Inaccurate Reporting of the Effects of Laetrile: Mistreatment of Ellison, Byar and Newell (1978) in Professional Papers

S Krashen

Citation

Abstract
There have been only two large-scale clinical studies on the impact of Laetrile (amygdalin) on cancer. In one of them, Ellison, Byar and Newell (1978) reviewed cases in which Laetrile was thought to be of help. Those who have cited this study have come to very different interpretations, however, some concluding that it showed that Laetrile was effective and some concluding that it was not. A review of Ellison et al. shows that even a conservative interpretation of their results suggests that condemnation of Laetrile may have been premature.

INTRODUCTION
The controversy over Laetrile (amygdalin), according to many, is settled a long time ago: Conventional wisdom is that Laetrile is quackery, and provides false hope to cancer patients who would be better off with traditional treatments. Maybe not. A review of the clinical evidence on Laetrile reveals a surprising number of individual case studies showing that Laetrile was effective, to at least some extent. Also surprising is the fact that there have been only two large-scale studies of the effect of Laetrile on cancer patients. One of them was done with patients “for which no standard treatment was known to be curative or to extend life expectancy” (Moertel et al. p. 201). The other was the National Cancer Institute “Special report on Laetrile” by Ellison, Byar, and Newell.

Ellison et al’s study only included cases “thought to have shown objective benefit from Laetrile” (p. 549) and the cases were reviewed by 12 expert oncologists.

Table one shows how Ellison et al. has been treated in reviews. The first five entries are from refereed journals, the sixth (NCI) is a website, the seventh is a report from the US Congress, and the eighth and ninth are books.

How Many Cases?
Six of the sources in Table 1 state that 66 to 68 cases were studied. Inspection of Table 1 in Ellison et al. reveals that there were indeed 68 cases presented to the panel, but 11 were eliminated because of “insufficient data” and 35 were “non-evaluable” (21 had “no followable disease,” 12 were...
undergoing other therapy at the same time, one did not have cancer, and another had more than one of the above reasons).
The actual number of cases considered by the panel was thus 22, not 66, 67 or 68, in agreement with Milazzo et al., Moss, Griffin, and part of the US Congress study.

The US Congress study mentions the Ellison et al. study on pages 28-29, where it is described as containing 67 cases “with sufficient information to be evaluated independently,” with six positive responses, as noted in table 1. In contrast, on page 105 of the same report, all the details of Ellison et al.’s table 1 are accurately presented, but the false impression given by the description on pages 28-29 is not corrected. The OTA report is about 300 pages and was the work of many contributors, which may explain the discrepancy.

**HOW MANY CASES WITH POSITIVE RESULTS?**

All of the authors in Table 1, except Moss and Griffin, concluded that only six cases in the Ellison et al. review showed positive results (two complete responses, four partial responses). Only Moss and Griffin noted that nine cases were classified as “stable disease,” that is, no increase or decrease in tumor size. Stabilization might have been due to the positive effect of Laetrile: If so, the number of positive cases would be 15, not six.

A conservative approach is to not include the nine cases of stable disease at all, given the fact that the Ellison et al. study included cases that had had a minimum of 30 days treatment with Laetrile, and given the considerable variability in rate of tumor growth. Eliminating the nine cases of stable disease from the analysis, we are left with six positive cases and seven cases of progressive disease.

**THE MYSTERIOUS THREE**

Only Moss and Griffin note that Ellison et al. describe other cases: “Three additional patients were judged to show increased disease-free interval, although Laetrile was used when there was no definite followable cancer” (p. 551). Ellison et al. then go on to say “The diagnosis in these three cases were Stage III testicular embryonal-cell carcinoma, State III ovarian adenocarcinoma, and a malignant tumor in an auxiliary lump node, probably metastatic” (p. 551). It is not clear whether these three cases were part of the 22. If we add them to the total, we obtain 9 positive cases out of 25. According to the conservative analysis, without the nine stable cases, we obtain 9 positive cases out of 16.

**CONCLUSION**

The most conservative approach is to eliminate the nine cases of stable disease and the three mystery cases, leaving six cases of positive response and seven of progressive disease. Even this conservative approach, however, is much more positive than the impression given by the reviewers in table one, with the exception of Moss and Griffin.

Moss and Griffin’s interpretations are the most lenient, both categorizing the nine stable cases and three mystery cases as success stories.

Of course, it is important to note that the cases reviewed by Ellison et al. were considered to be cases in which Laetrile was thought to have worked: Their review is therefore not a controlled study. But what we can conclude is that the Ellison et al. study has not been properly reported: Several reviewers erred in reporting the number of cases judged by the panel and all except Moss and Griffin ignored the stable cases and the three extra cases.

