Interobserver variability in the pathologic interpretation of endomyocardial biopsy results

JEFFREY G. SHANES, M.D., JALAL GHALI, M.D., MARGARET E. BILLINGHAM, M.D., VICTOR J. FERRANS, M.D., PH.D., JOHN J. FENOGLIO, M.D., WILLIAM D. EDWARDS, M.D., CHENG C. TSAI, M.D., JEFFREY E. SAFFITZ, M.D., PH.D., JEFFREY ISNER, M.D., SYLVIA FURNER, M.P.H., AND RAMIAH SUBRAMANIAN, M.D.

ABSTRACT  Controversy exists over the role of endomyocardial biopsy in evaluating patients with dilated cardiomyopathy, particularly in detecting myocarditis and in assessing prognosis. Interobserver variability, if high, could explain conflicting reports. To assess this possibility, we submitted biopsy specimens from 16 patients with dilated cardiomyopathy to seven cardiac pathologists. The same slides were independently reviewed by each and assessed for fibrosis, hypertrophy, nuclear changes on a 0 to 3+ scale, mean lymphocyte count per high-power field, and myocarditis. The prevalence of significant fibrosis ranged from 25% to 69%, hypertrophy from 19% to 88%, nuclear changes from 31% to 94%, and abnormal lymphocyte count from 0 to 38%. One or more pathologists diagnosed definite or possible myocarditis in 11 of the 16 patients. Of these 11 patients, three pathologists agreed about three and two pathologists agreed about five. Myocarditis was diagnosed by a single pathologist in three cases. We conclude that interobserver variability is high in interpreting biopsy specimens from patients with dilated cardiomyopathy and that quantitative and standardized methods are needed to increase diagnostic consistency.


ENDOMYOCARDIAL BIOPSY is a safe\(^1\) and widely utilized technique for evaluating patients with a variety of cardiac diseases. Use of this technique in patients who have undergone cardiac transplantation\(^2\) and in those receiving doxorubicin\(^3\) (Adriamycin) therapy is of proven clinical benefit and represents an important research tool.\(^4,5\)

However, use of this technique in patients with dilated cardiomyopathy remains controversial. Conflicting reports have been published describing the incidence of myocarditis in this disease entity\(^6-8\) as well as the therapeutic benefit of immunosuppressive therapy for myocarditis.\(^9,10\) Findings of fibrosis, hypertrophy, and nuclear abnormalities on biopsy have also been correlated with disease severity and assessment of prognosis with variable results.\(^11-15\) Nonuniform methods of pathologic interpretation might be an important reason for such conflicting data. The purpose of our study was to determine the interobserver variability that exists in the assessment of fibrosis, hypertrophy, nuclear changes, lymphocyte count, and diagnosed myocarditis.

Methods

Patients. Between August 1, 1982, to August 1, 1984, 64 endomyocardial biopsies were performed at the University of Illinois Hospital on patients with cardiomyopathy or idiopathic ventricular tachycardia. Pathologic specimens were selected from 16 of these patients for this study because their clinical presentation at the time of endomyocardial biopsy most strongly suggested myocarditis compared with the entire group in general. The patients' clinical characteristics are summarized in table 1. Eleven men and five women (mean age 43.8 years) were studied. The mean duration of symptoms before biopsy was 527 days; however, nine of the patients had symptoms less than 180 days. The mean ejection fraction was 22.3%. Four patients experienced a viral prodrome, four had a history of heavy alcohol use, and three had a positive gallium scan, a marker of inflammation.\(^16\) Three patients had coronary disease, but it was not severe enough to explain the degree of left ventricular dysfunction.

Procedure. All endomyocardial biopsies were performed under fluoroscopic guidance by percutaneous puncture of the right internal jugular vein; three to four specimens were obtained
from the right ventricular septum with the Stanford-Caves biop-
tome. When performed at the time of the complete right and left
heart catheterization and angiographic study, biopsies were
done at the beginning of the procedure before the administration
of contrast material. The specimens were obtained from various
sites along the right ventricular septum and immediately placed
into 3% paraformaldehyde or 10% buffered formalin. No com-
plications occurred in any patient.

**Histologic methods.** Tissue was embedded in paraffin and
cut into 5 μm sections. Each biopsy specimen was stained with
hematoxylin and eosin and Masson’s trichrome stain.

**Histologic evaluation.** Seven cardiac pathologists,* highly
experienced in the interpretation of endomyocardial biopsy
specimens, agreed to participate in this study. The same slides
were sent to each pathologist for review and the results were
recorded on a standard form.

Each slide was assessed for endocardial, interstitial, and peri-
vascular fibrosis, cellular hypertrophy, and specific nuclear ab-
normalities, including enlargement, hyperchromaticity, irregu-
larity, and the presence of nucleoli. A scale of 0 to 3 + was used
with 0 considered normal, 1 + mildly abnormal, 2 + moderately
abnormal, and 3 + severely abnormal. Scores were given for
each individual abnormality and total scores were obtained by
adding these scores together in each general category. For ex-
ample, a patient with 1 + interstitial fibrosis, 2 + perivascular
fibrosis, and 2 + endocardial fibrosis would have a total fibrosis
score of 5 + .

