another format; 3, complete rework and; 4, reject outright, redundant submission or unsuitable subject. Coded recommendations were placed on 4 × 4 tables (Table 1) and inter-rater reliability was calculated using the quadratic weighted Gwet’s AC1 coefficient [5] by means of the AgreetStat self-automated workbook for Excel available at http://agreetstat.com/agreetstat.

The weighted percentage of agreement was 0.81 (standard error 0.02) and Gwet’s AC1 coefficient 0.44 (standard error 0.06). Our results show a reasonably good, but not perfect, agreement between thoracic reviewers for EJCTS.

In a systematic review of the usefulness of the peer-review method for selecting top-quality manuscripts in medical journals, Jefferson et al. [6] concluded that, although widely used, peer review is largely untested and its effects are uncertain. Some alternatives to peer reviewing have been proposed. Some authors have suggested that ‘post-publication review-in which the article is assessed by the general readership of the journal instead of a small group of appointed reviewers—could potentially supplement or replace the peer-review process’ [7]. We cannot imagine how to finance the large number of pages that would be needed for a journal adopting such a policy.

Many a time reviewers are biased in their task by their own experience [8], the professional relevance of the authors’ names or institutions [9] and other variables influencing their decisions as published [10]. Older reviewers tend to reject a higher rate of manuscripts because they do not expect a high impact from many of them.

In spite of not being perfect, the peer-review process is currently the best available way for selecting high-impact articles [11]. From our point of view, the quality of the reviews is key to a proper selection of the best papers. Black et al. [12] studied the characteristics of good reviewers conducting a survey on editors and first authors receiving reviewers’ reports. The only significant factor associated with higher-quality ratings by both editors and authors was reviewers trained in epidemiology or statistics. Some authors [13] have promoted the use of a fictitious manuscript in the assessment of reviewers’ quality. In other journals, the quality of reviews is subjectively scored by the editors. Callaham et al. [14] studied the reliability of editors’ subjective quality ratings. They compared the editors’ ratings with reviewers’ performance on a fictitious paper. Highly rated reviewers reported twice as many flaws as poorly rated reviewers. Thus, it seems that subjectively scoring the reviewers and assigning papers to the best ones could improve the quality of the accepted manuscripts. How this...
attitude from the editors could represent an overload of manuscripts for the best reviewers has to be investigated.

In conclusion, the agreement between the thoracic reviewers for EJCTS is reasonably good, facilitating an accurate selection of manuscripts to be published. We are indebted to our reviewer colleagues for their help in selecting the most useful and scientifically sound papers to be published in this journal.

REFERENCES


Table 1: Distribution of subjects by reviewer and category

<table>
<thead>
<tr>
<th>Rater B</th>
<th>Accept</th>
<th>Revise</th>
<th>Rework</th>
<th>Reject</th>
<th>Total</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept</td>
<td>7</td>
<td>12</td>
<td>1</td>
<td>6</td>
<td>26</td>
<td>12.9%</td>
</tr>
<tr>
<td>Revise</td>
<td>8</td>
<td>37</td>
<td>21</td>
<td>28</td>
<td>94</td>
<td>46.5%</td>
</tr>
<tr>
<td>Rework</td>
<td>2</td>
<td>19</td>
<td>12</td>
<td>6</td>
<td>39</td>
<td>19.3%</td>
</tr>
<tr>
<td>Reject</td>
<td>3</td>
<td>15</td>
<td>7</td>
<td>18</td>
<td>43</td>
<td>21.3%</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>83</td>
<td>41</td>
<td>58</td>
<td>202</td>
<td>100%</td>
</tr>
<tr>
<td>Percent</td>
<td>[9.9%]</td>
<td>[41.1%]</td>
<td>[20.3%]</td>
<td>[28.7%]</td>
<td>[100%]</td>
<td></td>
</tr>
</tbody>
</table>