This is not a small point. Ellison et al. has been widely cited as counter-evidence to the claim that Laetrile is effective, with most reviewers claiming that Laetrile had a positive effect in only about 10% of the cases in which success was claimed, and they interpret this result as showing that Laetrile doesn’t work. Moss and Griffin, however, conclude that Laetrile was effective in over 80% of the cases, and my conservative interpretation is that there was a positive response in about half the cases. Condemnation of Laetrile, at least on the basis of the results of Ellison et al., may have been premature.

Despite these results, Ellison et al. betray some bias against Laetrile. They go to great lengths to speculate as to why a positive result occurred in six cases, suggesting that spontaneous regression might have taken place, and that concomitant therapy and psychological support may have played a role. They conclude that although they “cannot dismiss the possibility that the six patients … responded to Laetrile … the design of this study in no way allows us to draw this conclusion” (p. 552). Of great interest is the fact that they make no attempt to discuss why Laetrile might not have worked in all cases, even though there are obvious possibilities: Laetrile treatment could have been as short as 30 days and no details are provided on the kind or amount of Laetrile used or how it was administered. Also, like other treatments, effectiveness of Laetrile could depend on the kind of tumor. In addition, positive effects of Laetrile
Inaccurate Reporting of the Effects of Laetrile: Misreporting of Ellison, Byar and Newell (1978) in Professional Papers

reported in other studies, other than tumor size reduction, are not mentioned, such as relief of pain and improved appetite. Clearly, Laetrile is considered guilty until it is proven innocent.

**APPENDIX: A STATISTICAL ANALYSIS**

Herbert noted that Ellison et al. also reported on a “no treatment” group, and that a partial remission was recorded for one of the 24 cases. Ellison et al. did not do a statistical comparison of these 24 with the Laetrile-treated group because neither was randomly chosen (p. 550). Herbert, however, claimed that “the ‘partial remissions’ in the Laetrile group (four of 68) were not significantly greater than in the “no treatment” group (one of 24)” (p. 1135).

Ellison et al. were correct in not applying statistical tests because of the selection process used, but even if we allow testing, the results are not as Herbert claimed they are.

Herbert provided no details of the statistical test he used. Table 2 presents the results of a series of Fisher Exact Tests, designed to be used in 2 by 2 comparisons when sample sizes are small.

Laetrile 1 represents a comparison of the “no treatment” group with Laetrile-treated cases, assuming that the number of Laetrile-treated cases is in fact 68. As Herbert concluded, the difference between no treatment and Laetrile is not statistically significant. But one cannot claim a total of 68 cases, as argued here, as 46 cases could not be analyzed.

If we assume there were 22 cases (Laetrile 2), which counts “stable disease” cases as negative, the difference is on the borderline of statistical significance (p = .10, two tails).

If we assume 13 cases (Laetrile 3), the difference is easily statistically significant, that is, there were significantly more positive Laetrile-treated patients than positive no-treatment patients. Note also that the 13 case interpretation is conservative, as it does not include the three “mystery cases” and assumes that the “stable disease” cases provide no data one way or the other because of the possible short time duration. Recall that Moss and Griffin classified “stable disease” cases as positive, and included the three mystery cases.

### Table 2: Fisher exact tests applied to no treatment and laetrile-treated groups in ellison, byar, and newell

<table>
<thead>
<tr>
<th></th>
<th>positive</th>
<th>negative</th>
<th>p - one tail</th>
<th>p - two tails</th>
</tr>
</thead>
<tbody>
<tr>
<td>no treatment</td>
<td>1</td>
<td>23</td>
<td>0.41</td>
<td>0.67</td>
</tr>
<tr>
<td>Laetrile - 1</td>
<td>6</td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laetrile - 2</td>
<td>6</td>
<td>22</td>
<td>0.076</td>
<td>0.1</td>
</tr>
<tr>
<td>Laetrile - 3</td>
<td>6</td>
<td>7</td>
<td>0.004</td>
<td>0.004</td>
</tr>
</tbody>
</table>

A problem with this analysis is that the “no treatment” group might have included cases of “stable disease” as well. If so, the above analysis underestimates the percentage of positive cases in this group. Even if half were of this kind, and we eliminate these cases, the difference would still be significant for a one-tail test (p = .046) and quite close to significance for a two-tail test (p = .073).

**References**

Author Information

Stephen D. Krashen, Ph.D.
University of Southern California (Emeritus)