The pathologists were also asked to determine the mean lymp-
phocyte count in 10 high-power fields and to give the minimum
and maximum number of lymphocytes found for the 10 fields
examined. Finally, they were asked to comment on whether or
not myocarditis was found.

**Statistical methods.** Interobserver variability was assessed
by comparisons, between patients, of the coefficients of vari-
ation for each variable score. To further assess the influence
of the observer on each of the variable scores, analyses of variance
were conducted. The effect of observer on each of the variables
was determined.

*Margaret E. Billingham, M.D., Stanford University; Victor J. Fer-
rans, M.D., Ph.D., National Institutes of Health; John J. Fenoglio,
M.D., Columbia University; William D. Edwards, M.D., Mayo Clinic;
Cheng C. Tsai, M.D., St. Louis University; Jeffrey E. Saffitz, M.D.,
Ph.D., Washington University; Jeffrey Isner, M.D., Tufts-New En-
land Medical Center.

---

**Results**

Table 2 shows the median coefficient of variation for nuclear changes, fibrosis, hypertrophy, and lymphocyte count as graded by the seven pathologists. Good agreement among the pathologists would be indicated by a low median coefficient of variation. However, the lowest median coefficient of variation for the variables tested was 52.6 and the highest was 183.6. An analysis of variance indicated statistically significant interobserver variability for all variables except nuclear irregularity and interstitial fibrosis.

The effect of interobserver variability became less pronounced when we divided the same variables into significant vs nonsignificant abnormalities. Significant abnormalities were defined as those in which the individual score was ≥ 2 +, or the total score ≥ 4 + for fibrosis and ≥ 6 + for nuclear changes. A lymphocyte count was considered significant when the mean count was five or more lymphocytes per high-power field or the highest lymphocyte count in any field greater than 10 cells. Table 3 shows the prevalence per 100 patients of significant abnormalities for each vari-
able in the total patient population. It can be seen that the prevalence varied considerably by pathologist. The range for fibrosis was 25% to 69%, for hypertrophy 19% to 88%, for nuclear abnormalities 31% to 94%, and for a high lymphocyte count 0 to 38%. Only nuclear enlargement, irregularity, and cellular hypertrophy showed a statistically significant observer effect by an analysis of variance (p < .05).

Table 4 shows the findings in individual patients who were diagnosed as having myocarditis by each pathologist. Considerable variation in terminology
TABLE 3
Prevalence of significant abnormalities by pathologist: significant vs nonsignificant abnormalities

<table>
<thead>
<tr>
<th><strong>Variable</strong></th>
<th><strong>Pathologist</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Fibrosis</td>
<td></td>
</tr>
<tr>
<td>Any single score ≥2+</td>
<td>25</td>
</tr>
<tr>
<td>Total score ≥4+</td>
<td>19</td>
</tr>
<tr>
<td>Either/or</td>
<td>25</td>
</tr>
<tr>
<td>Hypertrophy</td>
<td>81</td>
</tr>
<tr>
<td>Nuclear abnormalities</td>
<td></td>
</tr>
<tr>
<td>Any single score ≥2+</td>
<td>88</td>
</tr>
<tr>
<td>Total score ≥6+</td>
<td>69</td>
</tr>
<tr>
<td>Either/or</td>
<td>88</td>
</tr>
<tr>
<td>Lymphocyte count ≥5 mean/HPF or ≥10 cells/any focus</td>
<td>0</td>
</tr>
</tbody>
</table>

HPF = high-power field.

*Data expressed as percentage of significant findings.

was used by the individual pathologists and included acute myocarditis, borderline myocarditis, chronic myocarditis, focal myocarditis, myocarditis, and resolving myocarditis. One pathologist believed that patient 3 had inadequate tissue and three pathologists thought that staining was inadequate in patient 13. Figure 1 is a hematoxylin and eosin–stained slide from patient 7, in whom a diagnosis of definite myocarditis was made by three pathologists; all three agreed that a high lymphocyte count was present.

Table 5 summarizes the percentage of patients with definite myocarditis (diagnosed myocarditis, acute myocarditis, and chronic myocarditis) compared with those patients with possible myocarditis (focal myocarditis, borderline myocarditis, and resolving myocarditis) and those with a significantly high lymphocyte count. One or more pathologists diagnosed definite or possible myocarditis in 11 or the 16 patients. Of these 11 patients, three pathologists agreed about three and two pathologists agreed about five. Myocarditis was diagnosed by a single pathologist in three cases.

The data were also analyzed for any correlation between the pathologists' scores and the patients' baseline clinical variables. We found that the degree of both perivascular fibrosis and total fibrosis score correlated inversely with the left ventricular ejection fraction, but no other correlation was observed.

Discussion

The clinical role of endomyocardial biopsy for evaluation of patients with dilated cardiomyopathy of both chronic and acute onset and in other clinical situations such as idiopathic ventricular tachycardia is uncertain. Some studies have demonstrated histologically proven myocarditis responsive to immunosuppressive therapy in some patients. However, the ability of this technique to detect myocarditis is variable. For example, some centers have reported a high incidence of myocarditis (63%) in patients with idiopathic dilated cardiomyopathy while other centers have found the incidence of myocarditis in dilated cardiomyopathy to be very low (2%). Even when myocarditis is found, the response to treatment with antiinflammatory drugs has also been variable. Although some centers have reported improvements in left ventricular function, others have not. Although differences in patient population or immunosuppressive regimens could potentially explain these conflicting results, nonuniform criteria as to what constitutes myocarditis and inconsistencies in the identification of histologic features of myocarditis could also explain these varying results. To test this hypothesis, we asked a panel of cardiac pathologists to independently interpret the same slides in a blinded fashion without knowledge of the patients’ clinical characteristics. We also believed it would be important to establish the interobserver variability that occurs when diagnosing other commonly reported findings of patients with dilated cardiomyopathy such as fibrosis, nuclear abnormalities, and hypertrophy. Only in re-
cent studies have investigators attempted to quantitatively correlate these abnormalities with prognosis and hemodynamic variables; this was not attempted in older studies. In addition, most centers performing endomyocardial biopsies do not have the capabilities for quantitative techniques and often report their findings in a qualitative or so-called semiquantitative fashion, as we asked our panel of pathologists to do.

Our results showed considerable interobserver variability for all data measured. This was particularly true for the diagnosis of myocarditis. The wide variability in terminology by the pathologists made our data somewhat difficult to quantify. However, by grouping these terms into three basic categories of definite, questionable, and no myocarditis, strong observer effects were demonstrated. The same observer effect was noted when quantifying lymphocyte counts or grading the severity of hypertrophy, nuclear abnormalities, and fibrosis on a four-point scale (0 to 3+). This effect, however, became less important statistically when we reclassified patients as having either mild or no abnormalities compared with moderate or severe abnormalities. However, even with this type of analysis, a strong trend toward observer bias was still seen.

Finally, we attempted to correlate scores for fibrosis, myocarditis, and other histologic findings with the baseline clinical characteristics of the patients such as ejection fraction and duration of disease. Only an inverse relationship between the ejection fraction and severity of fibrosis was noted. Importantly, no correlation was noted for patients with positive gallium scans and findings of myocarditis on cardiac biopsy. We believe that this study has important implications for delineating the clinical role of endomyocardial biopsy in the evaluation of patients with dilated cardiomyopathy. Clearly, one of the most important questions to be answered when this procedure is performed is whether or not myocarditis is present. Undoubtedly, reports of unexpected myocarditis in a large percentage of patients with dilated cardiomyopathy, idiopathic ventricular tachycardia, postpartum cardiomyopathy, and familial cardiomyopathy have

<table>
<thead>
<tr>
<th>Pathologist</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definite myocarditis</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>19</td>
<td>38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Possible myocarditis</td>
<td>6</td>
<td>19</td>
<td>0</td>
<td>13</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>No myocarditis</td>
<td>94</td>
<td>62</td>
<td>100</td>
<td>68</td>
<td>56</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>High lymphocyte count</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>13</td>
<td>38</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

*Data expressed as percentage of significant findings.*
encouraged the increased use of this procedure in hopes of finding a potentially treatable cause of heart failure. However, our study indicates that a diagnosis of myocarditis by one observer may not necessarily be in agreement with that of another observer of the same patient. This must be taken into account when interpreting published data, planning treatment modalities for patients with discovered myocarditis on biopsy, and planning clinical trials to assess the effects of immunosuppressive therapy on patients with diagnosed myocarditis.

Recently, a group of cardiac pathologists in Dallas established standard criteria for diagnosis of myocarditis.20 Hopefully, these criteria will be uniformly accepted and will help reduce confusion in the terminology of myocarditis. These criteria may also reduce interobserver variability, but only a follow-up study such as ours using these new criteria would establish this. Certainly, studies using objective criteria such as histochemical techniques to objectively recognize and quantitate inflammatory cells in biopsy material should reduce observer variability. Objective and quantitative data are already available and are being more widely used to measure fibrosis and hypertrophy and have increased our understanding of what these abnormalities mean in terms of cardiac function and prognosis.

Special thanks to Carol Davis and Frances Jefferson for their expert preparation of this manuscript and to Sharon Oates and Barbara Jane Dierenfeldt for their editorial assistance. We also wish to thank Dr. Jose R. Manaligod and Dr. Abel L. Robertson, Jr., from the Department of Pathology for making the slides available to us and coordinating shipping.

References

6. Parrillo JE, Aretz HT, Palacios I, Fallon JT, Block PC: The results of transvenous endomyocardial biopsy can frequently be used to diagnose myocardial diseases in patients with idiopathic heart failure: endomyocardial biopsies in 100 consecutive patients revealed a substantial incidence of myocarditis. Circulation 69: 93, 1984
Interobserver variability in the pathologic interpretation of endomyocardial biopsy results.

J G Shanes, J Ghali, M E Billingham, V J Ferrans, J J Fenoglio, W D Edwards, C C Tsai, J E Saffitz, J Isner and S Furner

Circulation. 1987;75:401-405
doi: 10.1161/01.CIR.75.2.401

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1987 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/75/2/401

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/